Mechanism of outer membrane protein assembly by the Bam complex

Dissertation Advisor: Professor Daniel Kahne David Tomasek

Mechanism of outer membrane protein assembly by the Bam complex

Abstract

β-barrel integral membrane proteins perform important roles in the outer membranes of Gram-negative bacteria, mitochondria, and chloroplasts. In Gram-negative bacteria, the β-barrel assembly machine (Bam) complex accelerates the folding and membrane integration of these proteins. Structural and biochemical studies of the Bam complex have led to hypotheses of how BamA, the central component of the complex that is itself a β-barrel protein, may interact with substrates to promote their assembly. However, an understanding of how this molecular machine accelerates folding in the absence of an energy source remains unclear. Characterization of folding intermediates trapped on the Bam complex is needed to develop a physical picture for how catalysis occurs.

This thesis describes biochemical and structural experiments aimed at revealing how the Bam complex promotes the assembly of outer membrane proteins. In Chapter 2, we use mutant substrates that become trapped on the machine during folding to show that folding is catalyzed within the interior of the BamA β-barrel. In Chapter 3, we generate and characterize a larger series of Bam complex substrates that are trapped at different stages of folding. These experiments demonstrate that folding occurs in a stepwise fashion from the C-terminus to the N-terminus of the substrate, with early stages occurring outside the membrane and late stages occurring within the membrane. In Chapter 4, the ability to trap substrate folding intermediates is exploited to obtain a structure of a substrate-bound Bam complex by cryo-electron microscopy. This structure reveals a network of interactions between BamA and the substrate, suggesting that BamA templates substrate folding and that intramolecular interactions within the substrate trigger its release once folding has finished.

These results provide insight into how the Bam complex assembles its substrates in an energy-independent manner. The findings here may guide further structural studies with substrate folding intermediates and may facilitate development of antibiotics that target Gram-negative pathogens by preventing proper function of the Bam complex.

Table of Contents

List of Figures

List of Tables

Acknowledgements

I am grateful for all the support I have received over the past six years. First, I would like to acknowledge my advisor, Professor Dan Kahne. When I first met with Dan, he proposed several projects that I found intriguing. I naïvely believed that these projects would be straightforward and decided to join the lab. Since then, Dan has provided an environment to me to develop independently as a scientist. I am appreciative of his encouragement, especially at times when failure seemed inevitable, and for helpful feedback when I needed it. Dan's ability to come up with creative explanations of data and think very far outside of the box has helped me begin to do the same. I am thankful that he has always been willing to listen to my thoughts, even if he disagrees with them at times.

Other faculty members at Harvard have also provided guidance to me during my years here. I would like to thank Professors Vlad Denic, Rachelle Gaudet, and Rich Losick for serving on my thesis committee. I am fortunate to have been able to receive feedback from this group of scientists, which sometimes helped determine the direction the project took. I appreciated having people outside the lab to convince me that maybe Dan was right at times, or to convince Dan that maybe I was right at other times.

My time at Harvard has been shaped by the other members of the Kahne lab, including other members of the Bam subgroup. The first person I met in the lab was James Lee, a fellow graduate student in my training program. I appreciated having had him nearby when I began working in the lab and was too intimidated to talk to anyone else. His experimental support has been a tremendous help to me, and his ideas have inspired my own work. He has been a great friend, and I appreciate all the conversations we have had both related and unrelated to science.

Within the Bam subgroup, another one of my early mentors was Joe Wzorek, a former postdoc in the lab. His ideas have led to a lot of the studies that make up this dissertation. I respect his ability to maintain a positive attitude even in the face of horrendous Western blots, a quality I

have failed to inherit. Although I have not yet completed the “to-do list” of experiments that he wrote up for me in 2017, I hope I have done enough to earn his respect.

I must also thank Mary May and Thiago Santos, other members of the Bam subgroup. They have been excellent collaborators and, together, we have finished up several projects for publication. More importantly, they are also people who are enjoyable to be around. I would also like to acknowledge Ina Meuskens, a summer student, for her contributions to the subgroup. Working with past and present members of the Bam subgroup has been a great experience, and I’ll be fortunate if I have such a great group of people to work with in the future.

Other members of the lab have also shaped my time at Harvard. Michael Mandler has been a constant source of amusement for me and his spontaneity is something I could learn from. I appreciate him introducing himself to me even when I initially kept my distance. We have also collaborated on several projects, although most of these did not amount to anything. Becca Taylor has also been a great friend during my time in the lab. I am grateful for the encouragement she has provided me, for brightening my day when I needed it (often through photos of cats), and for motivating me to do things outside the lab. I thank Karan Pahil for getting lunch with me many times and for our lively discussions. I appreciate his efforts in cheering me up when necessary and for also sharing photos of cats with me. I would also like to thank Fred Rubino and Katie Schaefer for being the main people I talked to on my side of the lab. I enjoyed the many conversations we had and for being given the opportunity to raise a hibiscus.

I appreciate everyone else who I have met in the Kahne lab and in the MCO Graduate Program for making Harvard an enjoyable place to work over these past six years. The memories I have made with all of you have shaped my time here, and I will forever be grateful for having had the opportunity to grow with you all. I wish you all nothing but happiness in your lives.

My collaborators outside the Kahne lab have also played an important role in shaping the project presented in this dissertation. In particular, Zongli Li and Shaun Rawson at the Harvard Cryo-EM Center for Structural Biology at Harvard Medical School performed cryo-EM

experiments described in this dissertation. I have learned a lot from them, and I appreciate their perseverance when our data looked terrible. I also appreciate the guidance provided by Professor Steve Harrison during the course of the project. This project would have gotten nowhere without their dedication.

I am also thankful for those who have kept the Kahne lab running during my time here: Helen Corriero, Mike Quinn, Rebecca Stillo, and Hannah Kelley. They have been very helpful in making the lab run properly (ensuring that we have money) and with fitting me into Dan's hectic schedule. Without their constant dedication and patience with Dan, the lab would rapidly fall into chaos.

I would like to thank my advisor from my time at Columbia University, Professor Filippo Mancia. Filippo gave me a chance to join his lab in 2012 to learn about protein purification and biochemistry. He took a chance by hiring someone who knew nothing about these areas, and I appreciate the opportunity he provided me. The skills I learned in his lab were critical for tackling the project presented in this dissertation, even if it still took me a long time to finish.

My family has also been supportive of me over these years. Even though they are unfamiliar with the process of obtaining a Ph.D. and scientific research in general, they rarely questioned why it was taking so long to finish the program, and instead allowed me to make my own decisions. They have given a lot for me to be where I am, and I will be always indebted to them.

Chapter 1: Introduction

Some sections of this chapter are adapted/reproduced from:
Lee J, Tomasek D*, Santos TMA*, May MD*, Meuskens I, Kahne D. Formation of a β-barrel membrane protein is catalyzed by the interior surface of the assembly machine protein BamA. *eLife* **8**, e49787 (2019).

1.1. The bacterial cell envelope

Bacteria are surrounded by a cell envelope that provides protection from the environment. Based on the structure of their cell envelopes, bacteria can be divided into two broad classes: Gram-positive and Gram-negative (**Figure 1.1**). Gram-positive bacteria have a thick peptidoglycan cell wall and a cytoplasmic membrane (**Figure 1.1, left**). In contrast, Gram-negative bacteria contain a thinner peptidoglycan cell wall and a second membrane called the outer membrane (**Figure 1.1, right**). The aqueous space separating the two membranes is called the periplasm. The unique outer membrane acts as a barrier to protect the cells from foreign agents including antibiotics, but also allows passage of molecules into and out of the cell that have roles in survival or virulence. This membrane has an asymmetric lipid composition with the glycolipid known as lipopolysaccharide in its outer leaflet and phospholipids in the inner leaflet. It also contains β-barrel membrane proteins and lipoproteins. The presence of this outer membrane results in the intrinsic resistance Gram-negative bacteria to many antibiotics that target Gram-positive bacteria[1–3].

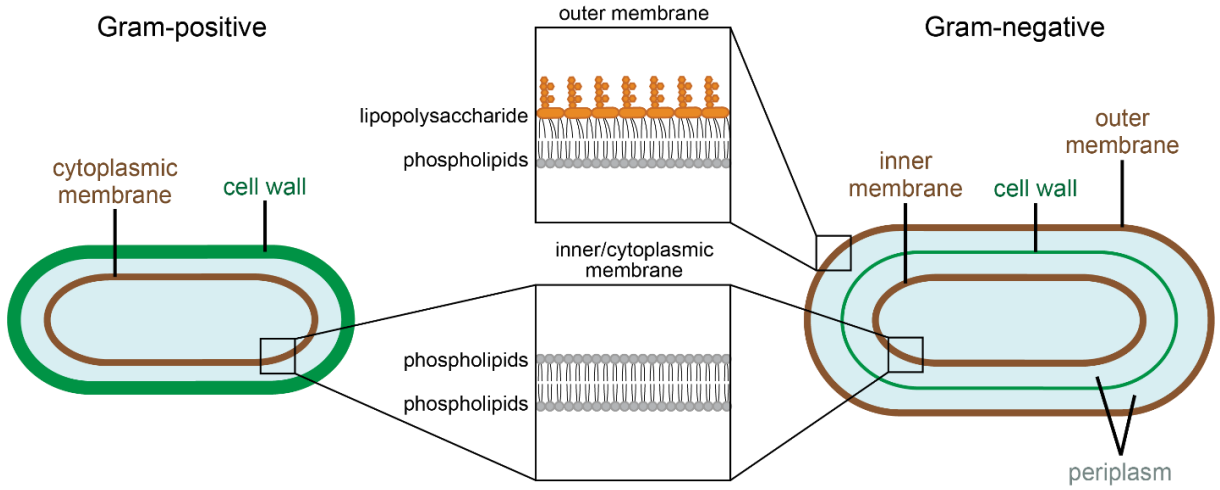

Figure 1.1. Structure of the bacterial cell envelope
Gram-positive bacteria (left) have a thick cell wall and a cytoplasmic membrane. In contrast, Gram-negative bacteria (right) have a thinner cell wall but contain an outer membrane in addition to the cytoplasmic (inner) membrane. The lipid composition of each membrane is different.

This thesis presents biochemical and structural experiments aimed at understanding how β-barrel integral membrane proteins are folded and inserted into the outer membrane of Gram-

negative bacteria. This chapter serves as an introduction to this area by outlining the biogenesis of the components of the outer membrane, including lipids and proteins.

1.2. Outer membrane lipid biogenesis

The outer membrane of Gram-negative bacteria is an unusual lipid bilayer due to its asymmetry, with phospholipids in the inner leaflet and lipopolysaccharide (LPS) in the outer leaflet. These lipids are generated at the cytoplasmic face of the inner membrane. Therefore, cells must not only have mechanisms to synthesize these lipids but must also have ways to transport them across the inner membrane and then deliver them to the outer membrane. Also, the balance of LPS and phospholipids between the two leaflets of the outer membrane must be regulated. This section discusses the basic biology of the composition and synthesis of the lipids of the outer membrane.

1.2.1. Lipopolysaccharide is found in the outer leaflet of the outer membrane

LPS is an amphipathic molecule containing a lipid A group, inner and outer core oligosaccharides, and the O antigen[4]. Within the sugars of the core, phosphate groups are present, which can interact with divalent cations such as magnesium, allowing efficient packing of LPS in the outer leaflet of the outer membrane. The sugars create a polar mesh around the cell, preventing entry of hydrophobic molecules, while the lipid A group prevents entry of polar molecules. Thus, the unique structure of LPS makes the membrane impermeable to many drugs.

1.2.1.1. Biosynthesis and translocation across the inner membrane

Given the complexity and location of LPS in the cell, many proteins are involved in its synthesis and transport to the outer membrane. LPS synthesis occurs in the cytoplasm and the process requires more than 100 genes[4–6]. Specifically, the lipid A and core oligosaccharide complex is produced in the inner leaflet of the inner membrane. The lipid A portion is synthesized

by the Lpx and Waa protein families, and glycosyltransferases then add sugar residues to form the core oligosaccharide portion. After their synthesis, the lipid A and core complex is then flipped by the ATP-binding cassette (ABC) transporter MsbA[7–10]. The O antigen is independently made in the cytoplasm and is transported to the periplasm, where it is joined to the rest of LPS by the O antigen ligase WaaL[11].

1.2.1.2. Transport of lipopolysaccharide to the outer membrane

After being flipped across the inner membrane by MsbA, LPS must then be transported to the outer membrane. This transport process must occur quickly, as millions of molecules need to be delivered to the cell surface per generation[12]. Within the past fifteen years, seven lipopolysaccharide transport (Lpt) proteins have been identified that perform this process (**Figure 1.2b**). In the first step, the $LptB_2FGC$ complex extracts LPS from the outer leaflet of the inner membrane and delivers it to the LptA component[13–15]. LptA connects the inner and outer membrane components by forming a membrane-to-membrane protein bridge that could consist of multiple copies of LptA[13,16–19]. Energy from ATP hydrolysis the dimer of LptB molecules is harnessed to perform this transport of LPS across the periplasm[20,21]. After transport across the periplasm, a two-protein complex consisting of LptD and LptE inserts LPS into the outer leaflet of the outer membrane[22,23]. All seven proteins are necessary for transport and reducing the levels of any results in inner membrane accumulation of LPS[14,15]. LptC, LptA, and LptD contain homologous β-jellyroll domains that allow their interaction to form the bridge for LPS delivery from the inner membrane to the outer membrane[18,24–26]. Importantly, this bridge must have a mechanism for shielding the acyl tails of LPS during transit to prevent aggregation, likely by utilizing a hydrophobic pocket of LptA[25].

Interestingly, LptD forms one of the largest β-barrels in the outer membrane, containing 26 β-strands[26–28]. LptD contains a separate protein, LptE, within its interior that facilitates folding and may play a role in the function of LptD[23,29–31]. The plug-and-barrel architecture of the LptDE

complex suggests that its assembly may have unique requirements. A further complexity is that LptD contains four cysteines: initially, upon synthesis of the protein, LptD forms two disulfide bonds between cysteines proximal in sequence. After folding of LptD, the cysteines are rearranged to form nonconsecutive disulfide bonds between cysteines distant in the primary structure[32,33]. LptD must adopt this properly folded state with the mature disulfide arrangement in order to form the transenvelope bridge that transports LPS[17]. The assembly of the LptDE complex is therefore of critical importance to the cell and some of the studies described in this thesis examine this assembly process.

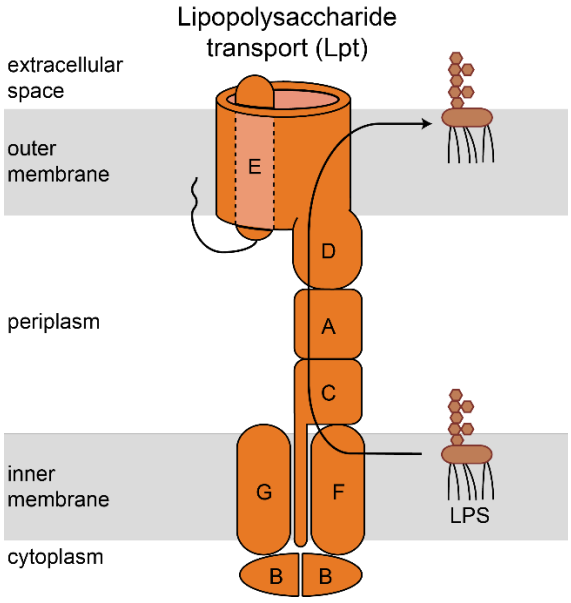

Figure 1.2. The lipopolysaccharide transport (Lpt) pathway delivers LPS to the outer membrane
The Lpt system consists of $LptB_2FGC$ in the inner membrane, LptDE in the outer membrane, and LptA, which connects the inner and outer membrane complexes. The Lpt pathway thus forms a membrane-to-membrane bridge for efficient delivery of lipopolysaccharide to the cell surface.

1.2.2. Phospholipids are found in the inner leaflet of the outer membrane

Phospholipids are essential for maintaining the integrity of the outer membrane. These lipids are most commonly glycerophospholipids, which are made up of a glycerol moiety, a phosphate group, variable head group, and two fatty acids. They must be transported to the outer membrane via anterograde transport, but mechanisms of retrograde transport also exist to permit the passage of lipids from the outer membrane to the inner membrane[34,35]. This implies that there

must exist molecular machines to facilitate this transport. Although the synthesis of phospholipids has been well characterized, defining the machines that perform the transport is a matter of ongoing work.

1.2.2.1. Biosynthesis and translocation across the inner membrane

Much work has been done to characterize the biosynthetic enzymes involved in phospholipid production[36,37]. In brief, a defining step in the synthesis of these lipids uses cytidine triphosphate and phosphatidic acid are to generate CDP-diacylglycerol (CDP-DAG)[37–39]. CDP-DAG is the common precursor for synthesizing different glycerophospholipids in bacteria, with phosphatidylethanolamine, phosphatidylglycerol, and cardiolipin being the most common. After phospholipids are synthesized in the inner membrane, they must be released from the membrane, transported across the periplasm, and then inserted into the outer membrane[40]. The mechanism of flipping is currently unknown, but presumably occurs via a transport protein.

1.2.2.2. Transport of phospholipids between the inner and outer membranes

Phospholipid transport has been shown to be bidirectional[34,35,41]. The mechanism of transport of phospholipids to the outer membrane (i.e., anterograde transport) remains unclear, but a system for transport would be expected to have a mechanism to shield acyl tails from water during the process. Many years ago, it was hypothesized by Bayer that lipid transport might occur between adhesion sites between the inner and outer membrane[42]; however, the presence of these “Bayer bridges” has been controversial and dismissed by some as artifacts[43]. It is thus assumed, as for LPS, that transport occurs via a molecular machine. No system has been conclusively demonstrated to promote anterograde transport of phospholipids to the outer membrane, the PbgA/YejM system may be involved in cardiolipin transport to the outer membrane, based on evidence that the complex is required for cardiolipin enrichment in the outer membrane of *Salmonella typhimurium*[44] or *Shigella flexneri*[45].

More work has been done to identify systems involved in retrograde transport of phospholipids. In particular, the Mla and Tol-Pal systems have been assigned a role in the process. The Mla system involves a complex of OmpC and MlaA at the outer membrane, the periplasmic protein MlaC, and the inner membrane ABC transporter MlaFEDB[46–48]. This system is believed to remove mislocalized phospholipids from the outer membrane and take them back to the inner membrane. Another pathway believed to play a role in phospholipid transport, the Tol-Pal system, consists of the TolQRA subcomplex in the inner membrane and the TolB-Pal subcomplex in the outer membrane[49,50]. Mutations within these proteins or their removal can cause outer membrane defects including the accumulation of excess phospholipids in the outer membrane[51]. However, it remains to be investigated whether this complex is the direct mediator of transport, or whether it can affect the function of a different system that directly interacts with phospholipids[40].

1.3. Outer membrane lipoprotein biogenesis

The biogenesis of outer membrane lipoproteins has been extensively studied[52,53]. As is true for any periplasmic or outer membrane protein, synthesis occurs in the cytoplasm, generating a protein with an N-terminal signal sequence[54]. This signal sequence allows post-translational targeting to the SecYEG translocon (described in more detail in section 1.4.2 below). After crossing the inner membrane, the pre-lipoprotein is modified by several enzymes. First, a diacylglycerol group is attached at the first cysteine after the signal sequence. The signal peptide is then cleaved at the cysteine, and, finally, the cysteine is then acylated again to form the mature lipoprotein[55–57]. This protein therefore contains a covalently-attached N-terminal lipid anchor, and this triacyl group anchors the otherwise soluble protein to the outer leaflet of the inner membrane.

After its synthesis, the nascent lipoprotein will be retained in the inner membrane if it contains an aspartate adjacent to the lipidated cysteine or transported across the periplasm to the outer membrane if it does not[58]. Transport to the outer membrane is performed by the Lol

(localization of lipoprotein) machine[52,53] (**Figure 1.3**). This system consists of five essential proteins, LolABCDE. In the transport process, the lipoprotein is first extracted from the inner membrane by the LolCDE ABC transporter system[59–61]. Energy from ATP hydrolysis allows transfer through LolC to the soluble chaperone LolA, which resides in the periplasm and interacts with the lipid tails to shield them from water and prevent aggregation[59,62,63]. LolA then delivers the lipoprotein to the outer membrane for transfer to the LolB lipoprotein via a "mouth-to-mouth" interaction between LolA and LolB[64,65]. Through an unclear mechanism, likely driven by the stability of the lipid tail in the membrane environment, LolB then releases the lipoprotein into the outer membrane[65–67], though how lipoproteins obtain different topologies (i.e., facing the periplasm or the extracellular space) is unknown[68].

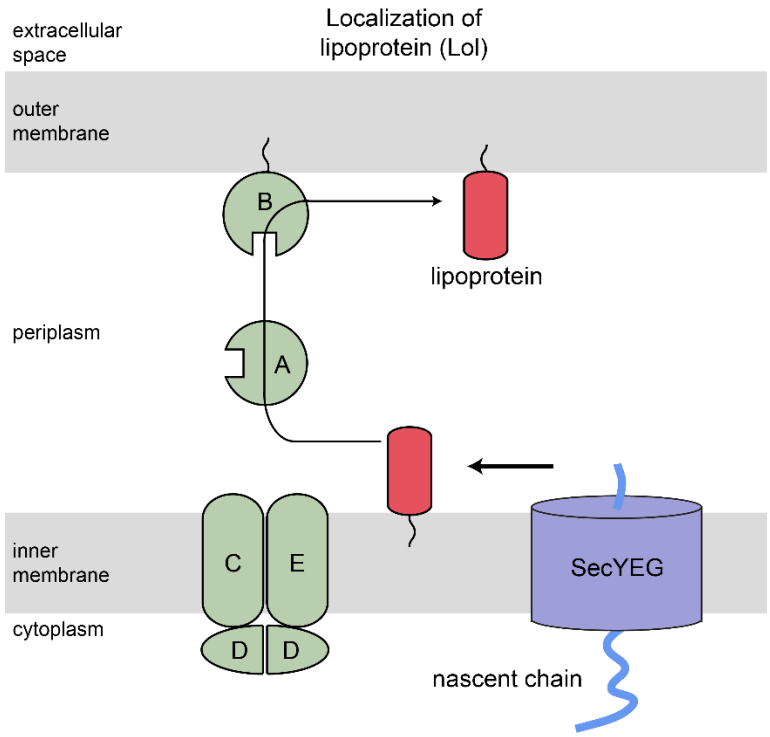

Figure 1.3. The Lol pathway transports lipoproteins to the outer membrane
The localization of lipoprotein (Lol) pathway extracts lipoproteins from the inner membrane and transports them to the outer membrane, where they become anchored via their N-terminal lipid tails.

1.4. β-barrel outer membrane protein biogenesis

In addition to containing different lipid compositions, the inner and outer membrane of Gram-negative bacteria each contain different types of integral membrane proteins. The inner membrane contains α-helical membrane proteins in which transmembrane-spanning helical bundles are packed together with hydrophobic side chains oriented outward in a way that permits hydrophobic interactions with membrane lipids. In contrast, the outer membrane contains β-barrel transmembrane proteins that perform a variety of functions important for cell survival. The work presented in this dissertation examines how these β-barrel proteins are folded and inserted into the outer membrane.

1.4.1. Structure and function of β-barrel outer membrane proteins

β-barrel transmembrane proteins, which comprise a variable number of antiparallel β-strands wrapped into a cylinder[69], are also found in the outer membranes of mitochondria and chloroplasts. These proteins are involved in a range of functions, performing roles as channels (porins), transporters, enzymes, receptors, or structural proteins. Their assembly is thus essential for maintaining the integrity of the cell envelope[70–73]. A β-stranded structure is thought to be stable in the membrane only as a completely folded β-barrel in which adjacent β-strands form hydrogen bonds, and the N- and C-terminal β-strands are also joined via hydrogen bonding to form a closed β-barrel. As in α-helical membrane proteins, the membrane-exposed exterior surface of β-barrels is hydrophobic, meaning that, along a given β-strand, side chains alternate between polar (facing inward towards the lumen of the barrel) and nonpolar (facing outward). Although they share a common architecture, β-barrel outer membrane proteins can range vastly in size from 8 to 36 strands from a single polypeptide (**Figure 1.4**). Despite their importance, the mechanism by which β-barrel transmembrane proteins are folded is not well understood.

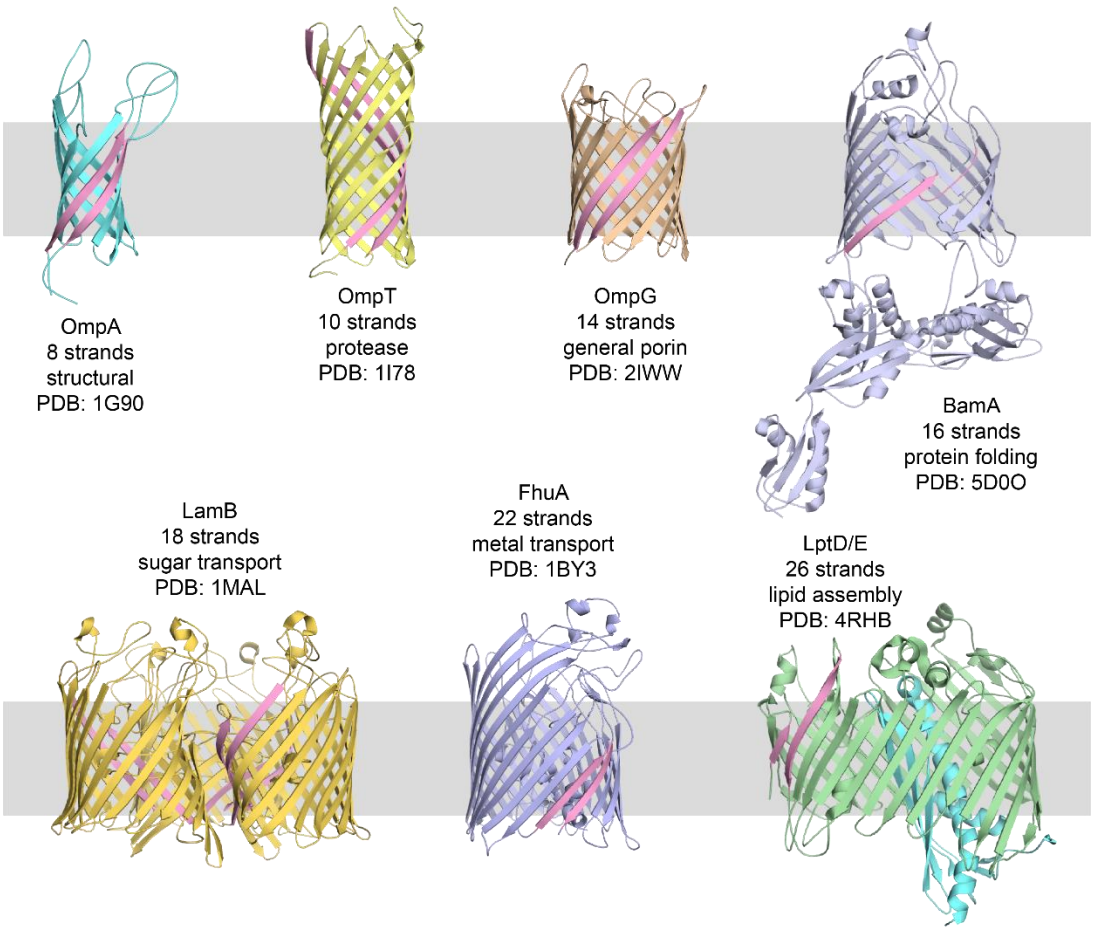

Figure 1.4. Outer membrane proteins are diverse in structure and function
β-barrel outer membrane proteins can vary in size from 8-36 strands. Some contain large soluble domains that reside in the periplasm or within the interior of the β-barrel, while others form oligomers. The first and last β-strands of each β-barrel are colored in salmon.

1.4.2. Biosynthesis and secretion

In Gram-negative bacteria, proteins destined for membrane integration or secretion are synthesized in the cytoplasm via ribosomes and are targeted to the inner membrane[74]. For α-helical membrane proteins interaction of a hydrophobic helical segment with the signal recognition particle (SRP) allows co-translational targeting to the Sec translocon for membrane insertion[75,76]. In contrast, β-barrels, as well as lipoproteins and soluble proteins to be secreted, are post-translationally targeted to the Sec translocon via an N-terminal signal sequence[75–78]. In this pathway, trigger factor binds to the signal sequence after translation, preventing interaction with the signal recognition particle[79–82]. The chaperone SecB then interacts with these proteins and

targets them to the Sec translocon[83,84]. After translocation into the periplasm by the Sec translocon using energy from the ATPase SecA[85], the signal sequence is cleaved[86].

1.4.3. Transport to the outer membrane by chaperones

Once in the periplasm, chaperones in several pathways exist to stabilize the nascent outer membrane proteins to protect them from aggregation. One pathway involves SurA, which is believed to be involved in the transport of most substrates. A second pathway, which involves the Skp and DegP proteins, can function when SurA is not present and may be important for promoting the assembly of substrates that have gone off-pathway[87].

It has been demonstrated that SurA preferentially binds substrates (in the form of extended, unfolded outer membrane protein-derived peptides) that contain the Ar-X-Ar motif, where Ar represents an aromatic residue[88,89]. Such sequences are found frequently in β-barrel outer membrane proteins. Multiple copies of SurA are believed to bind a single substrate, resembling beads on a string[90]. Although SurA is not essential, and outer membrane proteins can presumably use other chaperones for targeting to the outer membrane, the absence of SurA causes depleted levels of assembled β-barrels[87,91–93].

Another chaperone pathway involves the protein Skp (seventeen kilodalton protein). This protein forms a functional trimer and can be observed in complex with unfolded β-barrels[94–97]. The trimeric structure forms a hydrophobic cavity that can house unfolded substrates, preventing their aggregation by forming many weak and transient interactions with them[98,99].

Skp is believed to work in the same pathway as the degradation machinery DegP. DegP has a hexameric resting state, but upon binding an unfolded outer membrane protein, it can adopt a structure containing 12 or 24 monomers[100–102]. The cage can take in misfolded proteins and degrade them. Since DegP is not essential, most substrates probably do not require this protein, as would be expected if it handles the minority of substrates that fall off pathway. It has been

hypothesized that DegP may act as a chaperone in addition to its function as a protease, but a complete understanding of this function awaits further testing[101].

1.4.4. Folding and insertion into the outer membrane

Once unfolded outer membrane proteins are transported through the periplasm to the outer membrane, folding and membrane integration can occur. In Gram-negative bacteria, the β-barrel assembly machine (Bam) complex (**Figure 1.5**) accelerates the folding and membrane integration of these transmembrane proteins[72,73,78,103,104]. In *E. coli*, the Bam complex is composed of five proteins[73,103]. The core of the complex is BamA, an essential protein that belongs to the Omp85 superfamily of outer membrane proteins that function as protein translocation or assembly factors[105] and is conserved across all Gram-negative bacteria[106,107]. BamA contains five N-terminal soluble periplasmic polypeptide transport-associated (POTRA) domains and a 16-stranded C-terminal β-barrel transmembrane domain. The POTRA domains act as a scaffold that mediates interaction with four lipoproteins (BamB, BamC, BamD, and BamE) that are attached to the inner leaflet of the outer membrane with an N-terminal lipid anchor. Together, the periplasmic components of the complex create a protein vestibule beneath the membrane[108–111]. Given the membrane localization of BamA, this component is believed to perform a central role in substrate assembly. BamA and BamD are the essential components of the complex and are conserved in Gram-negative bacteria, but all components are needed for full activity of the *E. coli* Bam complex[104,106,107,112,113]. Importantly, the Bam complex must accelerate folding in the absence of an energy source, as no ATP is present in the periplasm. This machine must also accelerates folding of β-barrels containing vastly different amino acid sequences and numbers of β-strands[104,111,113–116]. Therefore, this machine must accelerate folding by exploiting features common to its diverse substrates.

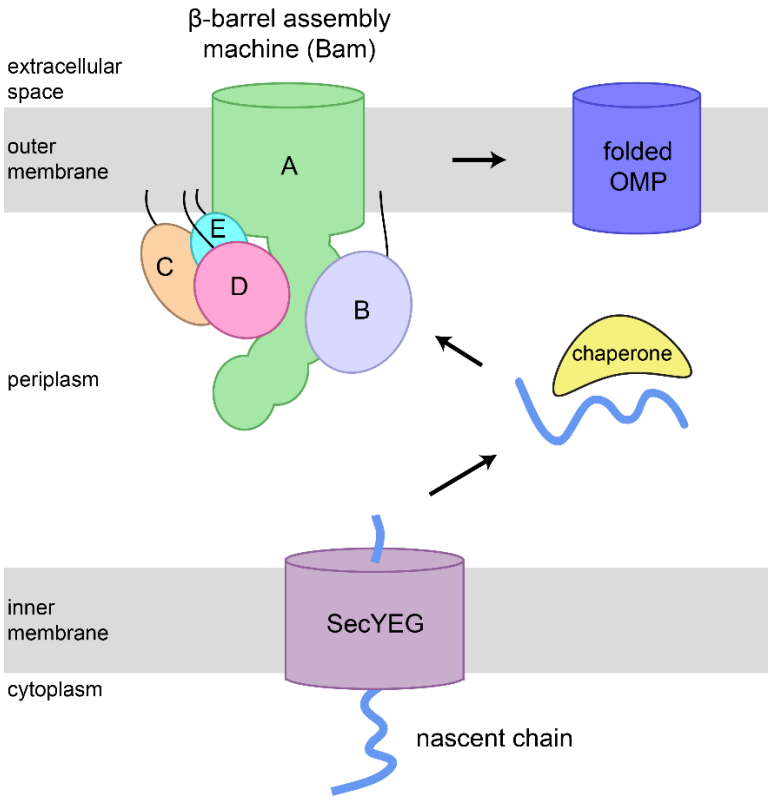

Figure 1.5. The Bam complex folds and inserts β-barrel transmembrane proteins into the outer membrane
The five-protein β-barrel assembly machine (Bam) receives unfolded β-barrels from chaperones and folds and inserts them into the outer membrane.

1.4.4.1. Identification of the Bam complex

The identification of the Bam complex occurred first through identification of the central component of the complex, BamA[72]. In this early work, BamA, originally called Omp85, was identified based on homology to Toc75, a component of the chloroplast protein-import machinery. In several Gram-negative species, Omp85 was found to be an essential protein, and its depletion resulted in accumulation of improperly assembled β-barrel outer membrane proteins.

Several years after identification of BamA, the full complex was discovered[71,73,103]. To do this, an *E. coli* strain was used that has outer membrane permeability effects due to a mutation in the LptD protein (*lptD4213*) (i.e., this strain has heightened sensitivity to antibiotics that normally do not pass through the outer membrane)[117]. A *bamB* null mutation was identified that restored the permeability barrier, resulting in reduced sensitivity to antibiotics. Introducing an affinity tag

on BamB and subsequent purification pulled down BamA, as well as BamC, BamD, and BamE. These proteins thus form a stable complex that can be isolated from cells.

1.4.4.2. Structures of isolated components of the complex

Many structures have been obtained of individual components of the Bam complex using X-ray crystallography and NMR spectroscopy. The first structural information regarding the Bam complex was of the soluble POTRA domains of BamA from *E. coli*[118–121]. In the earliest study[118], the crystallized construct contained POTRA domains 1 through 4 and a fragment of POTRA domain 5. Each individual POTRA domain contains about 75 residues with two antiparallel α-helices packed next to a three-stranded β-sheet (**Figure 1.6a**). Each POTRA domain is modular, and they are connected by short linkers. Interestingly, in the crystal structure, the fragment of POTRA 5 formed a β-stranded structure that interacted with a β-strand of POTRA 3 in a different monomer. In other words, the β-sheet of POTRA 3 could be extended by an additional β-strand (in a parallel arrangement), a mechanism termed β-strand augmentation[122,123]. Another similar study of the POTRA domains of BamA observed antiparallel β-strand augmentation[119]. Based on these works, it has been speculated that the Bam complex could interact with β-strands of substrates through this mode. Importantly, recognition by β-strand augmentation could allow this machine to process a range of substrates with diverse sequences, but which would all have β-strands. Initiation of β-sheet formation of the substrate could be catalyzed using the existing β-sheet of the POTRA domains within the machine.

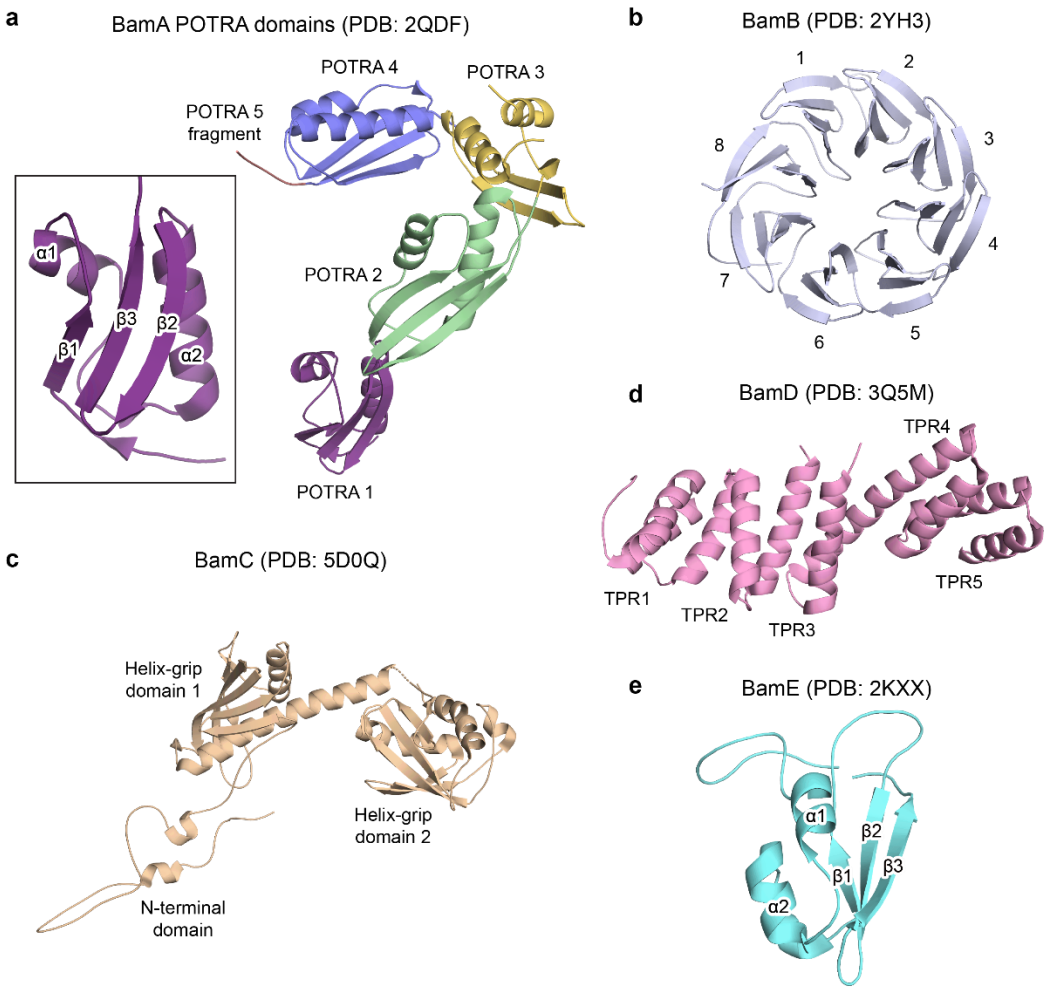

Figure 1.6. Structures of the soluble components of the Bam complex
a, First structure of the POTRA domains of BamA. The domains are labelled, and the inset shows a closer view of POTRA domain 1. **b**, Structure of BamB with β-propellers labelled. **c**, Structure of BamC. Note that the structure shown is from a structure of the full Bam complex (see below) but is shown here since it contains all domains of BamC, which are labelled. **d**, Structure of BamD with TPR motifs labelled. **e**, Structure of BamE showing the ααβββ fold.

After this initial work, full structures of the lipoprotein components were made available (**Figure 1.6b-e**). BamB was revealed to have an eight-bladed β-propeller fold with WD40-like motifs (**Figure 1.6b**). The role of BamB has not been well described, but it has been hypothesized the BamB serves as a scaffolding protein based on its structure and may modulate the flexibility of the POTRA domains of BamA with which it interacts. Structures of BamC have been reported, and these structures showed three domains: an unstructured N-terminal domain followed by two helix-grip domains (**Figure 1.6c**). Cellular microscopy studies have shown that these helix-grip

domains are surface-exposed in *E. coli*, but the functional significance of this remains unclear[124]. BamD is composed of five tetratricopeptide repeat (TPR) domains, which are often found to mediate protein-protein interactions (**Figure 1.6d**). BamD has been suggested to activate BamA and interact with substrates[31,125–129]. Finally, structures of BamE, the smallest of the Bam components, revealed an αββββ fold (**Figure 1.6e**). The function of BamE remains unclear, but it may modulate the structure of BamA in some manner to facilitate substrate folding. Interestingly, deletion of certain lipoproteins can have different effects on different substrates, indicating that some substrates may have different assembly requirements than others[130–132].

Structures of BamA that included the β-barrel domain were then published[133,134]. In one study, structures from *Neisseria gonorrhoeae* and *Haemophilus ducreyi* were reported[133]. Intriguingly, two different states of the β-barrel domain of BamA were observed: a laterally-closed state and a laterally-open state (**Figure 1.7**). In the closed state, the N- and C-terminal strands are joined, and this cavity is inaccessible from the membrane (**Figures 1.7a and 1.7c**). In the open state, a lateral opening is present where the N- and C-terminal strands would join, resulting in an entry into the central cavity of the barrel (**Figures 1.7b and 1.7d**). The open state is unusual since outer membrane proteins are typically found in the closed state with a substantial hydrogen bonding network between the first and last strands. Additionally, the hydrophobic belt around the BamA β-barrel was observed to be narrower along the C-terminal strand (~9 Å) compared to the opposite side of the barrel (~20 Å). Molecular dynamics simulations suggested that this narrowed hydrophobic belt was correlated with disorder of the membrane lipids in this region. The POTRA domains also appeared in different positions in the two structures. Taken together, a model was proposed in which BamA uses its lateral opening to interact with β-strands of substrates. BamA could also use the positioning of its POTRA domains to allow entry of substrates into the membrane while also disrupting lipids to facilitate the process. (See section 1.4.4.5 below for more detail about different models for folding.)

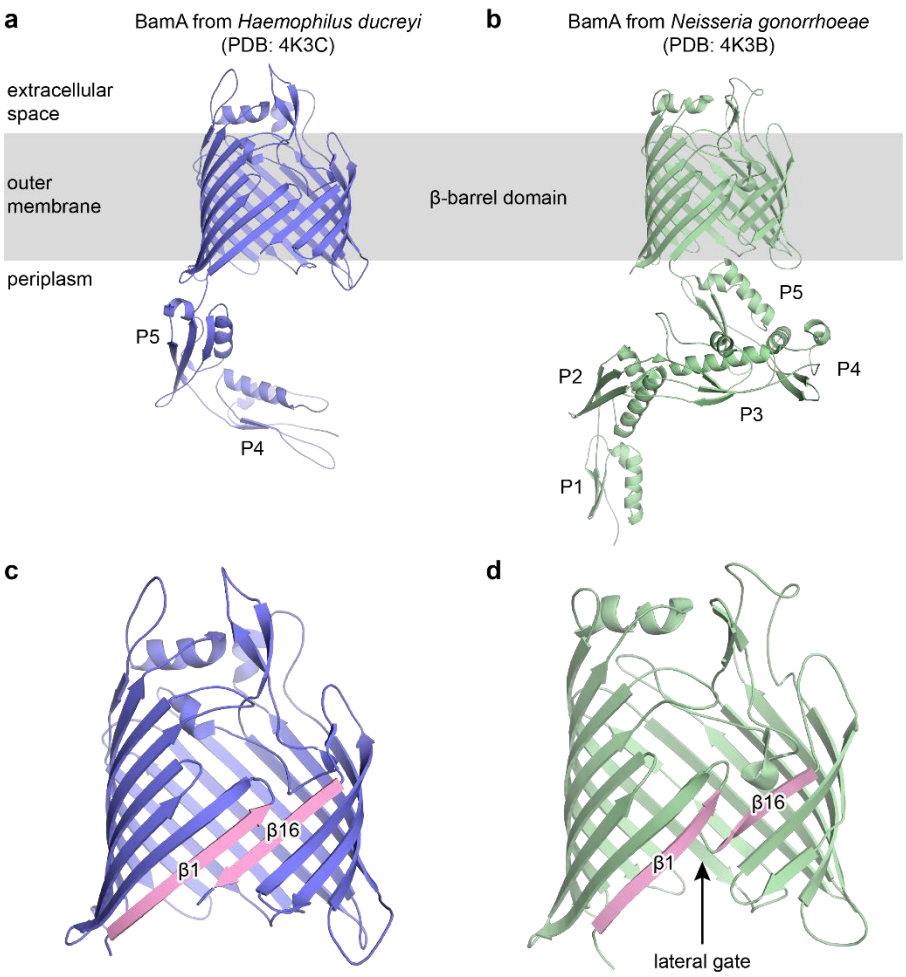

Figure 1.7. Structures of the β-barrel domain of BamA
a, Structure of BamA from *Haemophilus ducreyi*. The two POTRA domains included in the crystallized construct are labelled. **b**, Structure of BamA from *Neisseria gonorrhoeae*. All POTRA domains were present in the structure and are labelled. **c**, Enlarged view of the β-barrel domain of BamA from *Haemophilus ducreyi*. **d**, Enlarged view of the β-barrel domain of BamA from *Neisseria gonorrhoeae* showing the partially open junction between the first and last β-strands.

1.4.4.3. Structures of the full Bam complex

A significant step towards understanding the function of the Bam complex came in 2016 when four groups published five total structures of the complex[108–111]. Three of the structures contained all five proteins (PDB: 5D0O, 5AYW, 5LJO) and two structures lacked BamB (PDB: 5EKQ, 5D0Q). These structures revealed the overall arrangement of the lipoproteins on the POTRA domains of BamA. The general arrangement of the proteins is the same in each structure,

with BamD and the POTRA domains of BamA forming a periplasmic ring. It was speculated that substrates may enter this chamber as a part of the folding mechanism.

Despite the similarities among the structures, two distinct conformations of the complex could be discerned. These states differ in the position of the periplasmic ring structure relative to the BamA β-barrel and in the openness of the β-barrel. In one conformation, the periplasmic ring is located directly below the barrel, which is in a closed state (**Figure 1.8a-b**). In the second state, the ring is shifted so that the bottom of the barrel is blocked by POTRA 5, and the barrel is laterally open to the membrane (**Figure 1.8c-f**). Notably, this open state of BamA in the full complex is caused by outward rotation of the first seven β-strands, with smaller changes in the remaining β-strands (**Figure 1.9**). Additionally, the first three extracellular loops are flipped outward, allowing further opening of the barrel. These changes result in an opening to the membrane that is more obvious than in the isolated form of BamA (i.e., not within a Bam complex) described above.

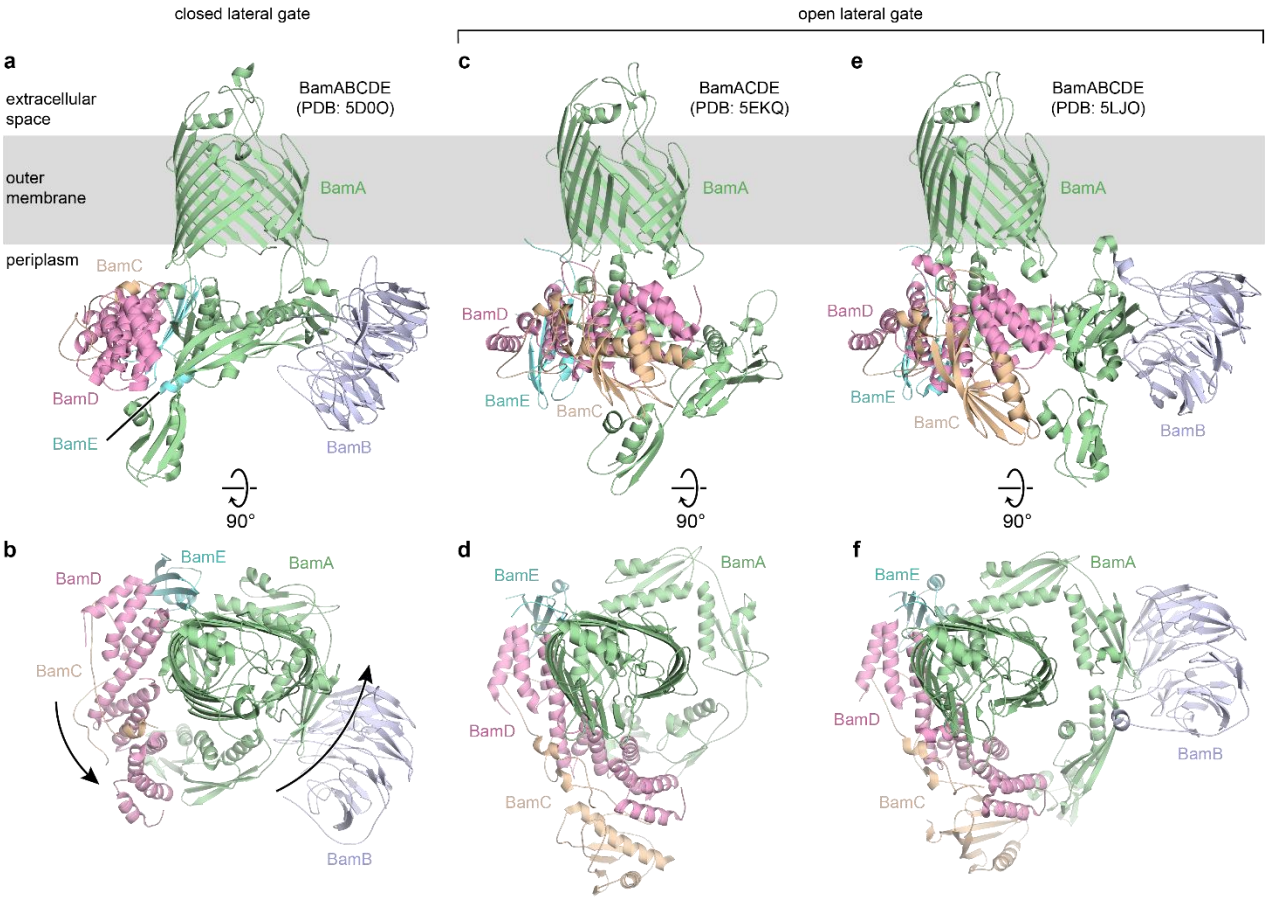

Figure 1.8. Structures of the Bam complex
a, Structure of the Bam complex in which the lateral gate of BamA is closed. **b**, Top-down view of complex in **a**. **c**, Structure of the Bam complex lacking BamB in which the lateral gate of BamA is open. **d**, Top-down view of complex in **c**. **e**, Structure of the Bam complex in which the lateral gate of BamA is open. **f**, Top-down view of complex in **e**.

Based on these structures of the Bam complex, it has been hypothesized that the complex containing the closed conformation of BamA may represent the resting state of the machine and the complex containing the open conformation may represent the state in which the machine is actively folding substrates. However, since all available structures were determined in the absence of substrates, support for this hypothesis is lacking.

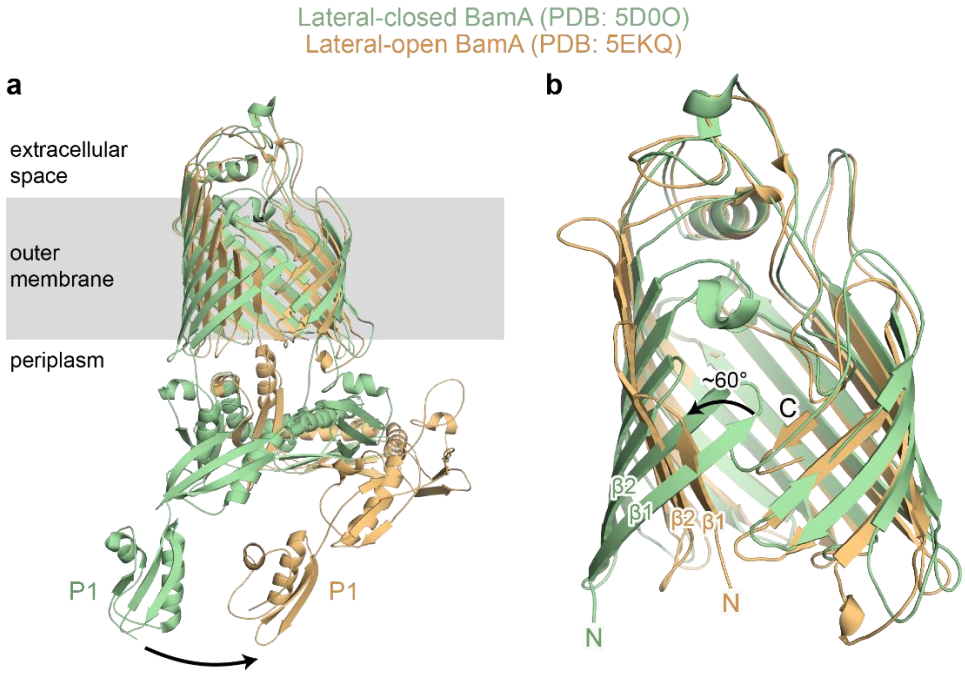

Figure 1.9. The β-barrel domain of BamA within the Bam complex
a, Structure of BamA from the Bam complex. **b**, Overlay of the β-barrel domain of BamA from a structure with a closed lateral gate and a structure with an open lateral gate. In the laterally-open structure, the initial β-strands are flipped outward approximately 60 degrees.

1.4.4.4. Mechanistic studies of the Bam complex

After identification of the components of the complex and some of the structural studies, a major step towards understanding the machine came when the function of the complex was reconstituted *in vitro* using purified components[104]. To do this, the complex was purified and

incorporated into liposomes and an unfolded substrate, OmpT, was added. The enzymatic activity of this substrate was used to report on its folding state. This study confirmed the function of the Bam complex in folding outer membrane proteins and, importantly, folding was able to proceed in the absence of an obvious energy source (e.g., ATP). This was anticipated since the machine resides at the outer membrane without any connection to the inner membrane but is unusual in terms of protein folding machines.

After the initial crystal structures of the BamA β-barrel were published, a study examined the role of lateral opening of BamA in its ability to fold substrates[135]. This study showed that, if the lateral gate was tethered closed, this resulted in cell death. Thus, the lateral opening of BamA is important for its function. However, it was unclear whether BamA must open laterally because the edges of BamA must interact directly with substrates, because substrates must simply pass through the opening, or because the opening allowed for disruption of lipids to permit entry of substrates into the membrane.

Studies have also focused on the role of the C-terminus of substrates in their folding. It has been known for many years that the C-terminal residue of β-barrels is important for their proper folding[136]. In one study, measurements of conductivity were used to show that BamA forms a pore, and the properties of the pore were assessed upon addition of unfolded outer membrane proteins peptides corresponding to their C-terminal sequences[137]. Interestingly, this study demonstrated that β-barrels contain species-specific C-terminal motifs that are recognized by BamA. Another study showed that that a C-terminal motif of substrates is important for binding a different component of the complex, BamD, and a role of this sequence in targeting to the Bam complex was proposed[138,139].

Other biochemical studies that have been performed have focused on characterizing slow-folding substrates that accumulate on the Bam complex. For example, intermediates in the folding of a group of outer membrane proteins known as autotransporters have been examined. Autotransporters are virulence factors that consist of an N-terminal extracellular (passenger)

domain and a C-terminal β-barrel domain (β domain) that becomes integrated into the outer membrane[125,140,141]. Early work showed that BamA, BamB, and BamD subsequently interacted with discrete regions of the β-barrel, with the lipoproteins BamB and BamD associating with the substrate for a longer amount of time, leading to a model in which the lipoproteins play an important role in the integration of the substrate into the membrane[125]. Later work showed that substrate folding is initiated in the periplasm and a model was proposed in which substrates fold outside the membrane[142], held by the N-terminal end of BamA, and then enter the membrane through a "swinging" motion mediated by BamA[143].

Other published work has focused on a different substrate of the Bam complex, LptD, and a variant called LptD4213 that stalls on the machine during its assembly[31,117]. LptD4213 lacks a 23-amino acid stretch within β-strand seven and extracellular loop four (LptD4213) and accumulates as a late stage folding intermediate. It was found that the substrate interacted directly with two essential β-barrel assembly machine (Bam) components, BamA and BamD, and its folding appeared to be templated by the lipoprotein, LptE, that resides within the folded β-barrel. Because BamD is a soluble protein, it was proposed that folding by the Bam complex begins in the periplasm at the membrane interface, similar to the work described above with autotransporters. Another mutation in the LptD protein, Y721D, has been shown to cause assembly defects due to an inability to properly interact with BamD; this defect could be rescued by making a compensatory mutation in BamA (F494L). This suggests that, normally, proper substrate binding to BamD induces a conformational change in BamA to allow membrane insertion[129]. Such activation of BamA by BamD has also been suggested in other studies[126].

The importance of extracellular loops has also been studied in the context of the assembly of BamA[127]. In the folded state BamA, extracellular loop 6 is buried within the β-barrel and contacts its interior wall. Mutations of V660 and R661 within this loop, which would prevent these contacts, caused periplasmic accumulation of BamA during its assembly. It was concluded that, as with

LptD, some folding therefore occurs outside the membrane environment and that extracellular loop burial within the forming β-barrel is important for proper folding.

Work with stalled substrates bound to Sam50, the mitochondrial homolog of BamA, has also demonstrated important principles of the assembly process[144]. Using cysteine crosslinking, it was demonstrated that peptide substrate fragments crosslink strongly to the edges of the Sam50 β-barrel[144]. Based on these experiments and the known structures of BamA, it was concluded that new β-strands were being added to the open junction of Sam50 at the N- and C-terminal β-strands.

1.4.4.5. Models for folding by the Bam complex

Structures and biochemical experiments described above have led to models for the mechanism of substrate folding[108–111,133]. One model, the "BamA-assisted" model, takes advantage of the fact that β-barrels have the ability to fold spontaneously *in vitro* into membrane bilayers or detergent molecules[145–150]. In spontaneous folding, an unfolded outer membrane protein binds to the surface of the membrane, and secondary structural elements begin to form (**Figure 1.10**). The protein then transitions into a molten globule state with formation of β-hairpins that penetrate the membrane until the extracellular side of the membrane is reached. In the BamA-assisted model (**Figure 1.11a**), BamA is proposed to disrupt membrane lipids at its seam, allowing substrates to fold and insert into the destabilized membrane through their intrinsic ability to self-assemble[133,151]. The driving force for folding would be the intrinsic thermodynamic stability of the folded protein, with the Bam complex reducing the activation energy for membrane integration by disrupting the membrane. Support for this model has been provided by molecular dynamics simulations showing thinning of the membrane at the seam of BamA[133], as well as *in vitro* biochemical experiments showing increased folding efficiencies in thinned or disrupted bilayers[150,152–155]. Although a substantial amount of *in vitro* work has been done to support this model, such studies are not performed in the asymmetric membrane present in living cells. *In vivo*

studies in the native membrane environment are required, assessing local membrane disruption at sites of β-barrel assembly *in vivo* would be difficult.

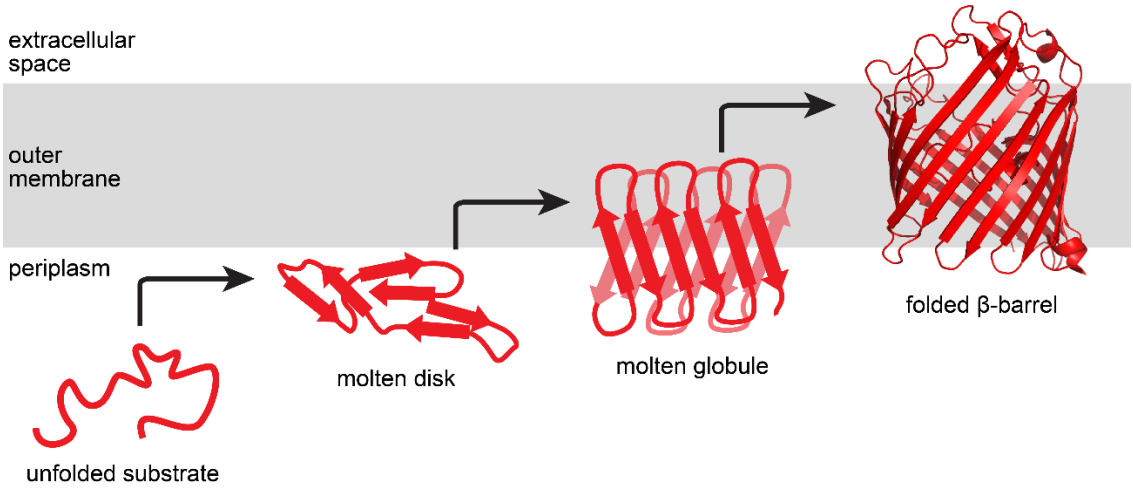

Figure 1.10. Model for the uncatalyzed assembly of outer membrane proteins.
Kinetic studies of OmpA folding in the absence of an assembly machine identified several intermediates in the process.

Alternatively, the "BamA-budding" model[133,135,144,156,157] is based on the striking finding in some of the structures of the complex that the BamA component is present in an open state with its N- and C-terminal β-strands unpaired[108,109,111,133]. The BamA-budding model suggests that BamA uses its open seam to hold substrates in a configuration that facilitates folding (**Figure 1.11b**). Specifically, the exposed edges could interact with β-strands of the substrate through β-strand augmentation. The substrate may be assembled by systematic insertion of β-strands or β-hairpins into the membrane with the exposed N- and C-terminal edges of BamA holding the opposite edges of the growing barrel by β-strand augmentation to satisfy the hydrogen bonding networks at each edge of the growing β-sheet. This could be initiated by interactions between the C-terminus of the substrate with the lateral opening of BamA, and β-strand addition would occur using the template formed by each strand previously added. Once the full substrate β-barrel has formed, it would be released from BamA through a "budding" event to exist as a separate β-barrel in the membrane. The exact mechanism of substrate release, however, is unclear. Studies demonstrating that lateral opening of BamA is required for proper function provide support for this

mechanism[111,135]. Molecular dynamics simulations have also demonstrated that the seam between β1 and β16 of BamA is dynamic[133,135]. Additionally, crosslinking experiments described above with Sam50, the mitochondrial homolog of BamA, have shown that its N- and C-terminal β-strands contact the edges of substrate proteins[144]. Finally, experiments testing the effects of closing the lateral gate of BamA with disulfides, described above, provides important evidence for the budding model, but does not rule out other possibilities.

Variations of the BamA-budding model propose that some or all folding occurs outside of the membrane environment. In these models, only the C-terminal edge of the substrate is held stably by BamA. In one variant of the budding model, the substrate is held by BamA, but is partially or fully exposed to the periplasm during folding, with full membrane integration occurring at a late stage of assembly[31,127,142,143,158]. The notion of folding outside the membrane is intriguing, and evidence has been provided through studies of the autotransporter EspP[125,140–143], LPS translocon LptD[31,129], and BamA[127]. Since similarities have been observed in the folding pathways of these proteins, despite their significant structural differences, this provides evidence that periplasmic assembly might be a general principle shared by diverse substrates.

It is important to note that none of the models have been demonstrated to apply to all outer membrane proteins; indeed, different substrates may follow different folding pathways depending on their size or other features. Whether Bam complex-mediated folding proceeds via direct interactions with substrates, or whether a destabilized membrane is all that is needed, is the subject of ongoing work.

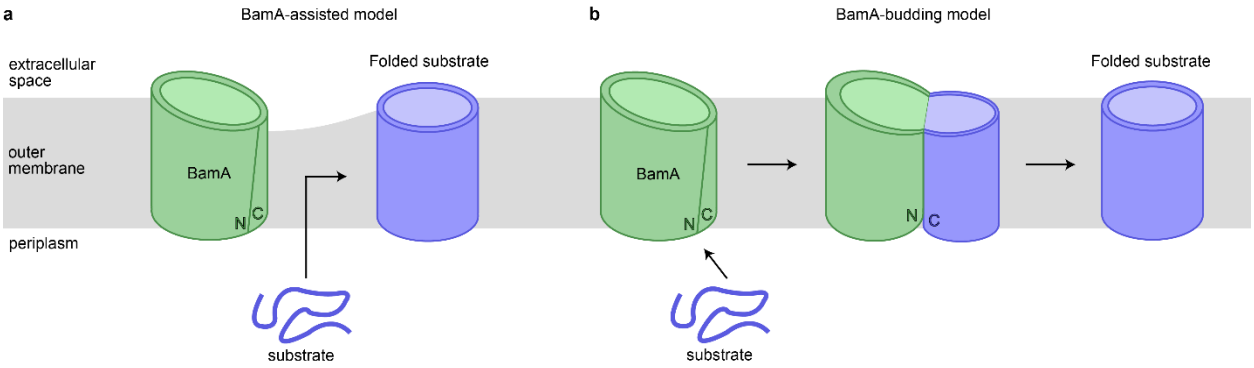

Figure 1.11. Current models for outer membrane protein assembly by the Bam complex.
a, In the BamA-assisted model, the role of BamA is to disrupt membrane lipids with its reduced hydrophobic belt around the junction between the N- and C-terminal β-strands. This destabilization or membrane "thinning" allows substrates to fold into the membrane using their intrinsic ability to do so. **b**, The BamA-budding model assigns a role to the lateral opening of BamA. The opening is believed to hold the ends of the substrate as they fold, forming a "super-barrel" consisting of the two proteins before the folded substrate buds off into the membrane.

1.5. Perspectives

Since the discovery the Bam complex, much effort has been dedicated to understanding how this molecular machine functions. After identification of the components, the activity of the complex was reconstituted, and structures of individual components and the full complex were obtained. However, all available structures of the Bam complex are in the absence of substrates. Although structures have revealed how the different components of the complex interact with each other, the significance of the two general conformations of the Bam complex is unclear. It also remains unclear how substrates interact with the machine, and how these interactions allow this machine to catalyze rapid and repeated assembly of substrates into β-barrels in the absence of an exogenous source of energy[104,115,116].

The properties of outer membrane proteins mean that a mechanism for their assembly must address several issues. For example, unlike α-helical proteins, which can laterally be inserted into the membrane as stable secondary elements with their hydrogen bonding networks satisfied, individual strands or hairpins would be unstable in the membrane if the hydrogen bonding groups at their edges are unsatisfied. In other words, the hydrogen bonding network of a β-barrel would only be satisfied once the entire barrel has formed and the N- and C-terminal β-

strands interact via hydrogen bonding. Thus, the mechanism of assembly must ensure that the edges of the substrate are satisfied. Additionally, the mechanism must account for the fact that there is no obvious energy source (e.g., ATP) in the periplasm. In other words, the Bam complex must have an energy-independent mechanism of folding and turning over to allow for further catalysis of folding. A mechanism for folding must explain how the Bam complex can accelerate folding but must also explain how substrates can be released. While the BamA-budding model is intriguing in light of recent biochemical experiments, the stable interactions between the edges of BamA and the substrate would seem to pose a problem: such a “super-barrel” structure would be highly stable, a large barrier would exist to break the interactions between BamA and the substrate, and budding would thus be difficult to imagine.

At the onset of the work described here, a detailed molecular mechanism for substrate folding did not exist. Previous studies have been almost exclusively biochemical or exclusively structural in nature. Future work will need to use both types of studies to examine how substrates interact with the machine, and how these interactions and conformational changes in the machine accelerate the folding while allowing efficient release once folding has completed. Isolation of multiple folding intermediates that accumulate on the machine during folding will be essential to gain snapshots of different stages of the process.

Chapter 2: Trapping LptD folding intermediates on the Bam complex and preliminary structural studies

Parts of this chapter are adapted/reproduced from:
Lee J, Tomasek D*, Santos TMA*, May MD*, Meuskens I, Kahne D. Formation of a β-barrel membrane protein is catalyzed by the interior surface of the assembly machine protein BamA. *eLife* **8**, e49787 (2019).

2.1. Introduction

The prevailing model for folding is based on structures showing an open seam in the BamA barrel where the N- and C-terminal strands interact. It has been suggested that β-hairpins in the substrate assemble at this open seam of BamA in what has been described as the "budding" model because the nascent substrate barrel grows into the membrane as new strands are added at the seam[135,144]. A budding model has also been proposed for the mitochondrial ortholog of BamA called Sam50. It has been demonstrated that peptide substrate fragments crosslink strongly to the N-terminus of the Sam50 barrel and more weakly to its C-terminus[144]. Based on these experiments and the known structures of BamA, it was concluded that new β-strands were being added at the seam in accordance with the budding model. An alternative model holds that an extensive region of β-sheet assembles in the periplasm at the POTRA domains of BamA with one end of the sheet held by one end of the seam[114,158]. These models have focused largely how folding is initiated with less attention paid to explaining how folding is completed and substrates released.

Here, we have studied the folding of a large β-barrel, LptD, and variants that fold more slowly. LptD, a component of the lipopolysaccharide transport machine[22,23,117,159], is one of the two essential β-barrel proteins in *Escherichia coli*, the other being BamA itself. LptD contains 26 β-strands, and must fold around a globular lipoprotein, LptE, which acts as a plug within the β-barrel[26,27,29,160]. LptD is useful as a model substrate for the Bam complex because it folds much more slowly (several orders of magnitude) than smaller β-barrel proteins[33,161] making the process of folding more accessible for study than for other substrates. Moreover, we have previously identified a variant of LptD lacking a 23-amino acid stretch within β-strand seven and extracellular loop four (LptD4213) that accumulates as a late stage folding intermediate that can complete folding[31]. Here, we take advantage of these slow folding substrates and *in vivo* crosslinking to identify contacts between folding intermediates and the Bam complex. The major conclusion from

these crosslinking experiments is that LptD, in the process of folding, forms extensive contacts with the concave interior wall of the BamA β-barrel.

Thus, in contrast to either the budding or periplasmic models for folding, our evidence indicates that folding is catalyzed in the interior of the BamA β-barrel. In agreement with earlier models we show that the substrate LptD is held at its C-terminus through a stable interaction with BamA. In our model, however, the β-hairpins do not form at the lateral gate but rather form inside the BamA β-barrel, generating an extensive β-sheet as folding proceeds from the C-terminus towards the N-terminus. Here we present evidence that release of the N-terminus from the interior wall of BamA ultimately allows substrate β-barrel insertion into the membrane. Importantly, we show that changes to residues in the interior of BamA can accelerate folding, leading us to conclude that it serves as an active site that catalyzes folding. Our results establish a model for β-barrel assembly by the Bam complex where the catalyst for β-strand formation is the interior surface of the BamA barrel.

2.2. Results

2.2.1. The interior surface and lateral gate of the BamA β-barrel form a binding site for substrates

We first sought to identify regions of the Bam complex that interact with LptD substrates. Individual components of the Bam complex can interact with substrates[116,125,126,140,141,153,162–164], but how these interactions facilitate folding is unknown. Because it is essential, is conserved across all Gram-negative bacteria, and is the only transmembrane component in the Bam complex, we decided to focus on the role of the BamA β-barrel. We substituted residues throughout the BamA β-barrel with the unnatural amino acid para-benzoyl phenylalanine (pBPA)[165] to capture interactions between BamA and the substrates LptD and LptD4213. LptD4213 stalls as a late-stage folding intermediate during its assembly[31]. This substrate was

chosen because we expected that the longer residence time of LptD4213 on the Bam complex would allow for more efficient crosslink formation. We introduced pBPA substitutions at the lateral gate of BamA, where the N- and C-termini of the β-barrel meet (**Figure 2.1a**), and within the sixth extracellular loop (L6) (**Figure 2.1b**) because both the lateral gate and L6 were previously proposed to be important in the mechanism of substrate assembly[135,144,166,167]. After photocrosslinking, we purified pBPA-containing BamA variants and assessed the presence of higher molecular weight adducts, representing BamA-substrate crosslinks, by immunoblotting.

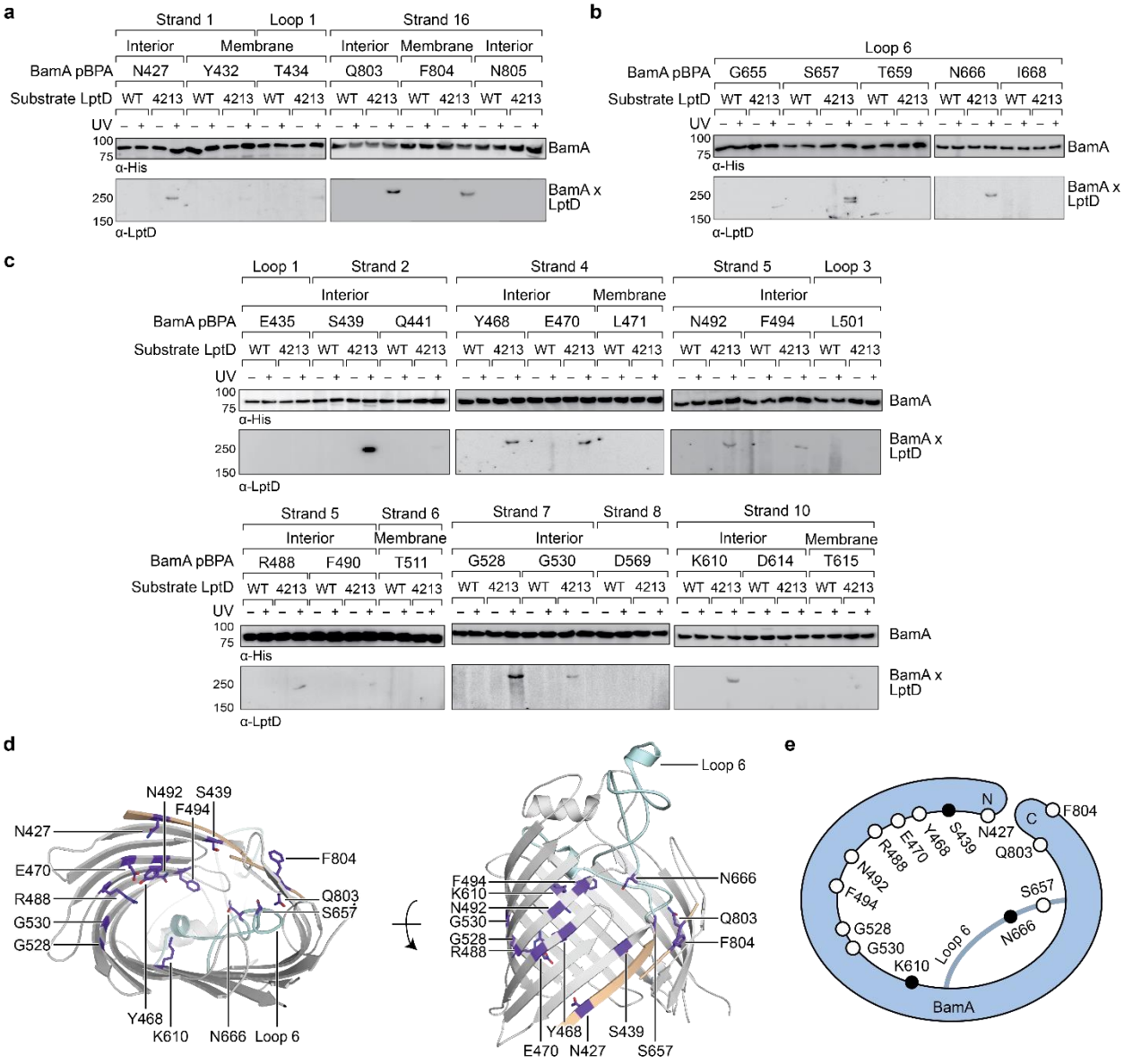

Figure 2.1. The interior wall and lateral gate of the BamA β-barrel form a substrate binding site
a-**c**, Residues at **a**, the lateral gate; **b**, extracellular loop 6; and **c**, the interior wall of BamA interact with substrate LptD during assembly. MC4100 and *lptD4213* (NR698) strains (expressing WT LptD or LptD4213, respectively) harboring the amber suppression system and expressing a His-tagged BamA (containing pBPA) were either left untreated or irradiated with UV light. Crosslinked adducts of BamA and substrate LptD/LptD4213 were identified by immunoblot analyses after Ni-NTA affinity purification. The orientation of the side chain of each residue in BamA substituted with pBPA is indicated (i.e., facing towards the membrane or interior of BamA). **d**, Specific sites in the BamA β-barrel that interact with substrate LptD. Residues substituted with pBPA that crosslink to substrate are colored in purple. The first and last β-strands are colored in tan while the L6 loop is colored in cyan. Images were generated in PyMOL using the crystal structure of the BamA β-barrel from the *E. coli* BamABCDE complex (PDB: 5D0O). **e**, Cartoon schematic of all sites in BamA that crosslink to substrate LptD. The view shown is the same as in the left panel of **d**. ***The immunoblots presented in this figure were obtained by Dr. James Lee.***

We identified three residues at the lateral gate of BamA that showed a UV-dependent crosslink to LptD4213, one in the N-terminal β-strand one (N427) and two in the C-terminal β-strand 18 (Q803 and F804) (**Figure 2.1a and Figure 2.1d-e**). Additionally, pBPA substitutions at two positions in L6 (S657 and N666) yielded crosslinks to LptD4213 (**Figure 2.1b and Figure 2.1d-e**). It is important to note that these two residues in L6 flank the VRGF motif, the most conserved sequence of residues across the entire Omp85 superfamily[167,168]. The crosslinks at the lateral gate and L6 provide the first demonstration that these regions in BamA interact directly with substrate but are consistent with other findings suggesting that these regions in the Omp85 superfamily members contact substrates[144,166].

If β-barrel substrates come in close contact with L6, which resides in the interior of BamA near the lateral gate, then substrates might insert directly at the lateral gate of BamA as previously suggested or might first enter the interior of the BamA β-barrel. The BamA interior surface has not been previously probed for interactions with substrates so we introduced pBPA at 18 positions spanning this surface (**Figure 2.1c**). Of those, we identified nine residues across β-strands two, four, five, seven and ten that showed UV-dependent crosslinks to substrate (**Figure 2.1c-e**). These nine residues and all but one of the four residues described above have their side chains oriented towards the interior of the BamA β-barrel (**Figure 2.1d**). The exception, F804, resides at the C-terminal edge of the lateral gate where the N- and C-terminal ends of the BamA β-barrel interact. Together, the residues that directly interact with substrate form an extensive surface that

includes a substantial portion of the interior wall of the BamA β-barrel (**Figure 2.1d-e**). Notably, several interacting residues, including G528, G530, and K610, are located directly opposite of the lateral gate. This is an important observation because it is not consistent with the proposal that the β-strands of substrates are sequentially added to the nascent β-barrel via interactions that occur between β-strands 1 and 16 at the BamA lateral gate[135,144]. Because our results show that the LptD4213 substrate forms interactions with the interior surface of BamA far from the lateral gate, substrates may begin to fold inside the BamA β-barrel before exiting through the lateral gate of BamA into the membrane.

2.2.2. BamA and BamD bind non-overlapping regions within the C-terminal strands of substrates

The identification of a large surface of the BamA interior wall that interacts with substrates implies that a large portion of the stalled LptD4213 substrate interacts with BamA. We focused on defining contacts to BamA from residues in the substrate LptD barrel. It has been proposed that there exist recognition sequences in substrates that can interact with either BamA[137] or BamD during assembly[129,139]. Using pBPA substitutions near the C-terminus of LptD4213 (**Figure 2.2a-b**), we identified residues in β-strand 24 that interact with BamD, but not BamA (**Figure 2.2a and Figure 2.2c**) and residues in β-strands 25 and 26 that interact with BamA, but not BamD (**Figure 2.2b-c**). Even at native expression levels, several of these pBPA substitutions efficiently crosslinked wild-type LptD to BamD or BamA. These results are consistent with previous findings that the C-terminus of β-barrel substrates associate strongly with the Bam complex[129,137–139]. The regions in the C-terminus of LptD that interact with BamA and BamD are spatially restricted and do not overlap, which would allow both components of the Bam complex to interact with the substrate simultaneously as has been previously suggested[125].

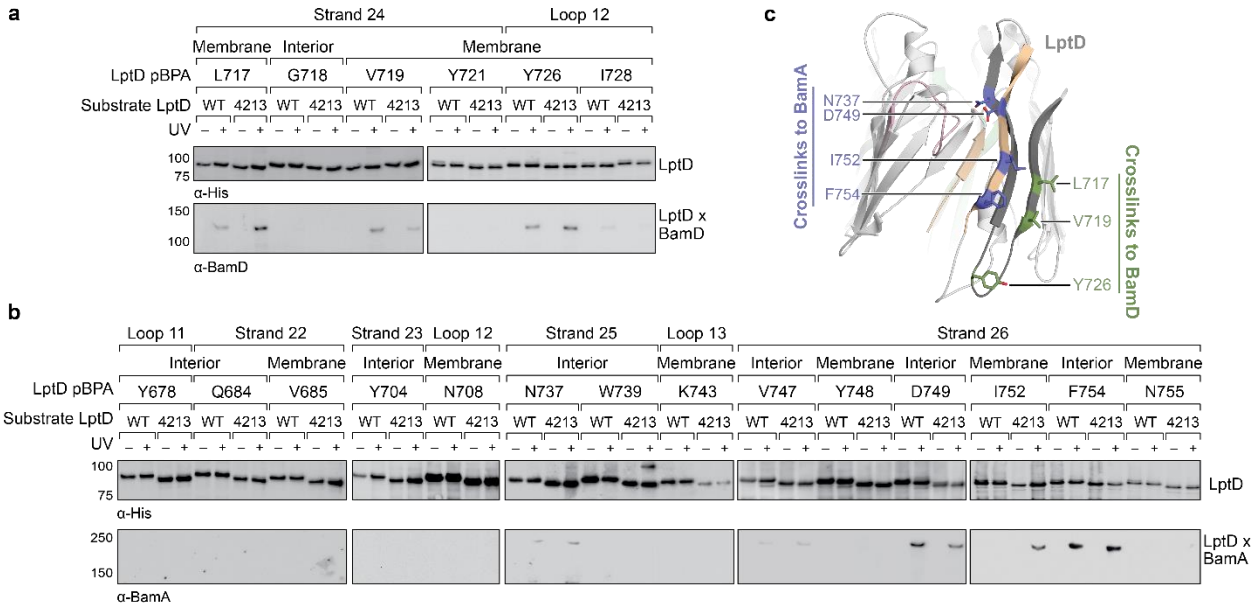

Figure 2.2. The C-terminal strands of LptD interacts with both BamA and BamD during assembly
a, The third-to-last β-strand and final periplasmic loop of substrate LptD interact with BamD during assembly. Crosslinked adducts of BamD and substrate LptD/LptD4213 were identified by immunoblot analyses after Ni-NTA affinity purification. The orientation of the side chain of each residue substituted with pBPA is indicated (i.e., facing towards the membrane or interior of the folded form of LptD). **b**, As in **a**, but showing crosslinking to BamA from the final two β-strands and the final extracellular loop of substrate LptD. **c**, Side view of LptD (gray) mapping the residues in the C-terminal strands that interact with the Bam complex. Residues in cyan and green form crosslinks to BamA and BamD, respectively. Extracellular loop 4 of LptD, part of the region deleted in LptD4213 (pink), interacts with the ends of the LptD β-barrel (tan). Images were generated in PyMOL using the crystal structure of *E. coli* LptD/E (PDB: 4RHB). ***The immunoblots presented in this figure were obtained by Dr. James Lee.***

2.2.3. β-strand formation takes place in the interior of BamA

Crystal structures of wild-type LptD[26,27] show that several residues in β-strands one and two contact the fourth extracellular loop (L4) of the protein. Deletion of this loop in LptD4213[117,159] causes this substrate to stall on the Bam complex as an open β-barrel[31], implying that L4 is important for stabilizing closure. Regarding LptD4213, we found that residues in β-strands one and two, which would contact L4 in the folded state, instead formed crosslinks to BamA in a UV-dependent fashion (**Figure 2.3a**). Because BamA interacts with residues that would be involved in maintaining the closure of substrate once folded (**Figure 2.3b**), one function of BamA during assembly may be to bind substrates to prevent their premature release.

Next, we introduced pBPA substitutions farther away from the N-terminus of LptD. We tested 12 additional residues in β-strands four through six and the associated extracellular loops, which are N-terminal of the 23-residue deletion in LptD4213 (**Figure 2.4**). The same substitutions were introduced into the wild-type substrate. Overall, the crosslinks to LptD4213 were more intense than for the wildtype substrate (**Figure 2.4a**), as expected given the longer residence time of LptD4213 on Bam. We also noted that the crosslink intensities from several residues in β-strands four, five, and six of LptD4213 varied in a periodic pattern (**Figure 2.4**). For example, in strand four crosslinking from residues W272 and F276 was more intense than from residue E275 (**Figure 2.4a, left**) and in strand five crosslinking intensities were greater for L289 and Y291 than for D290 (**Figure 2.4a, middle**). We did not observe periodicity in crosslinking intensities from strands one, two, and three (**Figure 2.3**). The periodicity or lack thereof implies that strands four, five and six have organized into a β-stranded structure but strands one, two, and three have not. Because we found an extensive surface within the interior of BamA that interacts with substrate (**Figure 2.1**), we propose that strands four, five, and six of the substrate are housed in the BamA barrel during the stall. Our data imply that BamA contains an extensive substrate binding site in its interior that can interact with substrates to chaperone β-sheet folding.

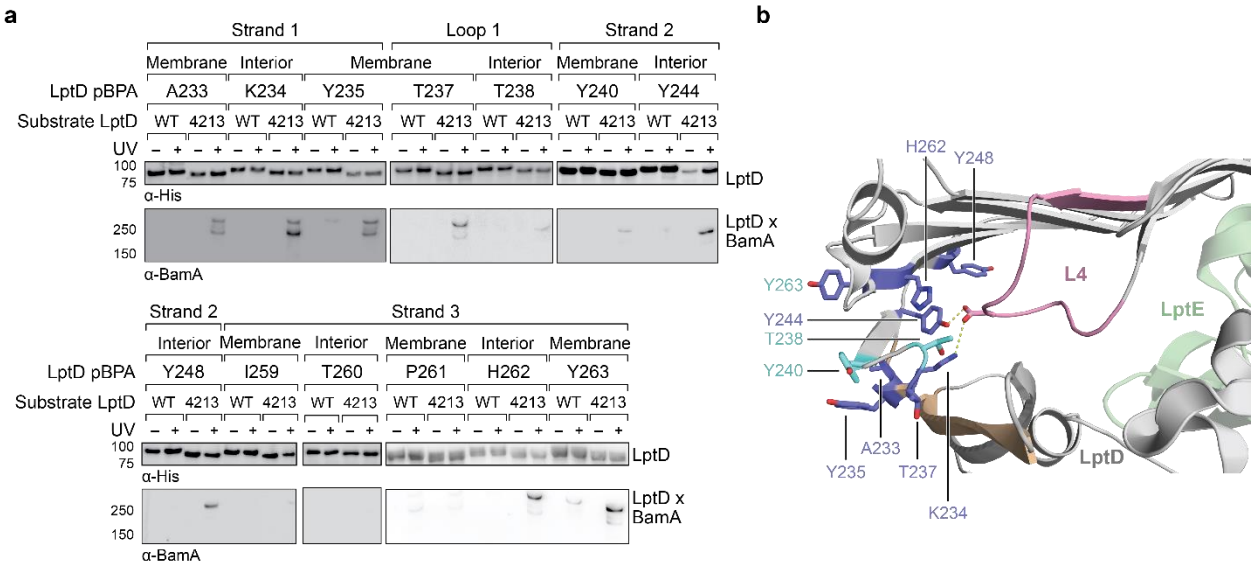

Figure 2.3. Crosslinking of the first three β-strands of substrate LptD to BamA
a, β-strands one, two, and three of substrate LptD interact with BamA. Crosslinking was tested as described in Figure 2.1. **b**, Top-down view of the region near the N- and C-termini of LptD showing residues in LptD4213 that form crosslinks to BamA during assembly but are important in mediating closure of LptD in the folded form. Residues K234, T238, and Y244 interact with L4 in folded LptD (within the region in pink, which is deleted in LptD4213). ***The immunoblots presented in this figure were obtained by Dr. James Lee and Mary May.***

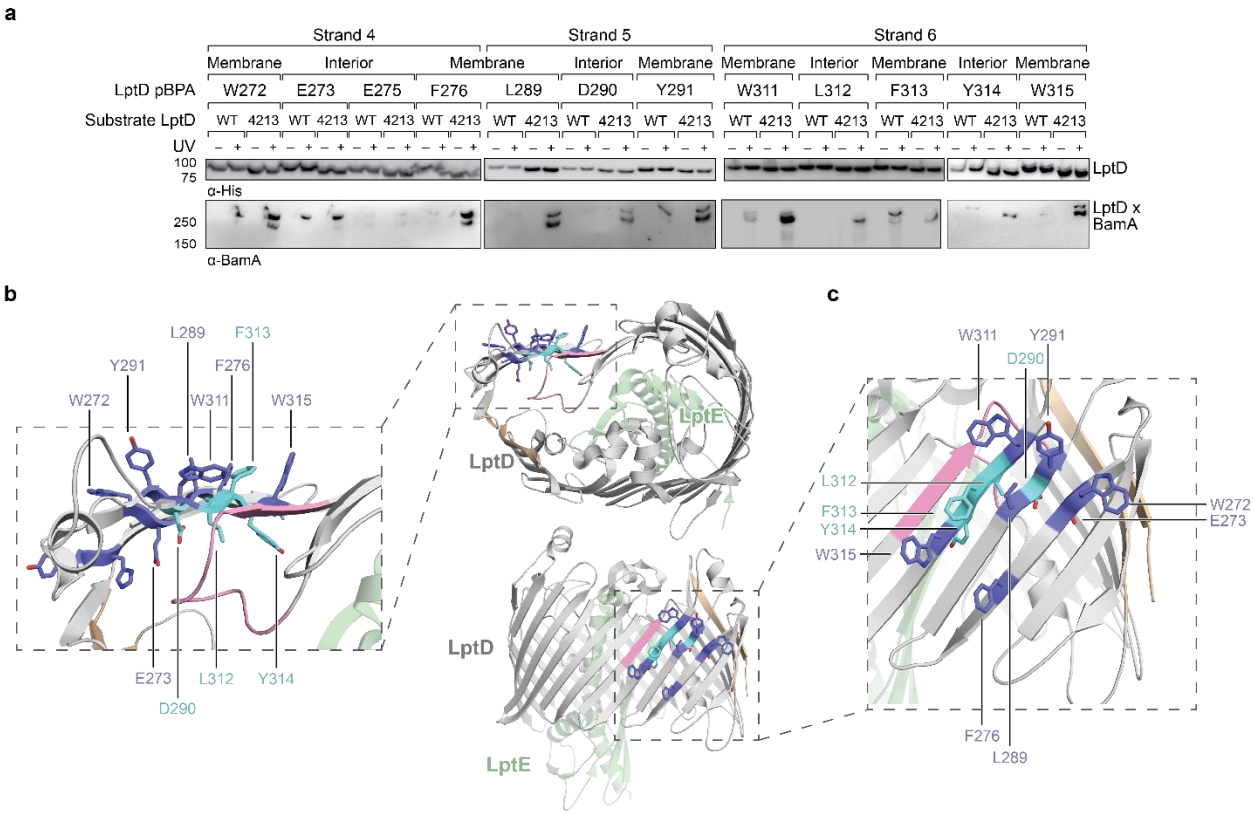

Figure 2.4. Crosslinking of β-strands 3, 4, and 5 of substrate LptD to BamA
a, β-strands four, five, and six of substrate LptD interact with BamA. Crosslinking was tested as described in Figure 2.1, but with pBPA substitutions in the N-terminal portion of substrate LptD/LptD4213. 'Membrane' and 'interior' specify where the indicated residues would face in the mature barrel. **b**, Top-down view of LptD showing that residues in at least three β-strands in the N-terminal region of the LptD barrel interact with BamA. Residues in LptD4213 that form strong crosslinks to BamA are shown in blue, while residues that form weak crosslinks are shown in cyan. The N- and C-terminal strands of LptD are indicated in tan, and LptE is shown in green. This color scheme is maintained in the rest of the figure. **c**, Side view of LptD showing crosslinking positions as depicted in **b**. Note that only crosslinks within a blot can be compared, and each blot includes only proximal residues. ***The immunoblots presented in this figure were obtained by Dr. James Lee and Mary May.***

Our *in vivo* photocrosslinking data show that the C-terminus of the substrate (**Figure 2.2**) and an extensive region of the N-terminus of the substrate both interact with BamA (**Figure 2.3 and Figure 2.4**). If this model is correct, it should be possible to form bidirectional crosslinks

between the corresponding regions of BamA and the substrate. We substituted residues in BamA and LptD4213 with cysteines and probed for crosslink formation between the two proteins after treatment with 1,4-bis(maleimido)butane (BMB, 11 Å linker) and purification of crosslinked adducts formed by the two proteins. We found that a cysteine placed near the N-terminal strand of BamA (S439C) formed crosslinks to cysteines introduced near the C-terminal strand of LptD4213 (E733C and N737C, **Figure 2.5a-b**). Therefore, this establishes that C-terminus of substrate LptD is held at the N-terminus of the BamA barrel. Additionally, a cysteine placed in the L6 loop of BamA (N666), which resides within the interior of BamA, formed crosslinks to cysteines introduced near the N-terminus of LptD4213 (L245C, H262C and Y291C), but not to cysteines introduced near the C-terminus (**Figure 2.5a-b**). These results establish the topology of substrate LptD on the Bam complex (**Figure 2.5c**) and demonstrate that a large part of the N-terminus of the substrate is held within the interior of BamA.

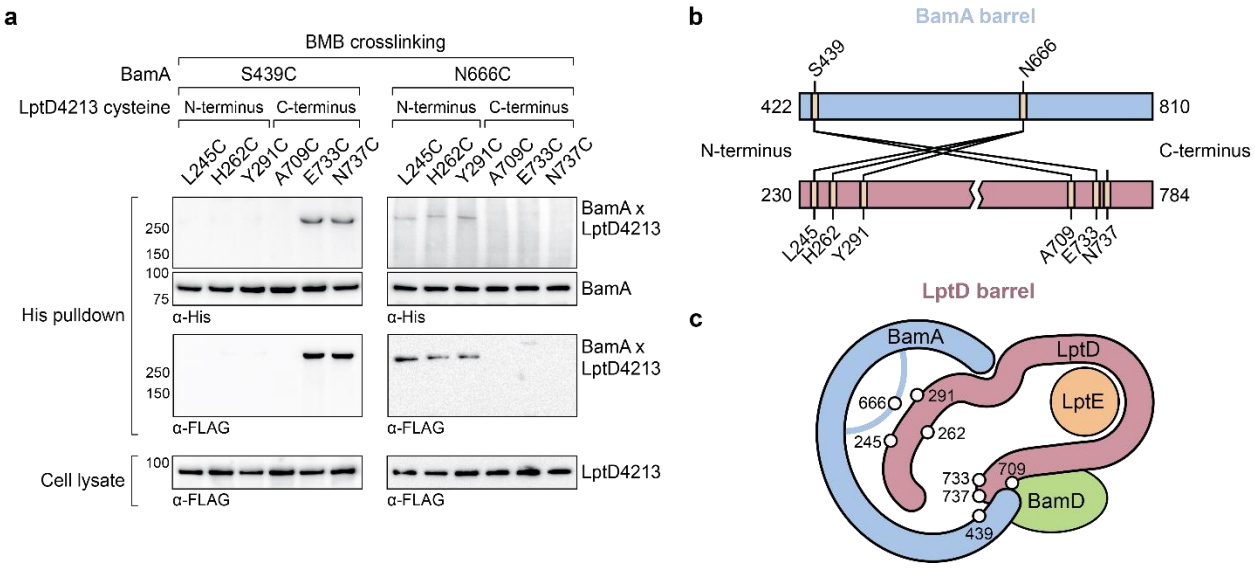

Figure 2.5. The N-terminal strands of the substrate are housed within the BamA β-barrel
a, The lateral gate and interior of the BamA β-barrel interacts with the ends of the LptD substrate. The N-terminal strands of BamA (S439) interacts with the C-terminal strands of substrate LptD, while L6 (N666), within the β-barrel of BamA, interacts with the N-terminal strands of substrate LptD. MC4100 strains expressing a His-tagged BamA cysteine mutant and a FLAG-tagged LptD4213 cysteine mutant were treated with the cysteine-cysteine crosslinker 1,4-bis(maleimido)butane (BMB). Crosslinked adducts of BamA and substrate LptD4213 were identified by immunoblot analyses after Ni-NTA affinity purification. Immunoblots are provided that show expression levels of cysteine-containing LptD4213 constructs from total cell lysates (bottom). **b**, Map of cysteine crosslinking between the BamA and LptD4213. **c**, Top-down cartoon representation of LptD engagement by BamA/D based on cysteine crosslinking data.

2.2.4. Substrate release from the interior wall of BamA facilitates β-barrel closure

Our data show that the interior wall of BamA interacts extensively with a large portion of the β-barrel substrate to chaperone folding. A model in which BamA catalyzes folding against its interior wall requires a rationale for how such an extensively bound substrate can be released once folding is complete. We and others have shown that no energy is required for folding on the Bam complex[104,111,113,114,116,169], and indeed, there is no ATP in the periplasm. The stalled complex provides a clue to the release mechanism. In fully-folded, wild-type LptD, L4 interacts with both the N- and C-termini of the β-barrel (**Figure 2.3b**)[26,27], suggesting that this loop stabilizes the released product relative to the late-stage intermediate that makes extensive contacts with the interior of BamA. When this loop is missing, as is the case in LptD4213, the interactions that normally stabilize the closed β-barrel are absent, and the relative energies of the BamA-associated intermediate and the fully closed β-barrel change such that the substrate stalls on the Bam complex at a late stage of folding.

This model predicts that mutations that destabilize the interactions between the N-terminus of LptD4213 and the interior of BamA may allow for entry into the membrane. An N274I amino acid substitution in LptD4213 was previously identified in a genetic selection for mutations that rescue permeability defects associated with LptD4213[71]. N274 is in β-strand four of the folded β-barrel (**Figure 2.6a, left**) with its polar side chain embedded within a nonpolar surface oriented towards the membrane (**Figure 2.6a, right**). Because our studies here identified crosslinks to BamA from β-strands flanking β-strand 4, we predicted that β-strand four itself would also interact with the interior wall of BamA. In β-strand 4, the side chains of residues 272 and 276 lie adjacent to that of residue 274 and are oriented towards the membrane (**Figure 2.6a**). Thus, pBPA substitutions at these positions should form crosslinks to BamA. To test this hypothesis, we replaced these residues with pBPA in both LptD and LptD4213 and probed for crosslinking to BamA (**Figure 2.6b**). Substitution at W272 and F276, two residues flanking N274, generated strong crosslinks to BamA in both wild-type LptD and the LptD4213 variant; weak crosslinks were

also observed for both proteins after substitution of N274 with pBPA. These results suggest that β-strand 4 of LptD is bound to the interior wall of BamA during folding. Furthermore, substitutions in β-strand four that disrupt polar contacts to the interior wall of BamA shorten the residence of the substrate at this binding site.

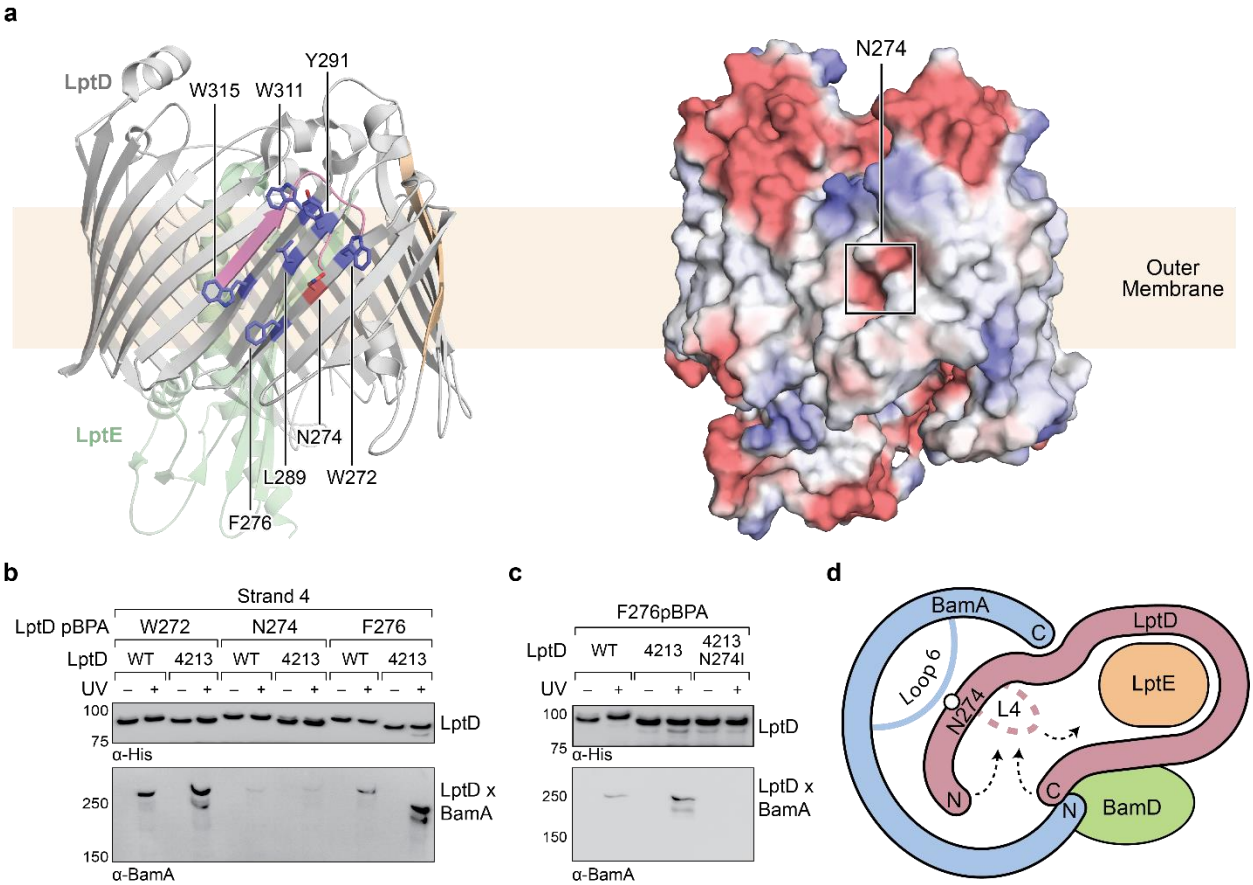

Figure 2.6. Substrate release from the interior wall of BamA allows β-barrel closure, triggering release from the Bam complex

a, N274 resides within a large hydrophobic patch that encompasses at least six β-strands at the N-terminal region of the LptD β-barrel. The left panel shows the structure of LptD in cartoon form. The color scheme is the same as in Figure 2.3, with N274 indicated in red, and six hydrophobic residues that crosslink strongly to BamA in blue. The right panel shows an electrostatic surface plot generated using APBS, presented in the same orientation as the cartoon (left). Colors in the electrostatic surface plot represent potential rather than crosslinking residues. Red represents negative potential, white represents neutral potential, and blue represents positive potential. **b**, The region around N274 directly interacts with BamA. **c**, The N274I mutation suppresses the folding defects associated with LptD4213, allowing release from BamA, as judged by a reduction in crosslinking efficiency. **d**, Model for substrate β-barrel closure and release from the Bam complex. N274I suppresses the folding defect associated with LptD4213 by facilitating release of the N-terminal strands of LptD from the interior of BamA. ***The immunoblots presented in this figure were obtained by Dr. James Lee and Dr. Thiago Santos.***

The weak crosslinks for the N274pBPA substitution suggested an explanation for how the N274I mutant suppresses the assembly defect of LptD4213. Because weaker crosslinking can indicate a shorter residence time, we speculated that the N274I suppressor weakens the affinity of the N-terminus of LptD for BamA. To test this possibility, we replaced residue 276, which is in the middle of β-strand four and faces outward, with pBPA in the LptD4213 N274I mutant. No crosslinking from this residue to BamA was observed (**Figure 2.6c**), consistent with a more transient interaction of β-strand four with the interior wall in the N274I variant. The N274I mutation evidently destabilizes the BamA-associated intermediate so that it is released more rapidly from the interior wall of BamA. Therefore, the ability of the N274I mutation to rescue folding is consistent with a release mechanism in which the intermediate is bound to BamA until folding has progressed to the point where the N- and C-termini are proximal. At this point, substrate β-barrel closure is promoted by interactions of the N- and C-terminal β-strands with loop 4, and this results in the breaking of interactions with the interior wall of BamA (**Figure 2.6d**).

2.2.5. A slow folding LptD mutant is rescued by a compensatory mutation in the interior wall of BamA

We sought to test our model for BamA binding and release of substrates through rational design of slow-folding mutants and identification of compensatory mutations that rescue these folding defects. Because we have proposed that L4 plays a crucial role in stabilizing the folded substrate, we deleted a single amino acid at the start of loop 4 (D330) with the expectation that this change would alter the disposition of the loop. Because the LptD$^{\Delta D330}$ substrate contains only a single amino acid deletion, it is more similar to wild-type LptD than the LptD4213 variant with its 23 amino acid deletion. Nevertheless, if the L4 loop is important for release because it stabilizes the folded β-barrel, then changing its orientation should affect folding.

We expressed either wild-type LptD, LptD4213, or LptD$^{\Delta D330}$ from a plasmid in otherwise wild-type *E. coli* and plated the strains on media with or without vancomycin to probe for changes

in outer membrane integrity. Because LptD forms the translocon that delivers lipopolysaccharide to the cell surface, defects in LptD assembly result in outer membrane defects that allow the entry of antibiotics that otherwise are ineffective against *E. coli*[23,71]. We found that expression of LptD$^{\Delta D330}$ increased susceptibility to vancomycin, consistent with impaired LptD assembly (**Figure 2.7a, top three lanes**). Since these strains also contain a second functional copy of LptD, the increased permeability conferred by LptD4213 and LptD$^{\Delta D330}$ is not a result of a lack of functional LptD translocons, but from a pore formed by the defective substrate. Next, we probed residence time of LptD$^{\Delta D330}$ on the Bam complex using crosslinking and observed that this variant formed strong crosslinks from the N-terminus of LptD$^{\Delta D330}$ to BamA, similar to LptD4213 and distinct from the wild-type LptD (**Figure 2.7b, left panel**). Therefore, LptD$^{\Delta D330}$, like LptD4213, has a longer residence time on BamA than wild-type LptD.

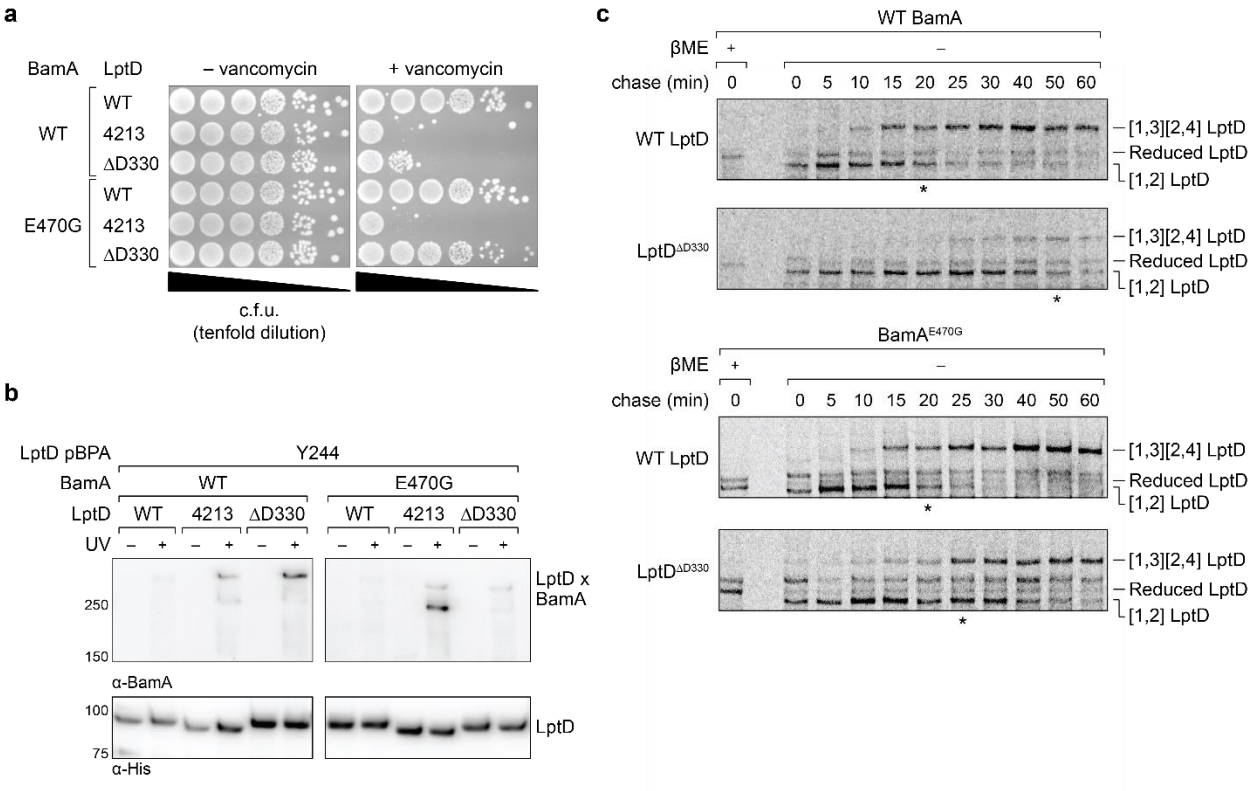

Figure 2.7. An assembly-defective LptD mutant is rescued by a compensatory mutation in the BamA interior wall

a, Expression of LptD$^{\Delta D330}$ confers outer membrane permeability defects, while changes in BamA can suppress LptD$^{\Delta D330}$ associated-defects. MC4100 or *bamA^{E470G}* cells were transformed with plasmids that express WT or mutant lptD alleles. Plating assays were performed on LB supplemented with 50 μg/mL

vancomycin. **b**, LptD$^{\Delta D330}$, like LptD4213, stalls on BamA during assembly. BamA^{E470G} alleviates stalling of LptD$^{\Delta D330}$, as judged by a reduction in crosslinking efficiency, but does not alleviate stalling of LptD4213. **c**, LptD$^{\Delta D330}$ is slow to mature into the functional disulfide bond configuration, while BamA^{E470G} alleviates LptD$^{\Delta D330}$ assembly defects. MC4100 or bamA^{E470G} cells expressing FLAG-tagged LptD(WT/ΔD330) were pulsed with [^{35}S]methionine and chased with cold methionine. Samples were subsequently immunoprecipitated using α-FLAG beads and analyzed by autoradiography. The asterisk below each autoradiograph represents the time point at which approximately 50% of the substrate has converted to the mature form (containing the [1,3][2,4] disulfide bond configuration). ***The plating assay, immunoblots, and autoradiographs presented in this figure were obtained by Dr. Thiago Santos.***

To evaluate the effect of the deletion of D330 on assembly kinetics, we pulse-labeled cells with [^{35}S]methionine and monitored the oxidation status of LptD over time[33]. In cells expressing wild-type LptD, 50% of the substrate is converted to the mature disulfide bonded state within 20 min. In cells expressing LptD$^{\Delta D330}$, we found that the LptD$^{\Delta D330}$ substrate folded more slowly than WT LptD, as judged by a slower conversion to mature LptD with its native disulfide bond configuration ([1,3][2,4] LptD) (**Figure 2.7c**, top two panels).

Finally, we sought to identify additional mutations that could compensate for the folding defect in LptD$^{\Delta D330}$. We have shown that an intragenic suppressor that increases the hydrophobicity of the N-terminus of substrate LptD can influence the affinity of LptD to BamA and facilitate substrate release. If this is true, we predicted that mutations in the interior wall of BamA that weaken interactions with substrate would also facilitate substrate release. Therefore, we screened positions in the interior wall of BamA that we showed contact substrate to identify a suppressor that would allow for more rapid release of the LptD$^{\Delta D330}$ substrate. We found that a single substitution in the interior of BamA, E470G, rescued the permeability barrier defects caused by LptD$^{\Delta D330}$ (**Figure 2.7, right bottom three lanes**). Moreover, BamA^{E470G} had a more transient association with LptD$^{\Delta D330}$ than with LptD4213 as judged by loss of crosslinking (**Figure 2.7b, right panel**). BamA^{E470G} did not compromise the ability of LptD or the LptD$^{\Delta D330}$ mutant to achieve the mature disulfide bond configuration (i.e., properly oxidized product) (**Figure 2.7c**). Finally, pulse-chase analysis showed that LptD$^{\Delta D330}$ was folded faster by BamA^{E470G} ($t_{1/2}$=25 min) than by wild-type BamA ($t_{1/2}$=50 min), and its folding by BamA^{E470G} was comparable to that of wild-type LptD by wild-type BamA (**Figure 2.7c**). The rate of folding of wild-type LptD was similar by wild-

type BamA and by BamA^{E470G}. These results confirm that interactions between substrate and the interior wall of BamA are important for folding and that changes in these interactions affect an important step in folding, which is release from the interior wall of BamA.

2.2.6. Structural studies of the Bam complex bound to the stalled substrate LptD4213

Given that we could biochemically characterize the interactions between LptD substrates and BamA, we sought to determine a structure of a substrate-bound Bam complex to gain further insight into the mechanism of folding. The biochemical results presented here suggest that a structure might reveal that LptD4213 interacts with the Bam complex such that the assembling N-terminal portion of its β-sheet is housed within the interior of the BamA β-barrel. We first developed an approach to purify the machine bound to LptD4213. This strategy involved expressing in *E. coli* all five components of the Bam complex (with His-tagged BamE), the lipoprotein LptE, Twin-Strep-tagged LptD4213, and the chaperone SurA. To prevent dissociation of the substrate during purification, the chemical crosslinker dithiobis(succinimidyl propionate) (DSP) was added after the induction period for protein overexpression. This crosslinker has a 12.0 Å spacer arm that bridges primary amines. We reasoned that using DSP would allow for stable association of the substrate on the Bam complex but could also restrict conformational flexibility in the substrate that may hinder structure determination.

After crosslinking, the substrate-bound Bam complex was purified in the detergent dodecyl maltoside (DDM) through standard protocols using the two affinity tags we introduced. Size-exclusion chromatography (**Figure 2.8a**), SDS-PAGE (**Figure 2.8b**), and negative-stain electron microscopy (**Figure 2.8c**) demonstrated that the crosslinking and purification approach was promising. Importantly, all protein components could be identified by SDS-PAGE without excessive contamination (**Figure 2.8b**). However, the yield of the purified complex was relatively low (approximately 0.02 milligrams of purified protein per liter of bacterial culture), preventing attempts at crystallization.

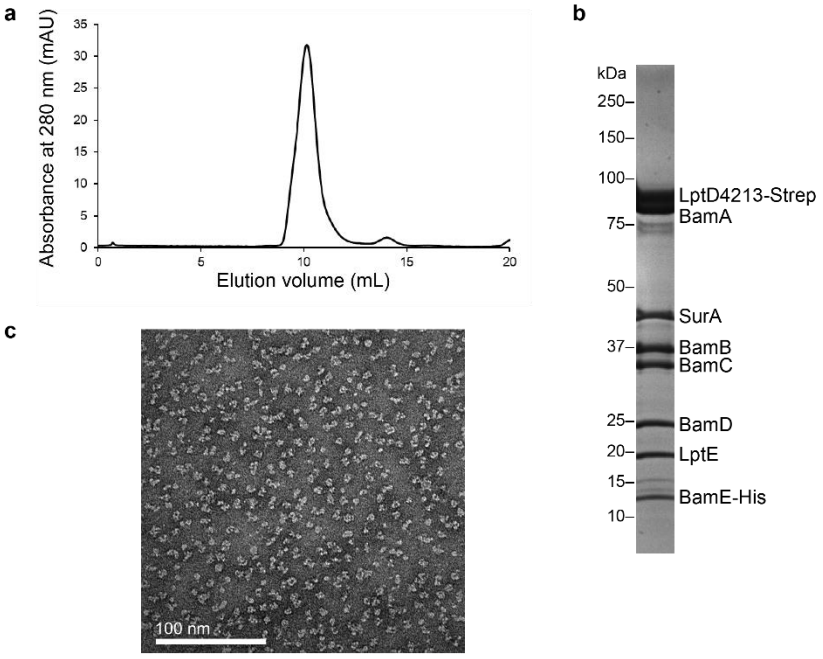

Figure 2.8. Purification of the Bam complex bound to LptD4213
a, Size-exclusion chromatography of the substrate-bound Bam complex in the detergent dodecyl maltoside. **b**, Purity of the sample shown by SDS-PAGE. **c**, Representative negative-stain electron micrograph shows relatively homogeneous particles.

Having been able to purify a substrate-bound Bam complex, we initiated cryo-electron microscopy experiments, working closely with Zongli Li at the Harvard Cryo-EM Center for Structural Biology at Harvard Medical School. We collected cryo-EM data and obtained a total of about 3,000 movies. From those, 2D classification and 3D classification protocols yielded six 3D classes (**Figure 2.9a**). Unfortunately, these 3D reconstructions were not of sufficient quality to confidently identify any components of the complex. However, the size of each 3D reconstruction was reasonable given the size of the apo state of the Bam complex (**Figure 2.9b**). Additionally, a region that may correspond to the location of the membrane portion embedded in the detergent micelle (represented by dotted lines) is visible in some reconstructions.

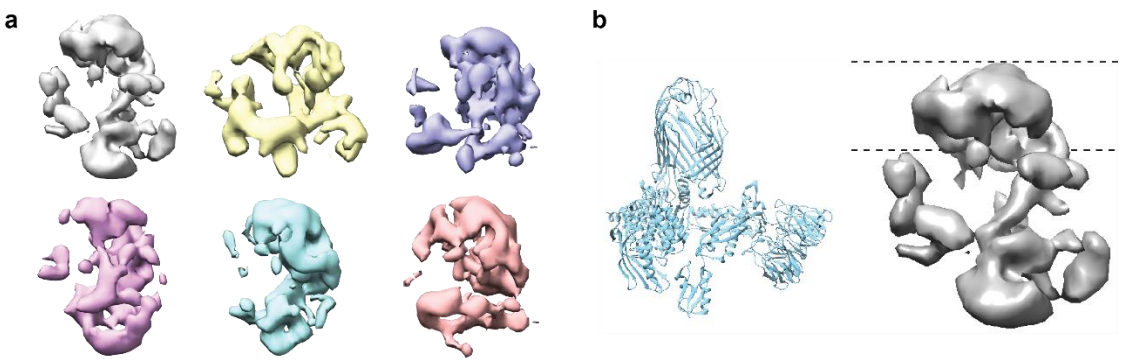

Figure 2.9. 3D classes of the substrate-bound Bam complex generated by cryo-EM
a, Six different 3D classes were obtained. **b**, An enlarged view of the gray reconstruction in **a** presented next to the Bam complex (PDB: 5LJO). A belt-like density corresponding to the membrane region may be present (represented by dotted lines).

2.3. Discussion

Here, we present evidence for how the Bam complex catalyzes folding of a β-barrel substrate. The main features of the model are that the C-terminal strand of the substrate is held by BamA while the N-terminal strands of the substrate fold inside the BamA barrel. In our model the concave interior wall of BamA serves as an active site for formation of β-strands. At a late stage of assembly release of the completed β-sheet from the interior wall allows the substrate N-terminus to pair with the substrate C-terminus and close the β-barrel. Substrate β-barrel closure is further assisted by intramolecular interactions within the substrate, with substrate loop residues playing a critical role by interacting to bring both ends of the β-barrel together. In support of this model, we have identified a large substrate-binding surface that encompasses the interior wall of BamA. The periodicity in substrate residues that crosslink to BamA implies that folding intermediates have substantial β-sheet structure when bound to the interior surface of BamA. We also found that changes to residues in the substrate that contact the BamA interior or changes to residues in the BamA interior surface itself that contact substrate can increase or decrease the rate of folding. It follows that the interior surface of the BamA barrel wall is the catalyst for β-barrel strand formation.

An interesting aspect of our folding model is that the BamA interior surface is relatively polar[133], and yet it promotes the folding of β-barrel substrates with a hydrophobic exterior by forming extensive contacts with the hydrophobic surface of the growing substrate β-sheet. We argue that this type of association makes sense in the context of understanding the kinetics of β-strand formation. First, the polar nature of the interior wall is likely necessary because the interior surface is exposed to water for at least some of the time and the cavity would not be stable if it was too nonpolar. Second, and more importantly, an extensively hydrophobic BamA interior surface would interfere with release of a hydrophobic β-sheet. That is, to ensure that substrate binding is not too strong as tight binding of a substrate would impede the rate of β-strand formation. Indeed, weak oriented binding of substrates is a hallmark of enzymatic catalysis. Third, release of bound water from the hydrophobic exterior surface of the substrate as it folds against the interior surface is likely the driving force for folding. Finally, the leading exposed edge of the β-sheet in the folding substrate contains unsatisfied hydrogen bonds and is therefore polar; a relatively polar interior may be required to stabilize this leading edge. Therefore, we propose that the apparent mismatch in polarity between the associating surfaces is in fact central to the enzymatic mechanism for how β-strand formation is catalyzed.

How does the interior wall accelerate the folding of LptD and other substrates by nucleating the formation of β-strands against the BamA interior wall? The barrier to formation of β-sheet structure involves formation of an extended, entropically disfavored conformation of the peptide chain. However, once one β-hairpin forms, successive addition of more β-hairpins is facilitated in two ways. First, the leading edge of the β-sheet serves as a preorganized template facilitating the simultaneous formation of multiple hydrogen bonds. Second, each new β-strand has the appropriate hydrophobic periodicity of side chains to benefit from the hydrophobic effect as it packs against the interior wall. In short, confining substrates within a cage overcomes the entropic barrier that would normally slow folding because successive peptide β-strands do not

have to sample as large a conformational space as they would in free solution, and moreover the new β-strands gain stability from their contacts to the cage itself.

Our model stands in contrast to the predominant model in the field of membrane β-barrel folding known as the budding model, which involves sequential insertion of individual β-hairpins at the seam of the BamA barrel[135,144]. This earlier model raises two important mechanistic problems. First, models in which hairpins are continuously added by forming hydrogen-bonding networks at the seam imply a substantial kinetic barrier for addition of each hairpin because the existing hydrogen-bonding networks must be disrupted to insert new hairpins. In contrast, in our model, the interior surface serves as a catalyst for β-strand formation without invoking continuous β-strand exchange with the machine. A second problem posed by the budding model is that, as the barrel grows, the N- and C-termini are displaced farther and farther apart, and this presents a problem for how barrel closure is achieved. No such problem arises in our model because confining the N-terminus of the growing sheet within the cage allows the ends of the substrate to remain close to each other as folding progresses.

Finally, we note that the mechanism that we have proposed for β-barrel folding by the Bam complex resembles the mechanism by which the GroEL/GroES chaperone system accelerates the folding of soluble proteins[170,171]. The GroEL/GroES chaperonin uses ATP hydrolysis to drive a series of conformational changes that open and close the cavity to promote binding and release of folding intermediates[172]. Like GroEL/GroES, we propose that BamA uses a hydrophilic cage to limit the conformational space substrates can sample. Unlike GroEL/GroES, ATP is not required to drive conformational changes in BamA between its open and closed forms because the BamA β-barrel can access the open form without energy[108,109,111]. Thus, the Bam complex accelerates the assembly of membrane β-barrel proteins by confining segments of folding substrates within the open BamA β-barrel to reduce the entropic cost of folding.

It is promising that the cryo-EM reconstructions we obtained are similar in size to what would be expected for a substrate-bound Bam complex. However, the overall quality of the

reconstructions is low, which is likely due to sample heterogeneity. One possibility for this is that the stalled substrate may be actually exist as a set of states in equilibrium rather than a single, static state. A related problem may be heterogeneity due to the use of the crosslinker (DSP) that generates crosslinks primarily between two lysine side chains. It is unclear where these crosslinks are being formed, and it is possible that the crosslinker stabilizes an undesired state that happens to have two lysine residues in proximity rather than the stalled state captured by photocrosslinking and cysteine crosslinking. Additional work is thus required to capture the substrate on the machine in a single state to enable structural studies.

Based on these results, biochemical and structural studies with a different substrate (BamA) were pursued to enable structural studies, as described in Chapters 3 and 4.

2.4. Materials and methods

2.4.1. Bacterial growth conditions

Unless otherwise noted, cultures were grown at 37°C and supplemented with the appropriate antibiotics and amino acids. Lysogeny broth (LB) and agar were prepared as described previously[129]. When appropriate, carbenicillin (50 μg/ml), chloramphenicol (30 μg/ml), and kanamycin (50 μg/ml) were used. para-Benzoylphenylalanine (pBPA; Bachem Americas) was used at 0.9 mM.

2.4.2. Bacterial growth conditions

Cloning of the mutant gene *bamA*^{E470G} into pDS132 was performed in *E. coli* DH5α λpir. The resulting plasmid, pDS132::*bamA*^{E470G}, was purified and transformed into *E. coli* MC4100 for allelic exchange. Cells of the recipient strain were plated on LB agar supplemented with chloramphenicol to select transformants that integrated the plasmid into their chromosome. Following overnight growth at 37°C, one colony was inoculated into LB, incubated at 37°C for 4

hr, diluted in 1× PBS, and plated on LB agar supplemented with 5% sucrose without NaCl. This step allowed selection of cells in which the integrated plasmid was excised from their chromosome. After overnight incubation at 37°C, about 50 colonies were streaked onto LB agar supplemented with chloramphenicol and on LB agar supplemented with 5% sucrose without NaCl. Clones that were resistant to sucrose and susceptible to chloramphenicol were screened by PCR and sequencing of the *bamA* locus.

2.4.3. Plasmid construction

Plasmids were constructed using traditional cloning methods and Gibson assembly. Vectors and insert DNA were generated by PCR with KOD DNA polymerase (Toyobo) and treated with DpnI (NEB). Constructs were initially transformed into NovaBlue competent cells (Sigma) by heat shock. All plasmids were verified by Sanger sequencing.

2.4.4. Site-specific *in vivo* photocrosslinking

Photocrosslinking experiments are based on techniques as previously described[29], with modifications. MC4100 strains harboring pSup-BpaRS-6TRN and pZS21*lptD*-His or pZS21His-*bamA* containing the TAG stop codon at the indicated positions were grown overnight, diluted 1:100 into 100 mL of the same media and grown to midlog phase. After normalization by optical density, each culture was split in half and directly irradiated with UV light at 365 nm for 10 min at room temperature. All samples were subsequently kept at 4°C. Samples were resuspended in 5 mL ice-cold TBS containing 1% Anzergent 3–14 (Anatrace), 100 µg/mL lysozyme, 1 mM PMSF, and 50 µg/mL DNase I, lysed by sonication, and centrifuged at 15,000 × g in a table-top centrifuge for 10 min. The supernatant was then passaged three times over Ni-NTA beads and then washed twice with 5 mL ice-cold TBS containing 0.02% Anzergent 3–14, and 20 mM imidazole. Samples were eluted with 1 mL ice-cold TBS containing 0.02% Anzergent 3–14 and 200 mM imidazole. Eluates were supplemented with 10% TCA by volume (100 µl) and incubated on ice for 30 min.

Precipitated proteins were pelleted at 18,000 x g for 10 min at 4°C. All samples were resuspended in 50 µL of SDS-PAGE buffer and incubated at 95°C for 10 min. 8 µL of each sample were separated on 4–8% SDS-PAGE gels and analyzed by immunoblotting.

2.4.5. *In vivo* BMOE chemical crosslinking

6×His BamA variants containing the S439C or N666C mutation were cloned into the pZS21 vector. Substrates with a C-terminal 3×FLAG tag (LptD4213 containing an N- or C-terminal cysteine) were cloned into the pTrc99a vector. MC4100 cells were transformed with one BamA-encoding plasmid and one substrate-encoding plasmid. The resulting strains were grown overnight in LB supplemented with 50 µg/mL carbenicillin, 50 µg/mL kanamycin, and 0.2% (w/v) glucose (37°C, 220 rpm). These overnight cultures were diluted 1:100 into 100 mL of fresh LB containing the same additives without glucose, and were grown (37°C, 220 rpm) to OD600 ~ 0.5. Cells were then collected by centrifugation (4200 x g, 10 min, 4°C). Cell pellets were resuspended in PBS (20 mM NaH_2PO_4 pH 7.2, 150 mM NaCl). TCEP-HCl (VWR) was then added at a final concentration of 2 mM, and cells were incubated on a rocking platform (20 min, room temperature). Cells were then centrifuged (5000 x g, 10 min, 4°C) and again resuspended in PBS. The cysteine-to-cysteine crosslinker 1,4-bis(maleimido)butane (BMB, Thermo Fisher Scientific) was added at a final concentration of 0.5 mM. After incubation on a rocking platform (40 min, room temperature), the crosslinking reaction was quenched via addition of L-cysteine hydrochloride monohydrate (Alfa Aesar) to a final concentration of 10 mM. Cells were centrifuged (5000 x g, 10 min, 4°C) and the pellets were frozen at −80°C prior to subsequent purification. In each sample, 6×His BamA (and any associated substrate) was purified, and Ni-NTA elutions were subjected to SDS PAGE and subsequent Western blotting. 6×His BamA was detected by using a penta-His (HRP) antibody (Qiagen). LptD4213−3×FLAG (substrate) that was pulled down with 6×His BamA was detected by using a monoclonal anti-FLAG M2-peroxidase (HRP) antibody (Sigma-Aldrich).

2.4.6. Analysis of antibiotic sensitivities

Plating of the strains was performed as previously described[127]. All strains were grown at 37°C to an OD600 of approximately 0.8. Cells were normalized to an OD600 of 0.1 and then subject to five serial 10-fold dilutions. 5 µL of the dilution series were plated on agar plates containing the indicated additive and incubated at 37°C for 18–20 hr.

2.4.7. Analysis of cellular protein levels

All strains were grown at 37°C to an OD600 ~0.5. The cells from a 1 mL sample were normalized to an OD600 of 0.3 and were collected by centrifugation at 10,000 x g for 10 min. The resulting cell pellets were resuspended in 100 µL of 1× SDS-sample buffer (+β-mercaptoethanol, β-ME) and incubated at 95°C for 10 min. The samples were separated on 4–20% SDS-PAGE gels and analyzed via immunoblotting. To analyze trimer assembly, samples were not boiled and were directly subject to SDS-PAGE.

2.4.8. Immunoblotting

Proteins were transferred from Tris-glycine polyacrylamide gels to PVDF membranes (Bio-Rad) for 15 min at a constant voltage of 25 V. Membranes were blocked with casein blocking buffer (Sigma) for 1 hr and incubated with primary antibodies in blocking buffer at 4°C overnight. Membranes were washed 3 times with TBST buffer (10 mM Tris·HCl pH 8.0, 150 mM NaCl, and 0.05% Tween-20), incubated with HRP-conjugated secondary antibodies for 1 hr at room temperature, and again washed 3 times with TBST buffer. Signal was detected using an Azure C400 imager (Azure Biosystems).

2.4.9. Pulse-chase analysis

Pulse-chase experiments were performed as previously described[33]. Briefly, a 5 mL culture was grown to an OD600 of ~0.5 in M63 minimal media supplemented with 18 amino acids

(minus methionine and cysteine) at 37°C. The culture was pulse-labeled with [^{35}S]methionine (100 μCi/mL final concentration) (American Radiolabeled Chemicals) for 2 min and then chased with cold methionine (5 mM) at 37°C. At the indicated time point during the chase, an 800 μL culture aliquot was transferred to a 1.5 mL tube containing 80 μL of TCA (70% in water) and incubated on ice for 20 min. Precipitated proteins were pelleted at 18,000 × g for 10 min at 4°C, washed with 700 μL ice- cold acetone, and then solubilized in 80 μL 100 mM Tris·HCl, pH 8.0, containing 1% SDS and 20 mM N-ethylmaleimide (Sigma). The sample was sonicated for 30 s to aid solubilization. Following that, 800 μL of ice-cold immunoprecipitation (IP) buffer (50 mM Tris·HCl, pH 8.0, containing 150 mM NaCl, 2% Triton X-100, and 1 mM EDTA) was added and the sample was centrifuged at 18,000 × g for 10 min at 4°C. 700 μL of the supernatant was transferred to another 1.5 mL tube containing 2.5 μL of anti-FLAG M2 magnetic beads 4 (Sigma). The beads were washed and preequilibrated with 3 × 1 mL IP buffer before use. The mixture was incubated on a rotary shaker for 1 hr at 4°C, and the beads were washed three times with 800 μL of ice-cold high-salt buffer (50 mM Tris·HCl, pH 8.0, containing 1 M NaCl, 1% Triton X-100, and 1 mM EDTA) and one time with 800 μL ice-cold 10 mM Tris·HCl, pH 8.0, using a magnetic separation rack (New England Biolabs). Sixty microliters of 2× SDS nonreducing sample buffer was then added to the beads and the mixture heated for 10 min at 100°C to elute the bound proteins. Fifteen microliters of eluted sample was applied to SDS/PAGE directly. For reduction of disulfide bonds, 0.5 μL β-ME (Sigma) was added to 20 μL eluted sample and heated for 5 min at 100°C before loading. Tris·HCl polyacrylamide gels (4–20%) were used for SDS/PAGE analysis (running conditions: 150 V for 120 min). The gel was then dried and exposed to phosphor storage screens for autoradiography. Signals were detected with a Typhoon FLA 7000 image analyzer (GE Healthcare) using ImageQuant TL (GE Healthcare).

2.4.9. Expression and crosslinking of substrate-bound Bam complex for cryo-EM

The plasmid pJH114 was used to express the five components of the Bam complex. LptD4213-2×Strep and SurA were cloned into the pBAD33 vector to generate pDT167. LptE was cloned into the pCL vector to generate pDT015. BL21(DE3) cells harboring pJH114, pDT015, and pDT167 were grown overnight (37°C, 220 rpm) in LB supplemented with 50 µg/mL carbenicillin, 30 µg/mL chloramphenicol, 50 µg/mL spectinomycin, and 0.2% glucose. This overnight culture was diluted 1:100 into 6 x 1.5 L of LB supplemented with 50 µg/mL carbenicillin and 30 µg/mL chloramphenicol. The resulting cultures were grown (37°C, 220 rpm) until an OD600 of ~0.7 was reached. At this point, the temperature was turned down to 30°C, and cells were allowed to continue shaking. After 20 min, protein expression was induced with isopropyl β-D-1-thiogalactopyranoside (IPTG, VWR) and L-(+)-arabinose (Alfa Aesar) at final concentrations of 0.2 mM and 0.1% (w/v), respectively. After three hours of additional shaking, cells were harvested via centrifugation (4,200 x g, 10 min, 4°C). Cell pellets were resuspended in PBS (20 mM NaH_2PO_4 pH 7.2, 150 mM NaCl) and incubated with 0.5 mM dithiobis(succinimidyl propionate) (DSP) for 45 minutes. Cells were centrifuged (5,000 x g, 10 min, 4°C) and the pellets were frozen at -80°C prior to subsequent purification.

2.4.10. Purification of substrate-bound Bam complex

Cell pellets expressing Bam-substrate complex as described above were thawed and resuspended in buffer containing 20 mM Tris-HCl pH 8.0, 150 mM NaCl, 100 µg/mL lysozyme (Sigma-Aldrich), 1 mM PMSF (Sigma-Aldrich), 50 µg/mL DNase I (Sigma-Aldrich), and 2.5 mM $MgCl_2$ (Sigma-Aldrich). Cells were lysed using an Emulsiflex C3 (Avestin) at a pressure of 10,000 to 15,000 psi. After lysis, cell debris was removed via centrifugation (5,000 x g, 10 min, 4°C). Membrane fractions were isolated via ultracentrifugation using a 45 Ti rotor (Beckman Coulter) (37,000 rpm, 45 min, 4°C) and an Optima XE-90 ultracentrifuge (Beckman Coulter). The membrane pellet was resuspended in buffer containing 20 mM Tris-HCl pH 8.0, 150 mM NaCl,

100 µg/mL lysozyme, and 1 mM PMSF. Membrane fractions were solubilized via incubation with 1% n-dodecyl-β-D-maltopyranoside (DDM, Anatrace) on a rocking platform (2 hours, 4°C). Unsolubilized material was then isolated via ultracentrifugation in a 70 Ti rotor (Beckman Coulter) (37,000 rpm, 30 min, 4°C). The supernatant, consisting of solubilized membrane proteins, was removed and supplemented with imidazole (pH 8.0) to a final concentration of 5 mM.

The supernatant was incubated with Ni-nitrilotriacetic acid (NTA) resin (Qiagen) that had been pre-washed with 10 CV buffer W1 (20 mM Tris-HCl pH 8.0, 150 mM NaCl, 10 mM imidazole pH 8.0, 0.05% DDM). After batch binding on a rocking platform (1 hour, 4°C), the resin was washed with 10 CV buffer W1. Elution was performed via addition of 5 CV buffer E1 (20 mM Tris-HCl pH 8.0, 150 mM NaCl, 200 mM imidazole pH 8.0, 0.05% DDM).

The Ni-NTA eluate was immediately incubated with Strep-Tactin XT Superflow resin (IBA Lifesciences) that had been pre-washed with 10 CV buffer W2 (100 mM Tris-HCl pH 8.0, 150 mM NaCl, 0.05% DDM). After batch binding on a rocking platform (1 hour, 4°C), the resin was washed with 15 CV buffer W2. Elution was performed via addition of 9 CV buffer E2 (100 mM Tris-HCl pH 8.0, 150 mM NaCl, 1 mM EDTA, 50 mM D-biotin, 0.05% DDM).

The Strep resin eluate was concentrated using an Amicon Ultra 4 mL 100 kDa molecular-weight cutoff centrifugal concentrator (EMD Millipore). The sample was then applied to an ÄKTA Pure (GE Healthcare Life Sciences) for purification via size-exclusion chromatography using a Superdex 200 Increase 10/300 GL column. The protein was eluted in buffer containing 20 mM Tris-HCl pH 8.0, 150 mM NaCl, and 0.05% DDM. After elution, protein corresponding to the center peaks of the chromatogram was concentrated to 5 mg/mL using an Amicon Ultra 0.5 mL 100 kDa molecular-weight cutoff centrifugal concentrator (EMD Millipore). A final yield of approximately 0.02 mg of complex per liter of bacterial culture could be obtained.

2.4.11. Electron microscopy data collection, image processing, and 3D reconstruction

Protein concentrated to 7 mg/mL was applied to cryo-EM grids that were then frozen. Electron microscopy data collection, image processing, and 3D reconstruction generation were performed in a manner similar to that described in sections 4.4.2 and 4.4.3 below.

2.4.12. Bacterial strains

Bacterial strains used in the experiments reported here are provided in Table 2.1.

Table 2.1. Bacterial strains used in chapter 2

Strains	Genotype	Source
MC4100	F^- *araD139* Δ*(argF-lac)U169 rpsL150 relA1 flbB5301 deoC1 ptsF25 rbsR thi*	[173]
NR698	MC4100 *lptD4213 (*Δ330-352)	[73]
NR1134	NR754 Δ*lptD::kan* pACYC184::*lptD* (*lptD* depletion strain)	[129]
JCM166	MC4100 *ara*$^{r/-}$ Δ(*λatt-lom*)::*bla* P_{BAD}*yaeT araC* Δ*yaeT*	[73]
DEK1	MC4100, *bamA*^{E470G} (*bamA*^{E470G} haploid strain)	[174]
DH5α λ*pir*	*supE4* Δ*lacU169* (φ80 *lacZ*ΔM15) *hsdR17 recA1 endA1 gyrA96 thi-1 relA1* λ*pir*	[175]
NovaBlue	*endA1 hsd*R17 (r_K^-, m_K^+) *sup*E44 *thi*-1 *rec*A1 *gyr*A96 *rel*A1 *lac* F′[*pro*A^+ B^+ *lac*I^q ZΔM15::Tn*10*]	Novagen
BL21(DE3)	*fhuA2 [lon] ompT gal (*λ *DE3) [dcm]* Δ*hsdS*	Novagen

2.4.13. Bacterial plasmids

Bacterial plasmids used in the experiments here are provided in Table 2.2.

Table 2.2. Bacterial plasmids used in chapter 2

Plasmid	Description	Source
pZS21	P_{Ltet}-dependent expression vector	[176]
pZS21::*His-bamA*	Encodes full-length BamA with an N-terminal His_8 tag	[118]
pZS21::*His-bamA*^{G655Am}	pZS21::*His-bamA* with G655Amber	[174]
pZS21::*His-bamA*^{S657Am}	pZS21::*His-bamA* with S657Amber	[174]
pZS21::*His-bamA*^{T659Am}	pZS21::*His-bamA* with G659Amber	[174]
pZS21::*His-bamA*^{N666Am}	pZS21::*His-bamA* with N666Amber	[174]
pZS21::*His-bamA*^{I668Am}	pZS21::*His-bamA* with I668Amber	[174]
pZS21::*His-bamA*^{N427Am}	pZS21::*His-bamA* with N427Amber	[174]
pZS21::*His-bamA*^{Y432Am}	pZS21::*His-bamA* with Y432Amber	[174]
pZS21::*His-bamA*^{T434Am}	pZS21::*His-bamA* with T434Amber	[174]
pZS21::*His-bamA*^{E435Am}	pZS21::*His-bamA* with E435Amber	[174]
pZS21::*His-bamA*^{S439Am}	pZS21::*His-bamA* with S439Amber	[174]
pZS21::*His-bamA*^{Q441Am}	pZS21::*His-bamA* with Q441Amber	[174]
pZS21::*His-bamA*^{Y468Am}	pZS21::*His-bamA* with Y468Amber	[174]

pZS21::*His-bamA*^{E470Am}	pZS21::*His-bamA* with E470Amber	174
pZS21::*His-bamA*^{L471Am}	pZS21::*His-bamA* with L471Amber	174
pZS21::*His-bamA*^{R488Am}	pZS21::*His-bamA* with R488Amber	174
pZS21::*His-bamA*^{F490Am}	pZS21::*His-bamA* with F490Amber	174
pZS21::*His-bamA*^{N492Am}	pZS21::*His-bamA* with N492Amber	174
pZS21::*His-bamA*^{F494Am}	pZS21::*His-bamA* with F494Amber	174
pZS21::*His-bamA*^{N501Am}	pZS21::*His-bamA* with N501Amber	174
pZS21::*His-bamA*^{T511Am}	pZS21::*His-bamA* with T511Amber	174
pZS21::*His-bamA*^{G528Am}	pZS21::*His-bamA* with G528Amber	174
pZS21::*His-bamA*^{G530Am}	pZS21::*His-bamA* with G530Amber	174
pZS21::*His-bamA*^{D569Am}	pZS21::*His-bamA* with D569Amber	174
pZS21::*His-bamA*^{K610Am}	pZS21::*His-bamA* with K610Amber	174
pZS21::*His-bamA*^{D614Am}	pZS21::*His-bamA* with D614Amber	174
pZS21::*His-bamA*^{T615Am}	pZS21::*His-bamA* with T615Amber	174
pZS21::*His-bamA*^{G655Am}	pZS21::*His-bamA* with G655Amber	174
pZS21::*His-bamA*^{S657Am}	pZS21::*His-bamA* with S657Amber	174
pZS21::*His-bamA*^{T659Am}	pZS21::*His-bamA* with G659Amber	174
pZS21::*His-bamA*^{N666Am}	pZS21::*His-bamA* with N666Amber	174
pZS21::*His-bamA*^{I668Am}	pZS21::*His-bamA* with I668Amber	174
pZS21::*His-bamA*^{Q803Am}	pZS21::*His-bamA* with Q803Amber	174
pZS21::*His-bamA*^{F804Am}	pZS21::*His-bamA* with F804Amber	174
pZS21::*His-bamA*^{N805Am}	pZS21::*His-bamA* with N805Amber	174
pZS21::*His-BamA*^{S439C}	pZS21::*His-bamA* with S439C	174
pZS21::*His-BamA*^{N666C}	pZS21::*His-bamA* with N666C	174
pTrc99a	$P_{trp/lac}$-dependent expression vector	177
pTrc99a::*lptD4213-FLAG*	Encodes LptD4213 with a C-terminal $FLAG_3$ tag	174
pTrc99a::*lptD4213-FLAG*^{L245C}	pTrc99a::*lptD4213-FLAG* with L245C	174
pTrc99a::*lptD4213-FLAG*^{H262C}	pTrc99a::*lptD4213-FLAG* with H262C	174
pTrc99a::*lptD4213-FLAG*^{Y291C}	pTrc99a::*lptD4213-FLAG* with Y291C	174
pTrc99a::*lptD4213-FLAG*^{A709C}	pTrc99a::*lptD4213-FLAG* with A709C	174
pTrc99a::*lptD4213-FLAG*^{E733C}	pTrc99a::*lptD4213-FLAG* with E733C	174
pTrc99a::*lptD4213-FLAG*^{N737C}	pTrc99a::*lptD4213-FLAG* with N737C	174
pZS21::*lptD-His*	Encodes full-length LptD with a C-terminal His_8 tag	129
pZS21::*lptD-His*^{A233Am}	pZS21::*lptD-His* with A233Amber	174
pZS21::*lptD-His*^{K234Am}	pZS21::*lptD-His* with K234Amber	174
pZS21::*lptD-His*^{Y235Am}	pZS21::*lptD-His* with Y235Amber	174
pZS21::*lptD-His*^{T237Am}	pZS21::*lptD-His* with T237Amber	174
pZS21::*lptD-His*^{T238Am}	pZS21::*lptD-His* with T238Amber	174
pZS21::*lptD-His*^{Y240Am}	pZS21::*lptD-His* with Y240Amber	174
pZS21::*lptD-His*^{Y244Am}	pZS21::*lptD-His* with Y244Amber	174
pZS21::*lptD-His*^{Y248Am}	pZS21::*lptD-His* with Y248Amber	174
pZS21::*lptD-His*^{I259Am}	pZS21::*lptD-His* with I259Amber	174
pZS21::*lptD-His*^{T260Am}	pZS21::*lptD-His* with T260Amber	174

pZS21::*lptD-His*^{P261Am}	pZS21::*lptD-His* with P261Amber	174
pZS21::*lptD-His*^{H262Am}	pZS21::*lptD-His* with H262Amber	174
pZS21::*lptD-His*^{Y263Am}	pZS21::*lptD-His* with Y263Amber	174
pZS21::*lptD-His*^{W272Am}	pZS21::*lptD-His* with W272Amber	174
pZS21::*lptD-His*^{E273Am}	pZS21::*lptD-His* with E273Amber	174
pZS21::*lptD-His*^{N274Am}	pZS21::*lptD-His* with N274Amber	174
pZS21::*lptD-His*^{E275Am}	pZS21::*lptD-His* with E275Amber	174
pZS21::*lptD-His*^{F276Am}	pZS21::*lptD-His* with F276Amber	174
pZS21::*lptD-His*^{L289Am}	pZS21::*lptD-His* with L289Amber	174
pZS21::*lptD-His*^{D290Am}	pZS21::*lptD-His* with L290Amber	174
pZS21::*lptD-His*^{Y291Am}	pZS21::*lptD-His* with Y291Amber	174
pZS21::*lptD-His*^{W311Am}	pZS21::*lptD-His* with W311Amber	174
pZS21::*lptD-His*^{L312Am}	pZS21::*lptD-His* with L312Amber	174
pZS21::*lptD-His*^{F313Am}	pZS21::*lptD-His* with F313Amber	174
pZS21::*lptD-His*^{Y314Am}	pZS21::*lptD-His* with Y314Amber	174
pZS21::*lptD-His*^{W315Am}	pZS21::l*ptD-His* with W315Amber	174
pZS21::*lptD-His*^{Y678Am}	pZS21::*lptD-His* with Y678Amber	174
pZS21::*lptD-His*^{Q684Am}	pZS21::*lptD-His* with Q684Amber	174
pZS21::*lptD-His*^{V685Am}	pZS21::*lptD-His* with V685Amber	174
pZS21::*lptD-His*^{Y704Am}	pZS21::*lptD-His* with Y704Amber	174
pZS21::*lptD-His*^{N708Am}	pZS21::*lptD-His* with N708Amber	174
pZS21::*lptD-His*^{M716Am}	pZS21::*lptD-His* with M716Amber	174
pZS21::*lptD-His*^{L717Am}	pZS21::*lptD-His* with L717Amber	129
pZS21::*lptD-His*^{G718Am}	pZS21::*lptD-His* with G718Amber	129
pZS21::*lptD-His*^{V719Am}	pZS21::*lptD-His* with V719Amber	174
pZS21::*lptD-His*^{Y721Am}	pZS21::*lptD-His* with Y721Amber	129
pZS21::*lptD-His*^{Y726Am}	pZS21::*lptD-His* with MY726Amber	129
pZS21::*lptD-His*^{I728Am}	pZS21::*lptD-His* with I728Amber	174
pZS21::*lptD-His*^{N737Am}	pZS21::*lptD-His* with N737Amber	174
pZS21::*lptD-His*^{W739Am}	pZS21::*lptD-His* with W739Amber	174
pZS21::*lptD-His*^{K743Am}	pZS21::*lptD-His* with K743Amber	174
pZS21::*lptD-His*^{V747Am}	pZS21::*lptD-His* with V747Amber	174
pZS21::*lptD-His*^{Y748Am}	pZS21::*lptD-His* with Y748Amber	174
pZS21::*lptD-His*^{D749Am}	pZS21::*lptD-His* with D749Amber	174
pZS21::*lptD-His*^{I752Am}	pZS21::*lptD-His* with I752Amber	174
pZS21::*lptD-His*^{F754Am}	pZS21::*lptD-His* with F754Amber	174
pZS21::*lptD-His*^{N755Am}	pZS21::*lptD-His* with N755Amber	174
pZS21::*lptD4213-His*	Encodes LptD4213 (Δ330-352) with a C-terminal His_8 tag	129
pZS21::*lptD4213-His*^{A233Am}	pZS21::*lptD4213-His* with A233Amber	174
pZS21::*lptD4213-His*^{K234Am}	pZS21::*lptD4213-His* with K234Amber	174
pZS21::*lptD4213-His*^{Y235Am}	pZS21::*lptD4213-His* with Y235Amber	174
pZS21::*lptD4213-His*^{T237Am}	pZS21::*lptD4213-His* with T237Amber	174
pZS21::*lptD4213-His*^{T238Am}	pZS21::*lptD4213-His* with T238Amber	174

pZS21::*lptD4213-His*^{Y240Am}	pZS21::*lptD4213-His* with Y240Amber	174
pZS21::*lptD4213-His*^{Y244Am}	pZS21::*lptD4213-His* with Y244Amber	174
pZS21::*lptD4213-His*^{Y248Am}	pZS21::*lptD4213-His* with Y248Amber	174
pZS21::*lptD4213-His*^{I259Am}	pZS21::*lptD4213-His* with I259Amber	174
pZS21::*lptD4213-His*^{T260Am}	pZS21::*lptD4213-His* with T260Amber	174
pZS21::*lptD4213-His*^{P261Am}	pZS21::*lptD4213-His* with P261Amber	174
pZS21::*lptD4213-His*^{H262m}	pZS21::*lptD4213-His* with H262Amber	174
pZS21::*lptD4213-His*^{Y263Am}	pZS21::*lptD4213-His* with Y263Amber	174
pZS21::l*ptD4213-His*^{R266Am}	pZS21::*lptD4213-His* with R266Amber	174
pZS21::*lptD4213-His*^{N269Am}	pZS21::*lptD4213-His* with N269Amber	174
pZS21::*lptD4213-His*^{W272Am}	pZS21::*lptD4213-His* with W272Amber	174
pZS21::*lptD4213-His*^{E273Am}	pZS21::*lptD4213-His* with E273Amber	174
pZS21::*lptD4213-His*^{N274Am}	pZS21::*lptD4213-His* with N274Amber	174
pZS21::*lptD4213-His*^{E275Am}	pZS21::*lptD4213-His* with E275Amber	174
pZS21::*lptD4213-His*^{F276Am}	pZS21::*lptD4213-His* with F276Amber	174
pZS21::l*ptD4213-His*^{L279Am}	pZS21::*lptD4213-His* with L279Amber	174
pZS21::l*ptD4213-His*^{L286Am}	pZS21::*lptD4213-His* with L286Amber	174
pZS21::*lptD4213-His*^{L289Am}	pZS21::*lptD4213-His* with L289Amber	174
pZS21::*lptD4213-His*^{D290Am}	pZS21::*lptD4213-His* with L290Amber	174
pZS21::*lpt4213D-His*^{Y291Am}	pZS21::*lptD4213-His* with Y291Amber	174
pZS21::*lptD4213-His*^{W311Am}	pZS21::*lptD4213-His* with W311Amber	174
pZS21::*lptD4213-His*^{L312Am}	pZS21::*lptD4213-His* with L312Amber	174
pZS21::*lptD4213-His*^{F313Am}	pZS21::*lptD4213-His* with F313Amber	174
pZS21::*lptD4213-His*^{Y314Am}	pZS21::*lptD4213-His* with Y314Amber	174
pZS21::*lptD4213-His*^{W315Am}	pZS21::*lptD4213-His* with W315Amber	174
pZS21::*lptD4213-His*^{Y678Am}	pZS21::*lptD4213-His* with Y678Amber	174
pZS21::*lptD4213-His*^{Q684Am}	pZS21::*lptD4213-His* with Q684Amber	174
pZS21::*lptD4213-His*^{V685Am}	pZS21::*lptD4213-His* with V685Amber	174
pZS21::*lptD4213-His*^{Y704Am}	pZS21::*lptD4213-His* with Y704Amber	174
pZS21::*lptD4213-His*^{N708Am}	pZS21::*lptD4213-His* with N708Amber	174
pZS21::*lptD4213-His*^{M716Am}	pZS21::*lptD4213-His* with M716Amber	174
pZS21::*lptD4213-His*^{L717Am}	pZS21::*lptD4213-His* with L717Amber	129
pZS21::*lptD4213-His*^{G718Am}	pZS21::*lptD4213-His* with G718Amber	129
pZS21::*lptD4213-His*^{V719Am}	pZS21::*lptD4213-His* with V719Amber	174
pZS21::*lptD4213-His*^{Y721Am}	pZS21::*lptD4213-His* with Y721Amber	129
pZS21::*lptD4213-His*^{Y726Am}	pZS21::*lptD4213-His* with MY726Amber	129
pZS21::*lptD4213-His*^{I728Am}	pZS21::*lptD4213-His* with I728Amber	174
pZS21::*lptD4213-His*^{N737Am}	pZS21::*lptD4213-His* with N737Amber	174
pZS21::*lptD4213-His*^{W739Am}	pZS21::*lptD4213-His* with W739Amber	174
pZS21::*lptD4213-His*^{K743Am}	pZS21::*lptD4213-His* with K743Amber	174
pZS21::*lptD4213-His*^{V747Am}	pZS21::*lptD4213-His* with V747Amber	174
pZS21::*lptD4213-His*^{Y748Am}	pZS21::*lptD4213-His* with Y748Amber	174
pZS21::*lptD4213-His*^{D749Am}	pZS21::*lptD4213-His* with D749Amber	174

pZS21::*lptD4213-His*^{I752Am}	pZS21::*lptD4213-His* with I752Amber	174
pZS21::*lptD4213-His*^{F754Am}	pZS21::*lptD4213-His* with F754Amber	174
pZS21::*lptD4213-His*^{N755Am}	pZS21::*lptD4213-His* with N755Amber	174
pZS21::*lptD4213-His*$^{F276Am, N274I}$	pZS21::*lptD4213-His* with F276Amber and N274I	174
pZS21::*lptD*$^{\Delta D330}$*-His*^{Y244Am}	pZS21::*lptD* $^{\Delta D330}$*-His* with Y244Amber	174
pET23/42	pET23a(+) with multiple cloning sites of pET42a(+), P_{T7}-dependent expression vector	73
pET23/42::*lptD*	Encodes full-length LptD	160
pET23/42::*lptD*^{N274I}	pET23/42::*lptD* with N274I	174
pET23/42::*lptD4213*	Encodes LptD4213 (Δ330-352)	174
pET23/42::*lptD4213*^{N274I}	pET23/42::*lptD4213* with N274I	174
pET23/42::*lptD*$^{\Delta D330}$	pET23/42::*lptD* with ΔD330	174
pET23/42::*lptD*$^{\Delta D330, N274I}$	pET23/42::*lptD*$_{\Delta D330}$ with N274I	174
pET23/42::*lptD-FLAG*	Encodes full-length LptD with a C-terminal $FLAG_3$ tag	33
pET23/42::*lptD4213-FLAG*	Encodes LptD4213 (Δ330-352) with a C-terminal $FLAG_3$ tag	174
pET23/42::*lptD*$^{\Delta D330}$*-FLAG*	pET23/42::*lptD-FLAG* with ΔD330	174
pDS132	*R6K ori*, *mobRP4*, *sacB*, Cam^R	178
pDS132::*bamA*	Encodes full-length BamA	174
pDS132::*bamA*^{E470G}	pDS132::*bamA* with E470G	174
pJH114	pTrc99a/3×FLAG-BamA–BamB–BamC–BamD-8×His–BamE	113
pDT015	pCL/LptE	This study
pDT167	pBAD33/LptD(Δ330-352)-2×Strep–SurA	This study

Chapter 3: Trapping BamA folding intermediates on the Bam complex

This chapter is adapted/reproduced from:
Tomasek D, Rawson S, Lee J, Wzorek JS, Harrison SC, Li Z, Kahne D. Structure of a nascent membrane protein as it folds on the β-barrel assembly machine. *In review* (2020).

3.1. Introduction

To understand how the Bam complex catalyzes β-barrel assembly, we have developed approaches to trap partially folded substrates on the Bam complex of *Escherichia coli*. In doing so, we have studied how structural features of β-barrels promote their folding and membrane integration by the Bam complex[31,127]. One substrate of the Bam complex that we have studied is BamA itself; before BamA can interact with the lipoproteins BamBCDE to form a Bam complex, its β-barrel domain must be properly folded by another Bam complex. We have demonstrated that mutations to the conserved VRGF motif of BamA, which is found in the sixth extracellular loop of the β-barrel domain, prevents efficient membrane integration of BamA and causes stalling on the Bam complex during folding[127]. We concluded that proper positioning of extracellular loops, which often bury into the interior of β-barrels, plays an important role in the mechanism of folding.

Here, we wished to perform a more systematic analysis of the role of extracellular loops in the folding process. To do so, we continued to study BamA as a substrate of the Bam complex ($BamA^S$), so that BamA within the machine ($BamA^M$) folds $BamA^S$. We constructed eight $BamA^S$ substrate variants, each of which contains a deletion of one of the eight extracellular loops. We used biochemical experiments to elucidate details of the mechanism of β-barrel assembly. Notably, we discovered that it possible to stall BamA at different stages of its assembly on the Bam complex based on the position of the deleted loop. We also found that folding of substrate BamA occurs from the C- to N-terminus.

3.2. Results

3.2.1. Generation of a series of mutant BamA substrates

Here, we generated a series of substrates ($BamA^S$), each of which lacked one of the eight extracellular loops of BamA (**Figure 3.1 and Figure 3.2**). In these constructs, no linker is provided to allow for a proper turn between the adjacent β-strands, and we reasoned that this would prevent

efficient β-hairpin formation. Stalling of a substrate should occur at the stage of folding when the deleted loop and the corresponding hairpin would normally be added to the forming β-barrel.

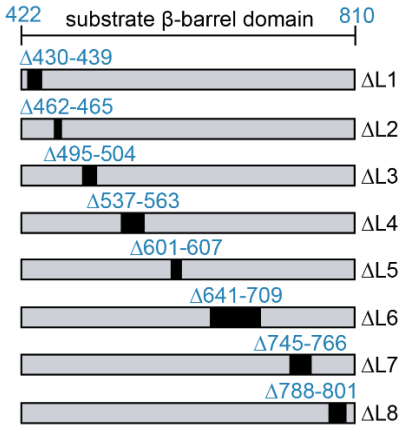

Figure 3.1. Schematic of engineered BamA substrates that contain extracellular loop deletions (BamA^S-ΔL)
The residues deleted in each substrate are indicated. For simplicity, only the β-barrel domain of the substrate is shown (residues 422-810 of BamA).

3.2.2. Substrates with C-terminal deletions are not membrane integrated

We then assessed susceptibility of the substrate variants to the periplasmic protease DegP to determine whether folding stalled in the periplasm before membrane insertion[127]. We reasoned that, if any of the substrates were susceptible to degradation by DegP, this would imply that those substrates have a defect early in folding prior to membrane integration. Removal of loops within the C-terminal half of BamA^S resulted in DegP susceptibility, but removal of loops within the N-terminal half did not (**Figure 3.3a-b**). Likewise, urea extraction experiments, which indicate whether a substrate is integrated in the membrane, showed that substrates lacking loops in the C-terminal half of the protein could be extracted by the denaturant, while substrates lacking N-terminal loops could not be extracted (**Figure 3.3c**). Both experiments are consistent with the idea that the C-terminal half of the BamA^S β-barrel, but not the N-terminal half, is needed for integration into the membrane and hence is assembled early in the folding process.

Next, we wanted to provide clues as to whether the substrates become trapped on the Bam complex during folding. We reasoned that, if substrates expressed from plasmids stall on

the machine, this would lead to reduced folding of endogenously-expressed, native β-barrel substrates, as fewer Bam complexes would be available for their folding. Indeed, we observed reduced levels of mature LptD when BamA^{S}-ΔL1 or BamA^{S}-ΔL2 were expressed, and reduced levels of OmpA when BamA^{S}-ΔL1 was expressed (**Figure 3.3d**). We also observed upregulation of DegP when BamA^{S}-ΔL1 was expressed, indicating the presence of cell stress during expression of this protein. These results suggest that at least BamA^{S}-ΔL1 or BamA^{S}-ΔL2 accumulate on the Bam complex.

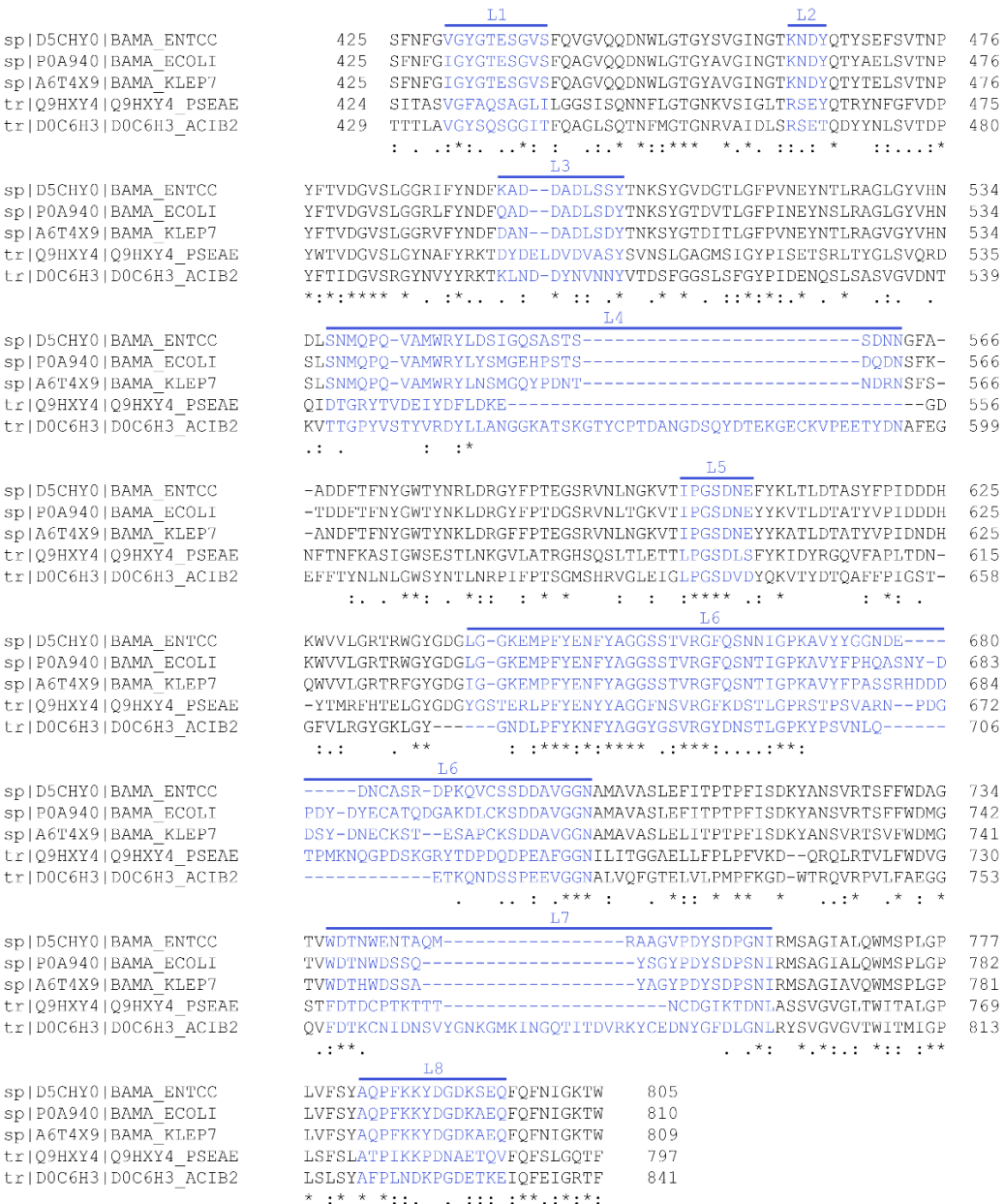

Figure 3.2. Alignment of BamA sequences
BamA sequences from five Gram-negative species are shown. The β-barrel domain is shown, which begins at approximately residue 425 in BamA from *E. coli*. The N-terminal soluble POTRA domains are omitted for simplicity. Locations of extracellular loops are indicated in blue. The boundaries of each extracellular loop are based on BamA from *E. coli* (PDB: 5LJO).

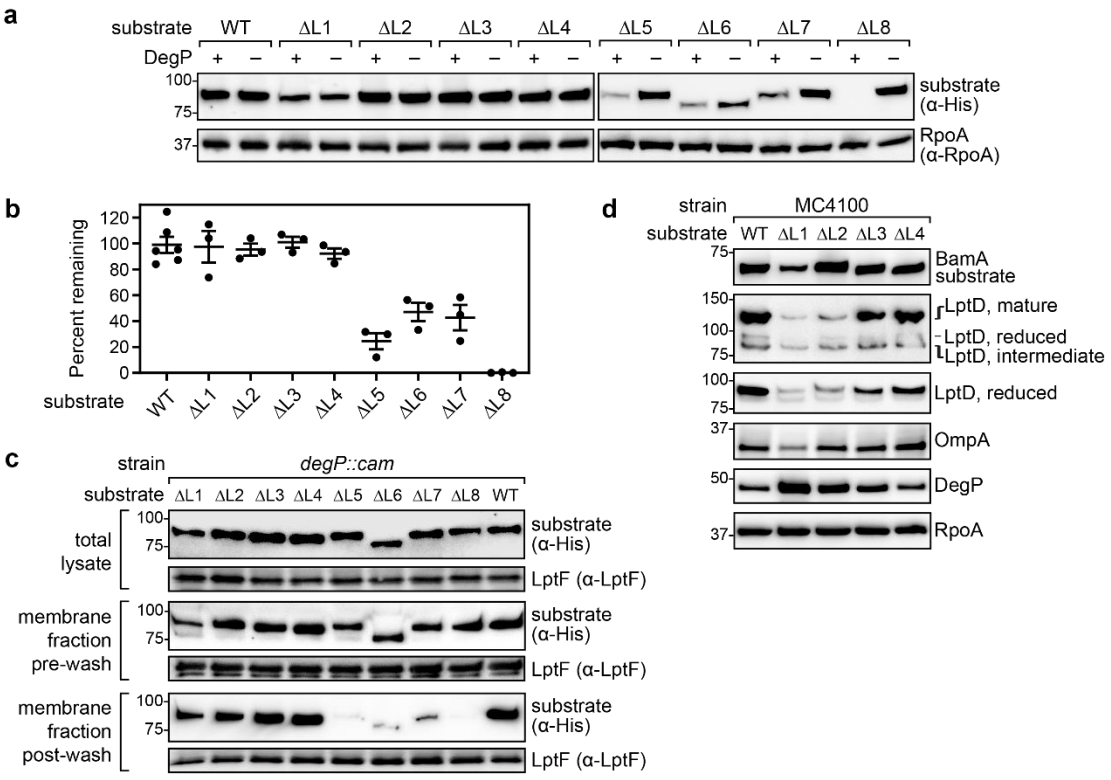

Figure 3.3. Assessments of membrane integration of substrates
a, Expression levels of 6×His-tagged substrates in strains with or without *degP* (MC4100 or MC4100 *degP::cam*, respectively). α-RpoA immunoblots are provided as loading controls. **b**, Quantification of the Western blot data shown in **a**, with expression levels for each substrate calculated as the percent that remains when *degP* is expressed. The plotted data represent mean ± SEM derived from quantification of immunoblotting data in **a** and additional independent replicates (n=6 for WT, n=3 for each mutant substrate). **c**, Urea extraction of 6×His-tagged BamA^{S}-ΔL substrates. Samples from total cell lysates (top), membrane fractions before urea incubation (middle), and membrane fractions that remained after urea incubation (bottom) were analyzed. α-LptF immunoblots are provided as loading controls. **d**, Analysis of expression levels of loop-deleted substrates and endogenous proteins.

3.2.3. Folding occurs at the lateral gate and within the lumen of BamA^{M}

To assess more directly whether substrates accumulate on the Bam complex during folding, we incorporated the photocrosslinkable amino acid para-benzoyl phenylalanine (pBPA) near the N-terminal strand of the seam in BamA^{M} at residue S439 (**Figure 3.4a**). Upon irradiation with UV light, this unnatural amino acid can form crosslinks to co-expressed substrates.

Substrates that fold slowly form stronger crosslinks than wild-type substrates because they have a longer residence time on BamA^{M}. We have previously shown that pBPA substituted at S439 can form crosslinks to the C-terminal region of LptD during its assembly by the Bam complex[174]. Here, we tested crosslinking to BamA^{S} in which POTRA domains 3, 4, and 5 were removed; this deletion prevents the substrates from forming Bam complexes if they are capable of finishing folding. We found that loop-deleted substrates crosslinked strongly to BamA^{M} containing the S439pBPA substitution (**Figure 3.4b**). Along with our experiments probing membrane integration, these results show that removal of extracellular loops in BamA^{S} leads to accumulation on the machine, but either outside or within the membrane, depending on the stage in the folding process at which they stall. Our findings are consistent with earlier reports showing that alterations to the extracellular loops of BamA can result in impaired cell growth[179,180].

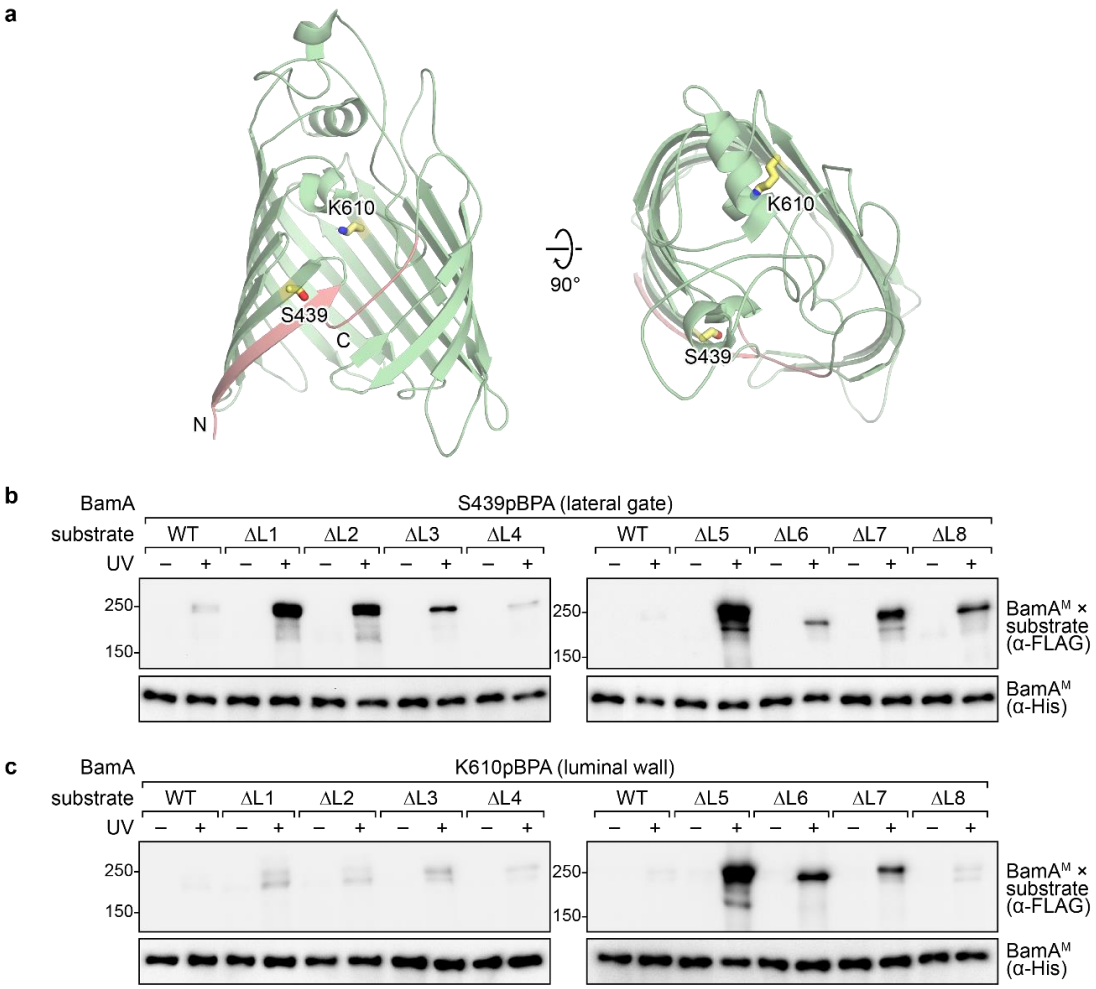

Figure 3.4. *In vivo* photocrosslinking between BamA^{M} and substrates
a, Structure of BamA showing positions that were substituted with pBPA (yellow sticks). The N- and C-terminal β-strands are indicated in salmon. **b**, *In vivo* photocrosslinking of 6×His-tagged full-length BamA^{M} S439pBPA to 3×FLAG-tagged substrates. The experiment is performed in the MC4100 *degP::cam* strain to ensure equal expression levels for all mutants. The substrates also contain deletion of POTRA domains 3, 4, and 5 (Δ172-421) to avoid the possibility of these substrates forming Bam complexes if they complete folding and ensure that only wild-type Bam complexes folding mutant BamA substrates are observed, rather than mutant Bam complexes folding wild-type BamA substrates. **c**, As in **b**, but with pBPA substituted at position 666.

Given our earlier work showing that N-terminal regions of assembling substrates can interact with the interior of the BamA^{M} β-barrel (see Chapter 2), pBPA was also substituted at position K610 (**Figure 3.4a**). This position is located within the interior wall of the BamA^{M} β-barrel and the side chain is oriented toward the aqueous lumen. Incorporation of pBPA at K610 resulted in crosslinks only to substrates containing deletions in the C-terminal half of the protein (ΔL5-L8) (**Figure 3.4c**). Therefore, although substrates containing both C- or N-terminal deletions can stall on the Bam complex, those with N-terminal mutations (ΔL1-L4) appear to stall in a state after they have exited the interior of the BamA^{M} β-barrel. This result establishes that a feature of LptD assembly—entry into the lumen of BamA^{M}—occurs during assembly of BamA^{S}; however, it remains to be determined whether assembly of β-hairpins of BamA^{S} occurs within this environment[174].

3.2.4. Folding begins outside the membrane and finishes within the membrane

To probe membrane integration with greater resolution, we reasoned that side chains within BamA^{S} that are oriented outward may be exposed to lipopolysaccharide (LPS) molecules in the membrane at late, but not early, stages of folding. We sought to incorporate the photocrosslinkable amino acid pBPA at sites in BamA^{S} that would be oriented toward the exterior of the folded protein.

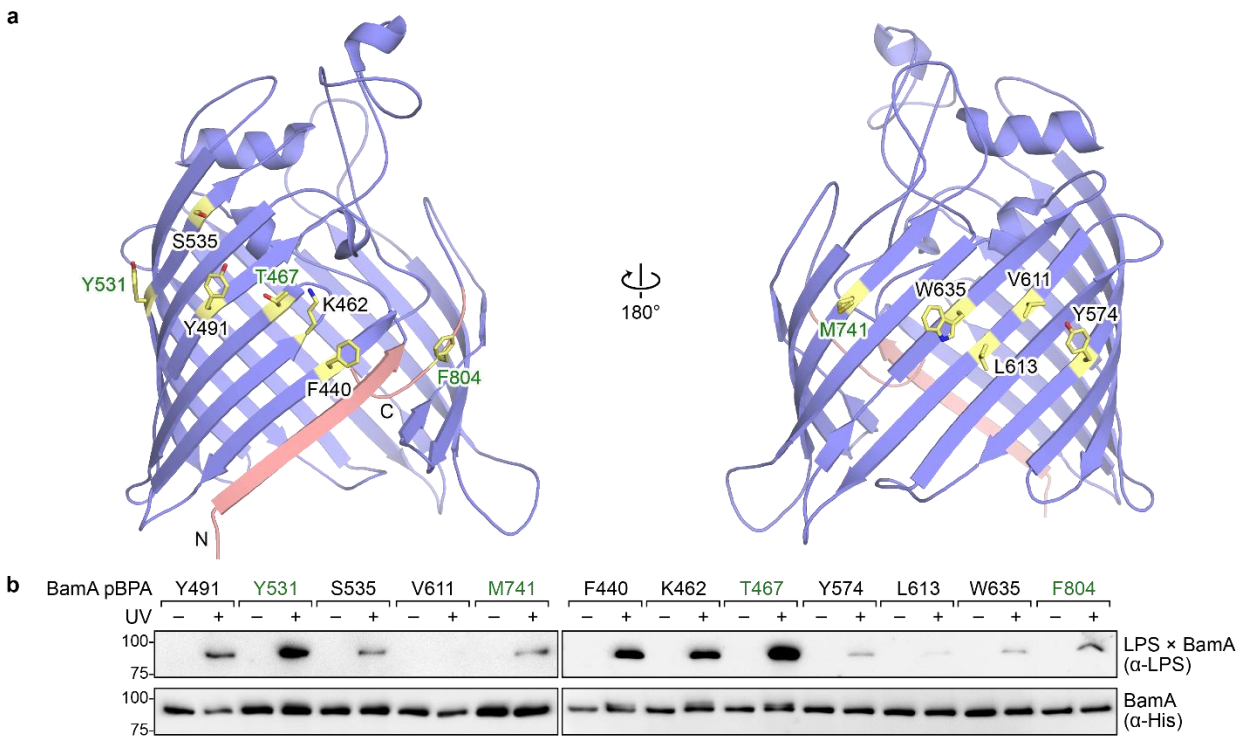

Figure 3.5. *In vivo* photocrosslinking of wild-type BamA to LPS
a, Residues in BamA substituted with the photocrosslinkable amino acid pBPA (yellow sticks) shown on the structure of the β-barrel domain of BamA (blue, PDB ID: 5D0O). These pBPA substitutions were used to observe crosslinks to LPS. The N- and C-termini of BamA are shown in salmon. All highlighted residues have side chains oriented outward towards the membrane environment. **b**, *In vivo* photocrosslinking of BamA to LPS. In **a** and **b**, the pBPA substitutions that were subsequently used to test crosslinking of stalled substrates to LPS (in Figure 3.6) are indicated in green.

We first identified positions in wild-type BamA that crosslink LPS in the folded state (**Figure 3.5**). Then, we tested for crosslinking within stalled substrates using a subset of these pBPA substitutions (T467, Y531, M741, F804). After irradiation of cells with UV light, we purified Bam-BamA^{S} complexes and analyzed crosslinks from BamA^{S} to both LPS and BamA^{M} (**Figure 3.6**). The results showed that the Y531pBPA, M741pBPA, and F804pBPA substitutions in BamA^{S}-ΔL1 yielded crosslinks to LPS, whereas the T467pBPA substitution yielded a crosslink to BamA^{M}. We did not detect crosslinks to LPS using pBPA substitutions within BamA^{S}-ΔL5 or BamA^{S}-ΔL8. This finding, combined with our results above, supports the notion that early stages of folding occur outside the membrane environment[31,127,142,143,181], while late stages of folding occur inside the membrane (compare LPS crosslinking intensity for ΔL1 and ΔL3 with that of ΔL5 and ΔL8).

Therefore, the substrate does not emerge from the BamA^{M} β-barrel in the manner envisioned in the BamA-budding model[133,135,144], as that model would suggest direct insertion of β-hairpins into the membrane during folding without any steps of folding occurring outside the membrane environment.

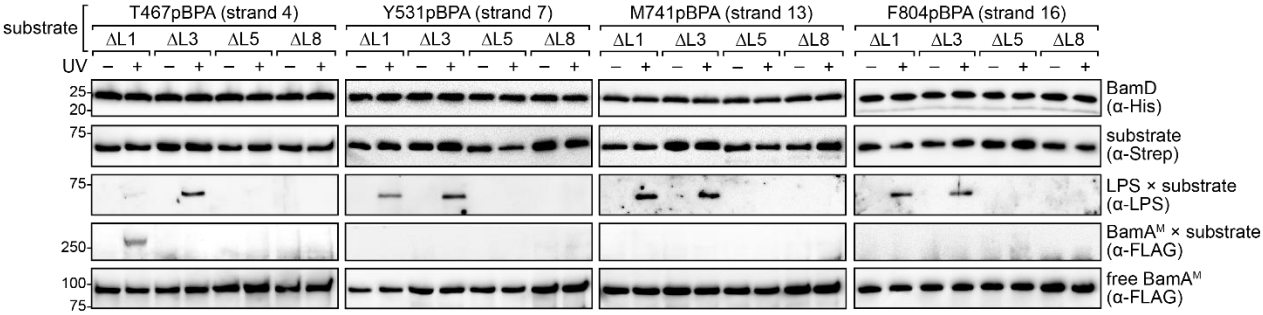

Figure 3.6. *In vivo* photocrosslinking of BamA^{S}-ΔL substrates to BamA^{M} or LPS
Substrates contained pBPA at positions T467, Y531, M741, or F804 and deletion of POTRA domains 3-5. Immunoblotting was performed using α-His, α-Strep, α-LPS, and α-FLAG antibodies to detect BamD (loading control), loop deletion substrates, substrate-bound LPS, and BamA^{M}, respectively. For α-FLAG immunoblot, a longer exposure (top, to detect crosslinks) and a shorter exposure (bottom) are shown.

3.2.5. Folding proceeds directionally from the C- to N-terminus of the substrate

We propose a model in which folding proceeds in a C- to N-terminal direction, consistent with earlier studies of different trapped substrates[143,144,174]. The C-terminus of the substrate, upon interacting with BamA^{M}, could create a new exposed edge to template additional strands. If folding depends on β-strand augmentation[118,122,123], folding will be directional, proceeding from the C-terminus to the N-terminus of the substrate and with each stage depending on successful completion of the previous stage. We probed directional folding by generating six BamA^{S} variants in which two extracellular loops had been removed rather than only one (**Figure 3.7a**). In each construct, removal of the loop closer to the C-terminus determined the overall susceptibility to protease digestion, which reports on membrane integration (**Figure 3.7b-c**). These differences in proteolytic susceptibility allow us to conclude that the region including loop 8 folds before that including loop 5, and so forth. Additionally, we tested urea extraction (**Figure 3.7d**) and crosslinking to S439pBPA within BamA^{M} (**Figure 3.7e**), and these experiments showed similar

results. This finding establishes C- to N-terminal directional folding and is consistent with a β-strand augmentation model.

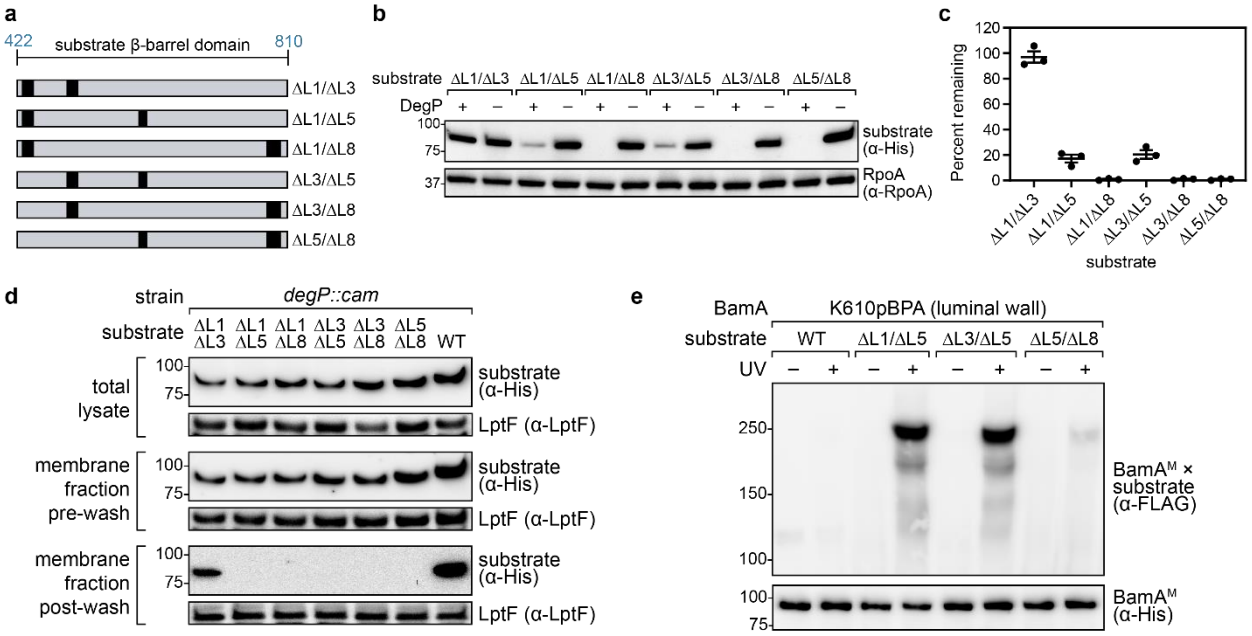

Figure 3.7. β-barrel substrates are folded directionally from the C- to N-terminus
a, Schematic of substrates in which two extracellular loops are deleted. Data shown are representative of results from two biological replicates. **b**, Expression levels of substrates containing two loop deletions in strains with or without *degP* (MC4100 or MC4100 *degP::cam*, respectively). α-RpoA immunoblot is provided as a loading control. **c**, Quantification of the Western blot data shown in **b**, with expression levels for each substrate calculated as the percent that remains when *degP* is expressed. The plotted data represent mean ± SEM derived from quantification of immunoblotting data in **b** and additional independent replicates (n=3 for each substrate). **d**, Urea extraction of 6×His-tagged substrates. Samples from total cell lysates (top), membrane fractions before urea incubation (middle), and membrane fractions that remained after urea incubation (bottom) were analyzed. α-LptF immunoblots are provided as loading controls. Data shown are representative of results from two biological replicates. **e**, *In vivo* photocrosslinking of 6×His-tagged full-length $BamA^{M}$ S439pBPA to 3×FLAG-tagged substrates. The experiment is performed in the MC4100 *degP::cam* strain to ensure equal expression levels for all mutants.

3.2.6. A substrate-engaged Bam complex can be expressed and purified

To pursue structural studies of a substrate-bound Bam complex, we needed first to select a substrate to capture on the machine. We chose the substrate with loop 1 removed ($BamA^{S}$-ΔL1) because it accumulates at a late stage of folding, when it has already integrated into the membrane but has not yet released from the Bam complex. We reasoned that using a substrate that stalls as a nearly folded protein would facilitate structural studies, as disordered regions in an unfolded substrate would likely hinder determination of a high-resolution structure. As in the

experiments reported above, here we deleted POTRA domains 3, 4, and 5 of the substrate to avoid the possibility of the substrate forming Bam complexes if it does complete folding. This allows us to ensure that we only pull down wild-type Bam complexes folding BamA-ΔL1 as the substrate, rather than also purifying complexes containing BamA-ΔL1 as the machine, as both complexes would contain the same two affinity tags. Importantly, the deletion of these POTRA domains does not prevent folding of an otherwise wild-type BamA substrate[163].

To stabilize the interactions between $BamA^{M}$ and the $BamA^{S}$-ΔL1 substrate, we engineered cysteines into the two proteins (**Figure 3.8**) that could be used to form disulfide bonds. Using disulfide bond formation, we reasoned that the substrate would remain associated with the Bam complex throughout purification. For initial testing, we introduced E800C, F802C, or F804C into $BamA^{M}$ and G460C, T467C, or A469C into $BamA^{S}$-ΔL1. The positions of these cysteine substitutions were selected based on reports that the C-terminal region of $BamA^{M}$ is in proximity to the N-terminal region of the substrate during folding[143,144,174].

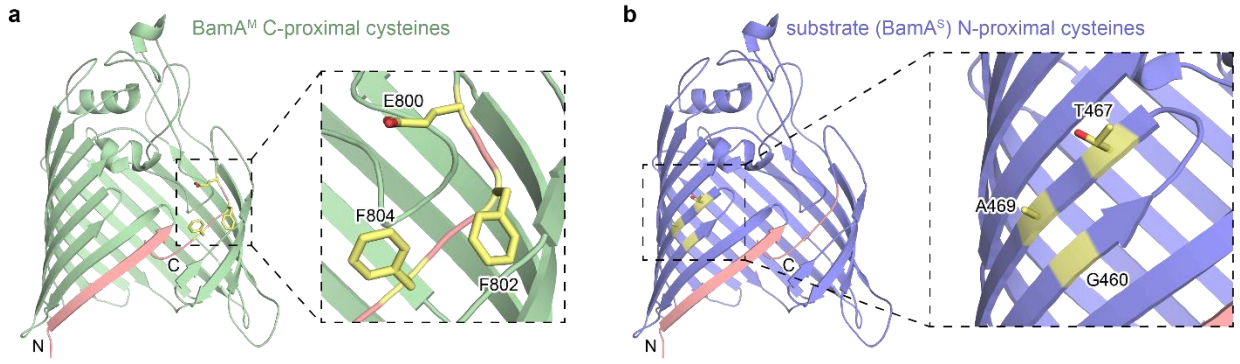

Figure 3.8. Positions of engineered cysteines in $BamA^{M}$ and $BamA^{S}$ for testing disulfide bond formation
a, Residues near the C-terminus of the β-barrel domain of $BamA^{M}$ that were substituted with cysteine (yellow sticks) shown on the structure of the β-barrel domain of BamA (green, PDB ID: 5D0O). The N- and C-terminal β-strands are labelled and shown in salmon. **b**, As in **a**, but showing residues near the N-terminus of the β-barrel domain of the substrate ($BamA^{S}$-ΔL1, blue) that were substituted with cysteine.

After creating a panel of cysteine mutants in $BamA^{M}$ and the substrate, we next tested for disulfide bond formation between the two proteins by co-expressing them (**Figure 3.9**). In this experiment, the presence of disulfide bond is detected as a high molecular weight adduct on the

α-FLAG immunoblot that represents the combined masses of the two proteins (**Figure 3.9a, middle immunoblot**). We observed disulfide bond formation between BamA^{M} and the substrate using most of the cysteine pairs tested. We also tested disulfide bond formation using BamA^{M} without any introduced cysteine (**Figure 3.9b**). In this control, we observed no disulfide bond formation between the proteins, demonstrating that the disulfides observed in Figure 3.9a are indeed formed between BamA^{M} and the substrate, and do not instead reflect non-specific association of the proteins. Importantly, when the same cysteine combinations were tested using substrates containing all five POTRA domains (instead of only POTRA domains 1 and 2), we observed similar disulfide bond formation to BamA^{M}. This control, with results presented in Figure 3.10 below, demonstrates that stalling of the BamA^{S}-ΔL1 substrate is caused by the deletion of extracellular loop 1 and not by the deletion of POTRA domains 3, 4, and 5.

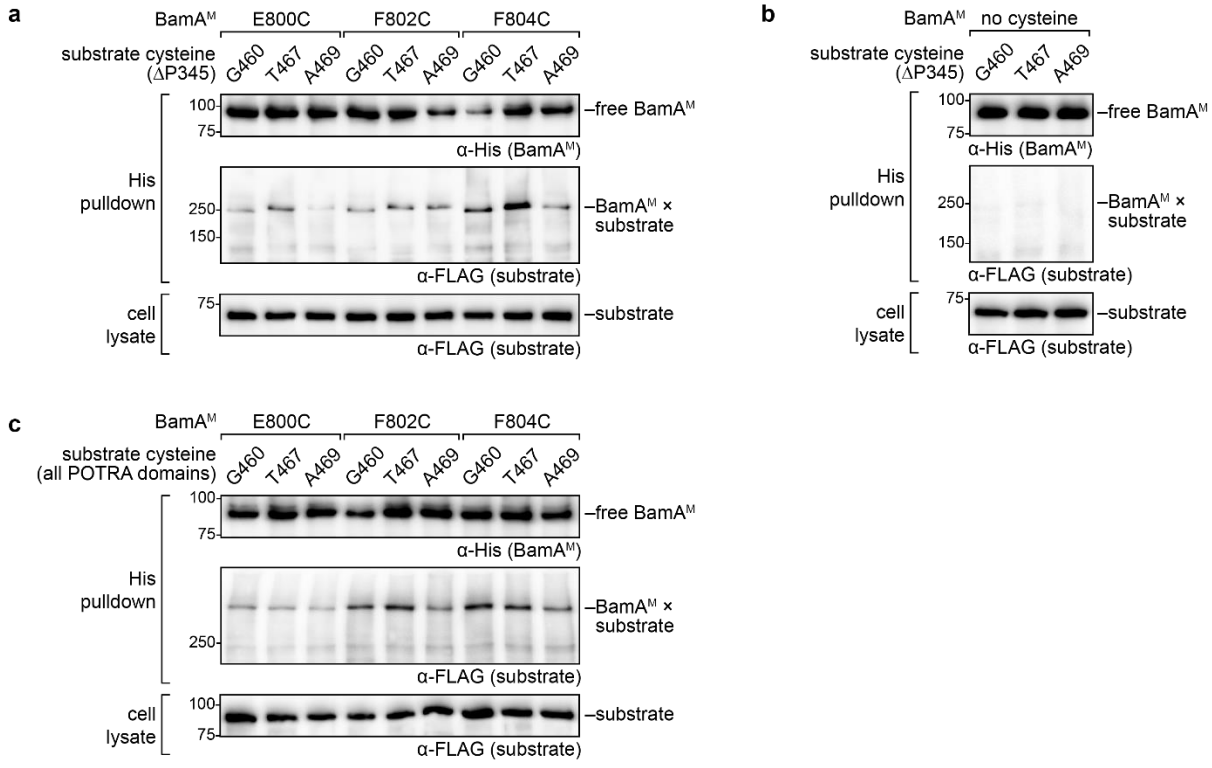

Figure 3.9. Disulfide bond formation between BamA^{M} and BamA^{S}-ΔL1
a, Disulfide bond formation between 6×His-tagged BamA^{M} and 3×FLAG-tagged BamA^{S}-ΔP345-ΔL1 containing cysteine substitutions shown. The presence of a Bam-substrate disulfide bond is detected as a high molecular weight adduct on the α-FLAG (middle) immunoblot. The adducts generated via disulfide bond formation were not present in an amount high enough for detection with the α-His antibody but could

be detected with the α-FLAG antibody. α-FLAG immunoblot of total cell lysates (bottom) shows that the substrates containing different cysteine substitutions are expressed at similar levels. **b**, As in **a**, but without any cysteine introduced into $BamA^{M}$. Disulfide bond formation is not observed between $BamA^{M}$ and $BamA^{S}$-ΔL1 when no cysteine is introduced into $BamA^{M}$. **c**, As in **a**, but with substrates containing all five POTRA domains.

While most of the cysteine pairs yielded disulfide bonds between $BamA^{M}$ and $BamA^{S}$-ΔL1, we observed relatively strong formation using the $BamA^{M}$(F804C) and $BamA^{S}$-ΔL1(T467C) pair. Thus, we chose to use this pair for large-scale purification of the substrate-bound complex for structural studies, and generated plasmids that could overexpress the Bam complex and the $BamA^{S}$-ΔL1 substrate containing these cysteine substitutions (**Figure 3.10a**). Specifically, we expressed BamABCDE from one plasmid and the substrate from another; in this setup, BamE within the machine contains an affinity tag (8×His) on BamE and the substrate contains another (Twin-Strep). As discussed above, the substrate we used for overexpression had POTRA domains 3, 4, and 5 (Δ172-421) removed to avoid the possibility of the substrate forming Bam complexes if it completes folding. Using heat modifiability as an assay for folding, we showed that substrate is indeed capable of folding under reducing conditions, though it does so slowly (**Figure 3.10b**). In this assay, unboiled cell lysates applied to SDS-PAGE can be used to calculate the fraction of each substrate that is folded *in vivo* (indicated below the blot). Since an otherwise identical construct that contained loop 1 folded nearly to completion, this showed that $BamA^{S}$-ΔL1 accumulates on the Bam complex (leading to reduced levels of folded protein) because of the deletion of loop 1 and not because of the introduction of the cysteine or deletion of the POTRA domains.

Next, we purified the $BamA^{M}$BCDE-$BamA^{S}$ complex using standard methods but including the addition of copper(II) sulfate and 1,10-phenanthroline as an oxidizing agent to promote disulfide bond formation. The complex was able to be purified to homogeneity as assessed by size-exclusion chromatography that yielded a monodisperse peak (**Figure 3.10c**). Analysis of peak fractions from size-exclusion chromatography using SDS-PAGE demonstrated the purity of

the complex (**Figure 3.10d**) with the majority of the BamA^{M} and substrate proteins forming disulfide bonds (left panel, compare the highest band to the two bands below representing free BamA^{M} and free substrate). The yield obtained was approximately 0.1 milligrams of purified complex per 1 liter of bacterial culture. Although this amount was relatively low, and precluded studies of the complex by X-ray crystallography, it was suitable for pursuing cryo-EM experiments (see Chapter 4).

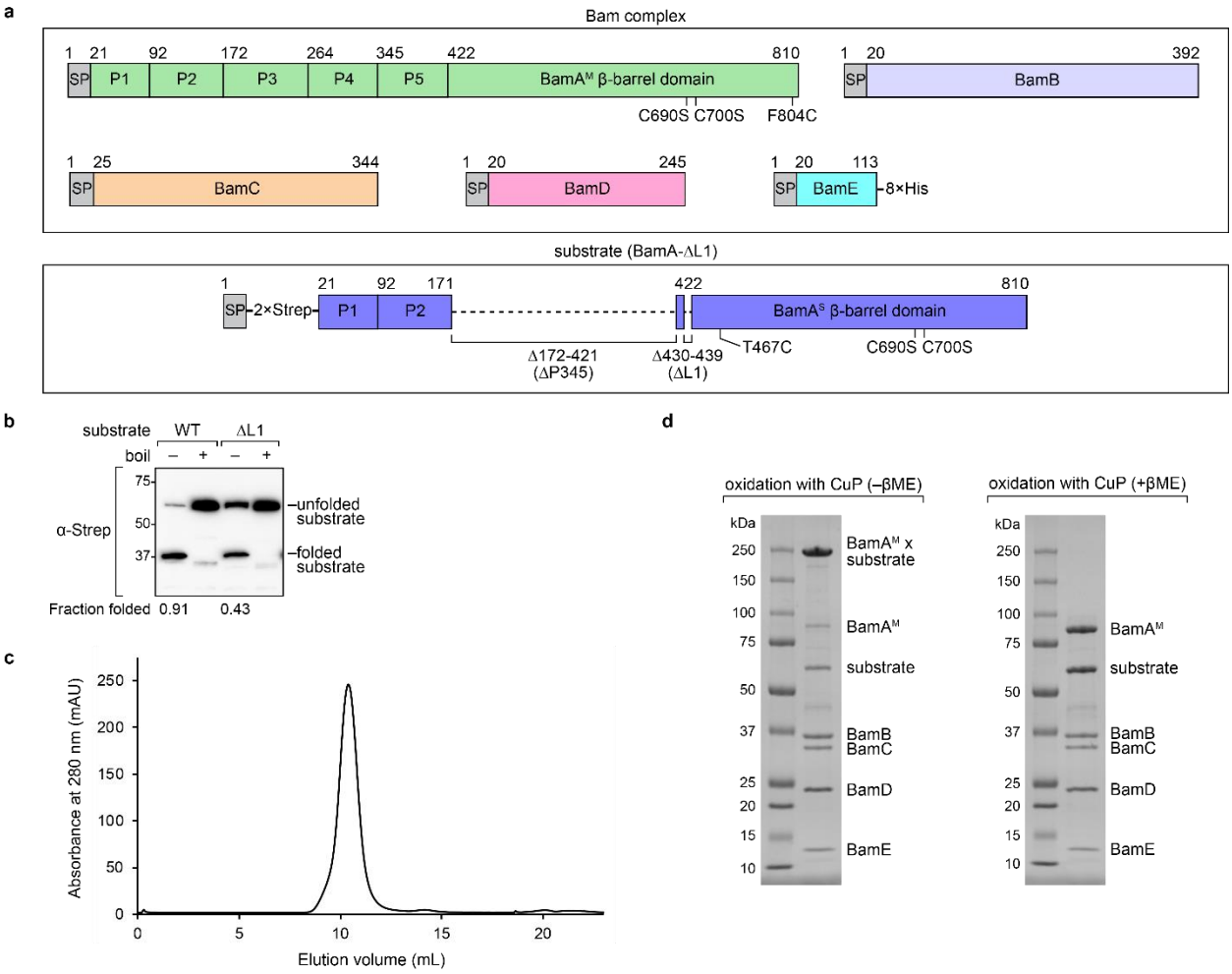

Figure 3.10. Expression and purification of a substrate-engaged Bam complex for cryo-EM
a, Constructs used for protein expression. Mutations and affinity tags introduced into each protein are indicated. SP represents signal peptide; P1 to P5 represent the POTRA domains of BamA. POTRA domains 3, 4, and 5 (Δ172-421) of the substrate are removed. **b**, Assessment of heat modifiability of substrate used for cryo-EM (ΔL1, right half of blot) that contains the T467C mutation and a substrate that contains loop 1 but is otherwise identical (WT, left half of blot). The samples in the unboiled lanes can be used to calculate the fraction of each substrate that is folded *in vivo* (indicated below the blot). BamA^{S}-ΔL1 accumulates on the Bam complex because of the deletion of loop 1. **c**, Representative size-exclusion chromatogram of the substrate-bound Bam complex in which disulfide bond formation between the N-

terminus of the substrate (BamA^{S}-ΔL1) and the C-terminus of BamA^{M} was induced with the oxidizing agent copper sulfate/1,10-phenanthroline (CuP). **d**, SDS-PAGE gels showing peak fractions from size-exclusion chromatography of the complex in **c**. The left gel shows the complex without addition of βME, while the right gel shows the complex after addition of βME to break the disulfide bond.

3.3. Discussion

Here we have assessed a panel of substrates (BamA^{S}), each of which had one extracellular loop removed. Characterization of these substrates was then performed with biochemical assays including protease accessibility, extractability with urea, *in vivo* photocrosslinking, and cysteine crosslinking. Based on these experiments, two distinct groups of substrates were characterized. In one group, a deletion was made within the second half of the protein, and this led to trapping of the substrate at an early stage of folding prior to membrane integration. In the second group, the deletion made was within the first half of the protein and caused trapping at a late stage of folding after membrane integration had occurred. Substrates within either group stalled during folding on the Bam complex, as interactions could be observed with the lateral gate of BamA^{M}. This demonstrates that the Bam complex can hold substrates at its lateral gate throughout folding. Interestingly, we observed interactions with the interior of the BamA^{M} β-barrel only for substrates stalled at the earlier stage of folding, suggesting that substrates are housed within this environment early and are then released into the membrane later in folding once a substantial amount of the β-barrel has been formed.

Given that our results suggested a directional mechanism of folding (i.e., starting at the C-terminus of the substrate and proceeding towards the N-terminus), we then performed experiments with substrates containing two extracellular loop deletions; since these substrates resembled the substrate containing the single deletion closest to the C-terminus, this provided support for a sequential, C- to N-terminal folding mechanism.

Having the ability to stall a substrate at different positions of assembly, we were in the position to use our findings to purify a substrate-bound Bam complex in sufficient amounts for

structural studies. Using disulfide bond formation, we were able to accomplish this and, in Chapter 4, cryo-EM experiments using this complex will be described.

3.4. Materials and methods

3.4.1. SDS-PAGE and immunoblotting

Homemade Tris-HCl 4-20% polyacrylamide gradient gels or Mini-PROTEAN TGX 7.5% precast gels (BioRad) were used with Tris-glycine running buffer. 2× SDS sample buffer refers to a mixture containing 125 mM Tris (pH 6.8), 4% (wt/vol) SDS, 30% (vol/vol) glycerol, 0.005% bromophenol blue, and 5% (vol/vol) β-mercaptoethanol. SDS-PAGE analyses were performed at 200 V for 45 to 60 min. To analyze purified protein complexes for cryo-EM, SDS-PAGE was performed and followed by staining with Coomassie Brilliant Blue (Alfa Aesar). Coomassie-stained SDS-PAGE gels were imaged using the “Gel” feature of an Azure Biosystems C400 imager. For Western blotting, proteins were transferred onto Immun-Blot PVDF membranes (Bio-Rad) and then incubated with appropriate antibodies. All HRP conjugates were visualized with the Amersham ECL Prime Western blotting detection reagent (GE Healthcare). Western blots were imaged using the “lowest” sensitivity setting of the “Chemi” feature of an Azure Biosystems C400 imager.

3.4.2. Analysis of cellular BamA levels

Derivatives of a plasmid containing 6×His-BamA cloned into the pZS21 vector (pSK476) were generated to contain single loop deletions (pDT175-pDT182), double loop deletions (pDT260-pDT265), or C-terminal kink mutations (pDT521, pDT536-pDT538, pDT550). MC4100 cells[173] and MC4100 *degP::cam* cells[127] were transformed with these plasmids. The resulting strains were grown in LB supplemented with 50 μg/mL kanamycin (for MC4100) or 50 μg/mL kanamycin and 30 μg/mL chloramphenicol (for MC4100 *degP::cam*) (37°C, 220 rpm). Once an

OD600 of ~1.0 was reached, the cells from a 1 mL sample were collected by centrifugation (5,000 x g, 10 min, 4°C). The resulting cell pellets were resuspended in 80 µL of a 1:1 mixture of 2× SDS sample buffer and buffer containing 20 mM Tris-HCl pH 8.0 and 150 mM NaCl. After boiling for 10 min, the samples were applied to SDS-PAGE and analyzed via Western blotting. 6×His-BamA was detected by using a penta-His (HRP) antibody (Qiagen). RpoA was detected using a mouse anti-RpoA (*E. coli* RNA Polymerase α) primary antibody (BioLegend, clone 4RA2) followed by a sheep anti-mouse (HRP) secondary antibody (GE Healthcare).

3.4.3. Membrane extraction with urea

Membrane extraction was performed in a manner as previously described[182]. MC4100 *degP::cam* strains were generated harboring the pZS21 vector encoding 6×His-BamA or derivatives containing single or double loop deletions. These strains were grown overnight in LB supplemented with 50 µg/mL kanamycin and 30 µg/mL chloramphenicol (37°C, 220 rpm). These cultures were used to inoculate 100 mL of LB with the same additives via 1:100 dilutions. The resulting cultures were grown (37°C, 220 rpm) until an OD600 of ~1.0 was reached. The cells were harvested via centrifugation (5,000 x g, 10 min, 4°C) and resuspended in 4 mL of resuspension solution (50 mM Tris-HCl pH 8.0, 10 µg/mL DNase). Cells were lysed via sonication (on and off at 10 second intervals for a total of 3 min on and 3 min off). The resulting lysate was centrifuged to pellet cell debris (5,000 x g, 10 min, 4°C). The supernatant was used to prepare normalized cell lysate suspensions to an A280 value of ~24 (assayed by using NanoDrop). An aliquot of each normalized lysate sample was removed, mixed 1:1 with 2× SDS sample buffer (containing β-mercaptoethanol), and boiled for 10 min. For each sample, 800 µL of the remaining lysate was transferred to 70.1 Ti ultracentrifuge tubes (Beckman Coulter) along with an additional 5 mL of 50 mM Tris-HCl pH 8.0. Membrane fractions were then isolated via ultracentrifugation (100,000 x g, 25 min, 4°C) using an Optima XE-90 ultracentrifuge (Beckman Coulter). The supernatant was discarded, and the membrane pellet was resuspended in 800 µL of 6 M urea.

The resuspensions were incubated at 37°C for 1 hour. An aliquot of each sample, representative of total membrane protein content, was removed, mixed 1:1 with 2× SDS sample buffer (containing β-mercaptoethanol), and boiled for 10 min. The post-wash membrane pellet was isolated from the remainder of each sample via ultracentrifugation (100,000 x g, 45 min, 20°C) after addition of 5 mL of 50 mM Tris-HCl pH 8.0. The pellets were resuspended in 500 µL of water. An aliquot of each sample was removed, mixed 1:1 with 2× SDS sample buffer (containing β-mercaptoethanol), and boiled for 10 min. The samples were applied to SDS-PAGE and analyzed via Western blotting. 6×His-BamA was detected by using a penta-His (HRP) antibody (Qiagen). LptF was detected using LptF antiserum followed by a donkey anti-rabbit (HRP) secondary antibody (GE Healthcare). The source of rabbit anti-LptF antiserum has been previously reported[18].

3.4.4. *In vivo* photocrosslinking of BamA to substrates

Photocrosslinking experiments are based on techniques as previously described[29,139], with modifications. The plasmid pSup-BpaRS-6TRN encodes an orthogonal tRNA and aminoacyl-tRNA synthetase to incorporate the unnatural amino acid pBPA at amber (TAG) stop codons[183]. This plasmid contains a chloramphenicol resistance cassette, so is not compatible with the MC4100 *degP::cam* strain. To circumvent this, we generated a variant of the pSup-BpaRS-6TRN plasmid that instead contains a spectinomycin resistance cassette. In brief, pSup-BpaRS-6TRN minus the chloramphenicol resistance cassette was amplified, and the spectinomycin resistance cassette was amplified from the pCDFDuet vector (EMD Millipore). The pSup-BpaRS-6TRN backbone and the spectinomycin resistance cassette were joined via Gibson assembly[184] to generate pSup-BpaRS-6TRN(spec^R) (pDT504).

6×His-BamA^M containing pBPA substitutions were cloned into the pZS21 vector to generate pJL77. 3×FLAG-tagged substrates were cloned into the pTrc99a vector to generate pDT209 (WT BamA^S); pDT201-pDT208 (single loop deletions); pDT268, pDT270, and pDT272

(double loop deletions); and pDT526, pDT558-pDT560, and pDT562 (C-terminal kink mutants). The 3×FLAG-tagged substrates also contain deletion of POTRA domains 3, 4, and 5 (Δ172-421) to avoid the possibility of these substrates forming Bam complexes if they complete folding. The deletion does not prevent folding of an otherwise wild-type BamA[163], and allows us to assess BamA^{S} mutants in terms of their ability to fold rather than function within a Bam complex.

MC4100 *degP::cam* strains were generated, each harboring pSup-BpaRS-6TRN(spec^{R}), pJL77, and one of the substrate-encoding plasmids. These strains were grown overnight in LB supplemented with 50 μg/mL carbenicillin, 50 μg/mL kanamycin, 50 μg/mL spectinomycin, and 0.2% (w/v) glucose (37°C, 220 rpm). These overnight cultures were diluted 1:100 into 100 mL of fresh LB containing the same additives minus glucose but supplemented with 0.8 mM H-p-Bz-Phe-OH (pBPA, Bachem), and were grown (37°C, 220 rpm) to OD600 ~0.35. After normalization by optical density, each culture was split in half, with one half used directly for irradiation with UV light at 365 nm for 10 min (on ice). Photocrosslinking was performed by using a UVP Blak-Ray B-100AP high-intensity UV lamp with a 100 W spot bulb. All cells were then pelleted by centrifugation (5,000 x g, 10 min, 4°C).

For protein purification, pellets were resuspended in 5 mL TBS (20 mM Tris-HCl pH 8.0, 300 mM NaCl) supplemented with 1% (w/v) Anzergent 3-14 (Anatrace), 20 mM imidazole (pH 8.0), 100 μg/mL lysozyme (Sigma-Aldrich), 1 mM PMSF (Sigma-Aldrich), and 50 μg/mL DNase I (Sigma-Aldrich). Cells were lysed via sonication (on and off at 10 second intervals for a total of 1 min and 30 sec on and 1 min and 30 sec off). The resulting lysate was centrifuged (10,000 x g, 10 min, 4°C). An aliquot of the supernatant was taken for analysis of 3×FLAG-BamA^{S} expression levels via SDS-PAGE and immunoblotting using a monoclonal anti-FLAG M2-peroxidase (HRP) mouse antibody (Sigma-Aldrich). The remainder of the supernatant was incubated with Ni-nitrilotriacetic acid (NTA) resin on a rocking platform (1 hour, 4°C). After removal of unbound proteins, the resin was washed twice with 50 CV (column volumes) TBS containing 0.02% Anzergent 3-14 and 20 mM imidazole (pH 8.0). Samples were eluted with 5 CV TBS containing

0.02% Anzergent 3-14 and 200 mM imidazole (pH 8.0). Eluates were supplemented with 10% TCA by volume and incubated on ice for 20 min. Precipitated proteins were pelleted via centrifugation (21,130 x g, 10 min, 4°C). All samples were resuspended in a 1:1 mixture of 2× SDS loading dye (containing β-mercaptoethanol) and 1 M Tris pH 8.0. After boiling for 20 min, each sample was analyzed by SDS-PAGE and Western blotting. 6×His-BamA^{M} was detected by using a penta-His (HRP) antibody (Qiagen). 3×FLAG-BamA^{S} that was pulled down with 6×His-BamA^{M} was detected by using a monoclonal anti-FLAG M2-peroxidase (HRP) antibody (Sigma-Aldrich).

3.4.5. *In vivo* photocrosslinking of BamA to LPS

6×His-BamA containing pBPA substitutions were cloned into the pZS21 vector to generate pDT411, pDT416, and pDT421-pDT430. MC4100 strains were generated, each harboring pSup-BpaRS-6TRN (containing a chloramphenicol resistance cassette) and one of the plasmids encoding pBPA-substituted BamA. These strains were grown overnight in LB supplemented with 50 μg/mL kanamycin and 30 μg/mL chloramphenicol. These overnight cultures were diluted 1:100 into 100 mL of fresh LB containing the same additives supplemented with 0.8 mM H-p-Bz-Phe-OH (pBPA, Bachem), and were grown (37°C, 220 rpm) to OD600 ~0.35. The protocols for UV irradiation and protein purification are the same as described above for *in vivo* photocrosslinking of BamA^{M} to substrates.

3.4.6. *In vivo* photocrosslinking of substrates to LPS or BamA^{M}, followed by DSP crosslinking to the Bam complex

In order to observe crosslinking from substrates to lipopolysaccharide or to BamA^{M}, a separate crosslinking protocol was devised to ensure that only substrates stalled on the Bam complex were assessed. In brief, we crosslinked the exterior surface of substrates to their surroundings (using incorporated pBPA) and then crosslinked stalled substrates to the Bam

complex (using an amine-to-amine crosslinker). This was followed by purification of Bam-substrate complexes and release of the substrate with reducing agent (which breaks the amine-to-amine crosslinks), allowing assessment of UV-dependent crosslinks by SDS-PAGE and Western blotting.

The pJH114 plasmid[113] was modified to encode 3×FLAG-BamA^{M} and BamD-8×His (while eliminating the 8×His tag on BamE), generating pDT340. 2×Strep-tagged BamA^{S} substrates containing deletion of POTRA domains 3-5, deletion of an extracellular loop (L1, L3, L5, or L8), and a pBPA substitution (at T467, Y531, M741, or F804) were cloned into the pCDF vector to generate pDT436-pDT439 (T467pBPA), pDT451-pDT454 (Y531pBPA), pDT471-pDT474 (M741pBPA), and pDT476-pDT479 (F804pBPA). BL21(DE3) strains were generated harboring pSup-BpaRS-6TRN, pDT340, and one of the substrate-encoding plasmids.

These strains were grown overnight in LB supplemented with 50 μg/mL carbenicillin, 50 μg/mL spectinomycin, 30 μg/mL chloramphenicol, and 0.2% (w/v) glucose (28°C, 220 rpm). These overnight cultures were diluted 1:100 into 100 mL of fresh LB containing the same additives, and were grown (37°C, 220 rpm) to OD600 ~0.35. At this point, isopropyl β-D-1-thiogalactopyranoside (IPTG, VWR) and 0.8 mM H-p-Bz-Phe-OH (pBPA, Bachem) were added to final concentrations of 0.1 mM and 0.8 mM, respectively. The strains were grown for an additional 90 minutes. After normalization by optical density, each culture was split in half, with one half used directly for irradiation with UV light at 365 nm for 10 min (on ice). All cells were then pelleted by centrifugation (5,000 x g, 10 min, 4°C). Pellets were resuspended in 20 mL PBS (20 mM NaH_2PO_4 pH 7.2, 150 mM NaCl) and the amine-to-amine crosslinker dithiobis(succinimidyl propionate) (DSP, Thermo Fisher Scientific) was added to a final concentration of 0.5 mM. After incubation on a rocking platform (30 to 60 min, room temperature), the crosslinking reaction was quenched via addition of Tris-HCl to a final concentration of 20 mM. Cells were centrifuged (5,000 x g, 10 min, 4°C) and the pellets were frozen at -80°C prior to subsequent purification.

Protein purification was performed similarly as described for testing site-specific *in vivo* crosslinking without DSP crosslinking, with a few differences. The Ni-NTA wash step was performed using 50 CV TBS containing 0.1% Triton X-100 (Sigma-Aldrich), 0.1% SDS (J. T. Baker), and 40 mM imidazole (pH 8.0). The Ni-NTA elution step was performed using 5 CV TBS containing 200 mM imidazole (pH 8.0) and no detergent. Instead of using TCA precipitation, each sample was concentrated using an Amicon Ultra 0.5 mL 10 kDa molecular-weight cutoff centrifugal concentrator (EMD Millipore). All samples were supplemented with an equal volume of 2× SDS loading dye (containing β-mercaptoethanol). After boiling for 10 min, each sample was analyzed by SDS-PAGE and Western blotting. 8×His-BamD was detected by using a penta-His (HRP) antibody (Qiagen). 2×Strep-BamA^{S} substrates were detected using a monoclonal Strep-Tag II (HRP) antibody (EMD Millipore). 3×FLAG-BamA^{M} was detected by using a monoclonal anti-FLAG M2-peroxidase (HRP) antibody (Sigma-Aldrich). Lipopolysaccharide was detected using a mouse monoclonal anti-lipopolysaccharide core primary antibody (Hycult Biotech, clone WN1 222-5) followed by a sheep anti-mouse (HRP) conjugate secondary antibody (GE Healthcare).

3.4.7. Assessing cysteine-to-cysteine crosslinking

6×His-BamA^{M} containing a cysteine substitution was cloned into the pZS21 vector to generate pDT511, pDT512, and pDT513. 3×FLAG-tagged substrates (BamA^{S} containing deletion of POTRA domains 3-5, deletion of loop 1, and a cysteine substitution) were cloned into the pTrc99a vector to generate pDT514-pDT516. MC4100 cells were transformed with one BamA^{M}-encoding plasmid and one substrate-encoding plasmid. The resulting strains were grown overnight in LB supplemented with 50 μg/mL carbenicillin, 50 μg/mL kanamycin, and 0.2% (w/v) glucose (37°C, 220 rpm). These overnight cultures were diluted 1:100 into 100 mL of fresh LB containing the same additives minus glucose, and were grown (37°C, 220 rpm) to OD600 ~0.5. Cells were then collected by centrifugation (4,200 x g, 10 min, 4°C). Cell pellets were resuspended

in PBS (20 mM NaH_2PO_4 pH 7.2, 150 mM NaCl). TCEP-HCl (VWR) was then added at a final concentration of 2 mM, and cells were incubated on a rocking platform (20 min, room temperature). Cells were then centrifuged (5,000 x g, 10 min, 4°C) and again resuspended in PBS. The cysteine-to-cysteine crosslinker 1,2-bis(maleimido)ethane (BMOE, AstaTech) was added at a final concentration of 0.5 mM. After incubation on a rocking platform (40 min, room temperature), the crosslinking reaction was quenched via addition of L-cysteine hydrochloride monohydrate (Alfa Aesar) to a final concentration of 10 mM. Cells were centrifuged (5,000 x g, 10 min, 4°C) and the pellets were frozen at -80°C prior to subsequent purification. To test disulfide bond formation in the absence of crosslinker, cells were frozen after the initial centrifugation step. In each sample, 6×His-$BamA^M$ was purified and crosslinked adducts were detected by Western blotting as described for site-specific *in vivo* photocrosslinking experiments. β-mercaptoethanol was not added to samples used to assess disulfide formation in the absence of crosslinker.

3.4.8. Assessing heat modifiability of substrates

2×Strep-$BamA^S$ containing deletion of POTRA domains 3-5 and the C690S, C700S, and T467C mutations was cloned into the pTrc99a vector to generate pDT534. Another plasmid containing the same mutations, but with deletion of loop 1 (Δ430-439) was cloned to generate pDT535. MC4100 cells were transformed with pDT534 or pDT535. The resulting strains were grown overnight in LB supplemented with 50 μg/mL carbenicillin and 0.2% (w/v) glucose (37°C, 220 rpm). These overnight cultures were diluted 1:100 into 5 mL of fresh LB containing the same additives minus glucose, and were grown (37°C, 220 rpm) to OD600 ~0.4. Cells were then collected by centrifugation (4,200 x g, 10 min, 4°C) and resuspended in 2× SDS loading dye (containing β-mercaptoethanol). Each sample was divided into two, with one half boiled for 5 min and the other half remaining unboiled. All samples were analyzed by SDS-PAGE and Western blotting. 2×Strep-$BamA^S$ was detected using a monoclonal Strep-Tag II (HRP) antibody (EMD Millipore).

3.4.9. Expression and crosslinking of substrate-bound Bam complex for cryo-EM

The plasmid pJH114 was modified such that $BamA^M$ contained the C690S and C700S substitutions, in addition F804C, generating pDT517. 2×Strep-$BamA^S$ containing deletion of POTRA domains 3-5, deletion of loop 1, the C690S and C700S substitutions, and T467C, was cloned into the pBAD33 vector to generate pDT518. BL21(DE3) cells harboring pDT517 and pDT518 were grown overnight (37°C, 220 rpm) in LB supplemented with 50 µg/mL carbenicillin, 30 µg/mL chloramphenicol, and 0.2% glucose. This overnight culture was diluted 1:100 into 3 x 1.5 L of LB supplemented with 50 µg/mL carbenicillin and 30 µg/mL chloramphenicol. The resulting cultures were grown (37°C, 220 rpm) until an OD600 of ~0.7 was reached. At this point, the temperature was turned down to 30°C, and cells were allowed to continue shaking. After 20 min, protein expression was induced with isopropyl β-D-1-thiogalactopyranoside (IPTG, VWR) and L-(+)-arabinose (Alfa Aesar) at final concentrations of 0.2 mM and 0.1% (w/v), respectively. After three hours of additional shaking, cells were harvested via centrifugation (4,200 x g, 10 min, 4°C). Cell pellets were resuspended in PBS (20 mM NaH_2PO_4 pH 7.2, 150 mM NaCl) and incubated with 0.3 mM copper(II) sulfate (Acros Organics) and 0.3 mM 1,10-phenanthroline (Sigma-Aldrich) for 30 minutes. Cells were centrifuged (5,000 x g, 10 min, 4°C) and the pellets were frozen at -80°C prior to subsequent purification.

3.4.10. Purification of substrate-bound Bam complex

Cell pellets expressing Bam-substrate complex as described above were thawed and resuspended in buffer containing 20 mM Tris-HCl pH 8.0, 150 mM NaCl, 100 µg/mL lysozyme (Sigma-Aldrich), 1 mM PMSF (Sigma-Aldrich), 50 µg/mL DNase I (Sigma-Aldrich), and 2.5 mM $MgCl_2$ (Sigma-Aldrich). Cells were lysed using an Emulsiflex C3 (Avestin) at a pressure of 10,000 to 15,000 psi. After lysis, cell debris was removed via centrifugation (5,000 x g, 10 min, 4°C). Membrane fractions were isolated via ultracentrifugation using a 45 Ti rotor (Beckman Coulter) (37,000 rpm, 45 min, 4°C) and an Optima XE-90 ultracentrifuge (Beckman Coulter). The

membrane pellet was resuspended in buffer containing 20 mM Tris-HCl pH 8.0, 150 mM NaCl, 100 µg/mL lysozyme, and 1 mM PMSF. Membrane fractions were solubilized via incubation with 0.75% n-dodecyl-β-D-maltopyranoside (DDM, Anatrace) and 0.5% glyco-diosgenin (GDN, Anatrace) on a rocking platform (2 hours, 4°C). Unsolubilized material was then isolated via ultracentrifugation in a 70 Ti rotor (Beckman Coulter) (37,000 rpm, 30 min, 4°C). The supernatant, consisting of solubilized membrane proteins, was removed and supplemented with imidazole (pH 8.0) to a final concentration of 5 mM.

The supernatant was incubated with Ni-nitrilotriacetic acid (NTA) resin (Qiagen) that had been pre-washed with 10 CV buffer W1 (20 mM Tris-HCl pH 8.0, 150 mM NaCl, 10 mM imidazole pH 8.0, 0.02% GDN). After batch binding on a rocking platform (1 hour, 4°C), the resin was washed with 10 CV buffer W1. Elution was performed via addition of 5 CV buffer E1 (20 mM Tris-HCl pH 8.0, 150 mM NaCl, 200 mM imidazole pH 8.0, 0.02% GDN).

The Ni-NTA eluate was immediately incubated with Strep-Tactin XT Superflow resin (IBA Lifesciences) that had been pre-washed with 10 CV buffer W2 (100 mM Tris-HCl pH 8.0, 150 mM NaCl, 0.02% GDN). After batch binding on a rocking platform (1 hour, 4°C), the resin was washed with 15 CV buffer W2. Elution was performed via addition of 9 CV buffer E2 (100 mM Tris-HCl pH 8.0, 150 mM NaCl, 1 mM EDTA, 50 mM D-biotin, 0.02% GDN).

The Strep resin eluate was concentrated using an Amicon Ultra 4 mL 100 kDa molecular-weight cutoff centrifugal concentrator (EMD Millipore). The sample was then applied to an ÄKTA Pure (GE Healthcare Life Sciences) for purification via size-exclusion chromatography using a Superdex 200 Increase 10/300 GL column. The protein was eluted in buffer containing 20 mM Tris-HCl pH 8.0, 150 mM NaCl, and 0.02% GDN. After elution, protein corresponding to the center peaks of the chromatogram was concentrated to 5 mg/mL using an Amicon Ultra 0.5 mL 100 kDa molecular-weight cutoff centrifugal concentrator (EMD Millipore). A final yield of approximately 0.1 mg of complex per liter of bacterial culture could be obtained.

3.4.11. Bacterial strains

Bacterial strains used in the experiments reported here are provided in Table 3.1.

Table 3.1. Bacterial strains used in chapter 3

Strains	Genotype	Source
MC4100	F⁻ *araD139* Δ*(argF-lac)U169 rpsL150 relA1 flbB5301 deoC1 ptsF25 rbsR thi*	[173]
NR779	MC4100 *degP::cam*	[127]
NovaBlue	*endA1 hsd*R17 (r_K^-, m_K^+) *sup*E44 *thi*-1 *rec*A1 *gyr*A96 *rel*A1 *lac* F′[*pro*A⁺ B⁺ *lac*I^q ZΔM15::Tn*10*]	Novagen
BL21(DE3)	*fhuA2 [lon] ompT gal (λ DE3) [dcm] ΔhsdS*	Novagen

3.4.12. Bacterial plasmids

Bacterial plasmids used in the experiments here are provided in Table 3.2.

Table 3.2. Bacterial plasmids used in chapter 3

Plasmid	Description	Source
	pSup-BpaRS-6TRN	[183]
pDT175	pZS21/6×His-BamA(Δ430-439)	This study
pDT176	pZS21/6×His-BamA(Δ462-465)	This study
pDT177	pZS21/6×His-BamA(Δ495-504)	This study
pDT178	pZS21/6×His-BamA(Δ537-563)	This study
pDT179	pZS21/6×His-BamA(Δ601-607)	This study
pDT180	pZS21/6×His-BamA(Δ641-709)	This study
pDT181	pZS21/6×His-BamA(Δ745-766)	This study
pDT182	pZS21/6×His-BamA(Δ788-801)	This study
pDT201	pTrc99a/3×FLAG-BamA(Δ172-421/Δ430-439)	This study
pDT202	pTrc99a/3×FLAG-BamA(Δ172-421/Δ462-465)	This study
pDT203	pTrc99a/3×FLAG-BamA(Δ172-421/Δ495-504)	This study
pDT204	pTrc99a/3×FLAG-BamA(Δ172-421/Δ537-563)	This study
pDT205	pTrc99a/3×FLAG-BamA(Δ172-421/Δ601-607)	This study
pDT206	pTrc99a/3×FLAG-BamA(Δ172-421/Δ641-709)	This study
pDT207	pTrc99a/3×FLAG-BamA(Δ172-421/Δ745-766)	This study
pDT208	pTrc99a/3×FLAG-BamA(Δ172-421/Δ788-801)	This study
pDT209	pTrc99a/3×FLAG-BamA(Δ172-421)	This study
pDT260	pZS21/6×His-BamA(Δ430-439/Δ495-504)	This study
pDT261	pZS21/6×His-BamA(Δ430-439/Δ601-607)	This study
pDT262	pZS21/6×His-BamA(Δ430-439/Δ788-801)	This study
pDT263	pZS21/6×His-BamA(Δ495-504/Δ601-607)	This study
pDT264	pZS21/6×His-BamA(Δ495-504/Δ788-801)	This study
pDT265	pZS21/6×His-BamA(Δ601-607/Δ788-801)	This study
pDT340	pTrc99a/3×FLAG-BamA–BamB–BamC–BamD-8×His–BamE	Modified from [113]
pDT411	pZS21/6×His-BamA(Y491pBPA)	This study
pDT416	pZS21/6×His-BamA(Y531pBPA)	This study
pDT421	pZS21/6×His-BamA(M741pBPA)	This study
pDT422	pZS21/6×His-BamA(F440pBPA)	This study
pDT423	pZS21/6×His-BamA(K462pBPA)	This study
pDT424	pZS21/6×His-BamA(T467pBPA)	This study
pDT425	pZS21/6×His-BamA(S535pBPA)	This study
pDT426	pZS21/6×His-BamA(Y574pBPA)	This study
pDT427	pZS21/6×His-BamA(V611pBPA)	This study

pDT428	pZS21/6×His-BamA(L613pBPA)	This study
pDT429	pZS21/6×His-BamA(W635pBPA)	This study
pDT430	pZS21/6×His-BamA(F804pBPA)	This study
pDT436	pCDF/2×Strep-BamA(Δ172-421/Δ430-439/T467pBPA)	This study
pDT437	pCDF/2×Strep-BamA(Δ172-421/Δ495-504/T467pBPA)	This study
pDT438	pCDF/2×Strep-BamA(Δ172-421/Δ601-607/T467pBPA)	This study
pDT439	pCDF/2×Strep-BamA(Δ172-421/Δ788-801/T467pBPA)	This study
pDT451	pCDF/2×Strep-BamA(Δ172-421/Δ430-439/Y531pBPA)	This study
pDT452	pCDF/2×Strep-BamA(Δ172-421/Δ495-504/Y531pBPA)	This study
pDT453	pCDF/2×Strep-BamA(Δ172-421/Δ601-607/Y531pBPA)	This study
pDT454	pCDF/2×Strep-BamA(Δ172-421/Δ788-801/Y531pBPA)	This study
pDT471	pCDF/2×Strep-BamA(Δ172-421/Δ430-439/M741pBPA)	This study
pDT472	pCDF/2×Strep-BamA(Δ172-421/Δ495-504/M741pBPA)	This study
pDT473	pCDF/2×Strep-BamA(Δ172-421/Δ601-607/M741pBPA)	This study
pDT474	pCDF/2×Strep-BamA(Δ172-421/Δ788-801/M741pBPA)	This study
pDT476	pCDF/2×Strep-BamA(Δ172-421/Δ430-439/F804pBPA)	This study
pDT477	pCDF/2×Strep-BamA(Δ172-421/Δ495-504/F804pBPA)	This study
pDT478	pCDF/2×Strep-BamA(Δ172-421/Δ601-607/F804pBPA)	This study
pDT479	pCDF/2×Strep-BamA(Δ172-421/Δ788-801/F804pBPA)	This study
pDT504	pSup-BpaRS-6TRN(spec^{R})	Modified from [183]
pDT511	pZS21/6×His-BamA(E800C)	This study
pDT512	pZS21/6×His-BamA(F802C)	This study
pDT513	pZS21/6×His-BamA(F804C)	This study
pDT514	pTrc99a/3×FLAG-BamA(Δ172-421/Δ430-439/G460C)	This study
pDT515	pTrc99a/3×FLAG-BamA(Δ172-421/Δ430-439/T467C)	This study
pDT516	pTrc99a/3×FLAG-BamA(Δ172-421/Δ430-439/A469C)	This study
pDT517	pTrc99a/BamA(C690S/C700S/F804C)–BamB–BamC–BamD–BamE-8×His)	Modified from [113]
pDT518	pBAD33/2×Strep-BamA(Δ172-421/Δ430-439/T467C/C690S/C700S)	This study
pJL077	pZS21/6×His-BamA(S439pBPA)	This study
pSK476	pZS21/6×His-BamA	[73]

Chapter 4: Structure of a substrate-engaged Bam complex

This chapter is adapted/reproduced from:
Tomasek D, Rawson S, Lee J, Wzorek JS, Harrison SC, Li Z, Kahne D. Structure of a nascent membrane protein as it folds on the β-barrel assembly machine. *In review* (2020).

4.1. Introduction

To understand how the Bam complex accelerates folding of outer membrane proteins, we developed an approach to purify this machine bound to a stalled substrate (see Chapter 3). Here, we report a structure of the Bam complex folding BamA, determined by cryo-electron microscopy. The BamA catalyst ($BamA^{M}$, for $BamA^{machine}$) forms an asymmetric hybrid β-barrel with the BamA substrate ($BamA^{S}$). The N-terminal edge of $BamA^{M}$ has an extensive antiparallel interface with the C-terminal edge of $BamA^{S}$ by hydrogen bonding; the other edges of $BamA^{M}$ and $BamA^{S}$ are close to each other but curl inward and do not pair. Six hydrogen bonds in a membrane environment make the interface between the two proteins very stable. This stability allows C- to N-terminal directional folding without premature substrate release but creates a high kinetic barrier to release once folding has finished. Features at each end of the substrate may overcome the barrier and promote release by stepwise exchange of hydrogen bonds. This mechanism of substrate-assisted product release explains how the Bam complex can maintain a stable interface with the substrate during folding and then turn over rapidly when folding is complete.

4.2. Results

4.2.1. Structure determination of a substrate-engaged Bam complex

After successfully purifying a substrate-engaged Bam complex using $BamA^{S}$-ΔL1 as the substrate (see Chapter 3), we performed negative-stain electron microscopy experiments to confirm the homogeneity of the sample. We subsequently collected cryo-EM data using a Titan Krios cryo-electron microscope in collaboration with Dr. Zongli Li and Dr. Shaun Rawson at the Harvard Cryo-EM Center for Structural Biology at Harvard Medical School. 4,097 cryo-EM movies were obtained and approximately 2 million particles were picked (**Figure 4.1**). After rounds of 2D and 3D classification, 3D refinement was performed with a subset of approximately 223,000 particles. This refinement led to a structure of a substrate-engaged Bam complex, with $BamA^{S}$-

ΔL1 as the substrate, to an overall resolution of 4.1 Å (**Table 4.1 and Figure 4.2**). The cryo-EM map contains density for all five components of the Bam complex and for the substrate (**Figure 4.3**). All β-strands of the $BamA^{M}$ β-barrel were resolved, as were all β-strands of the $BamA^{S}$ β-barrel except β1 and β2. Extracellular loop 1, which we removed in the substrate, would normally connect these two β-strands. The two POTRA domains that were included in $BamA^{S}$ could not be resolved.

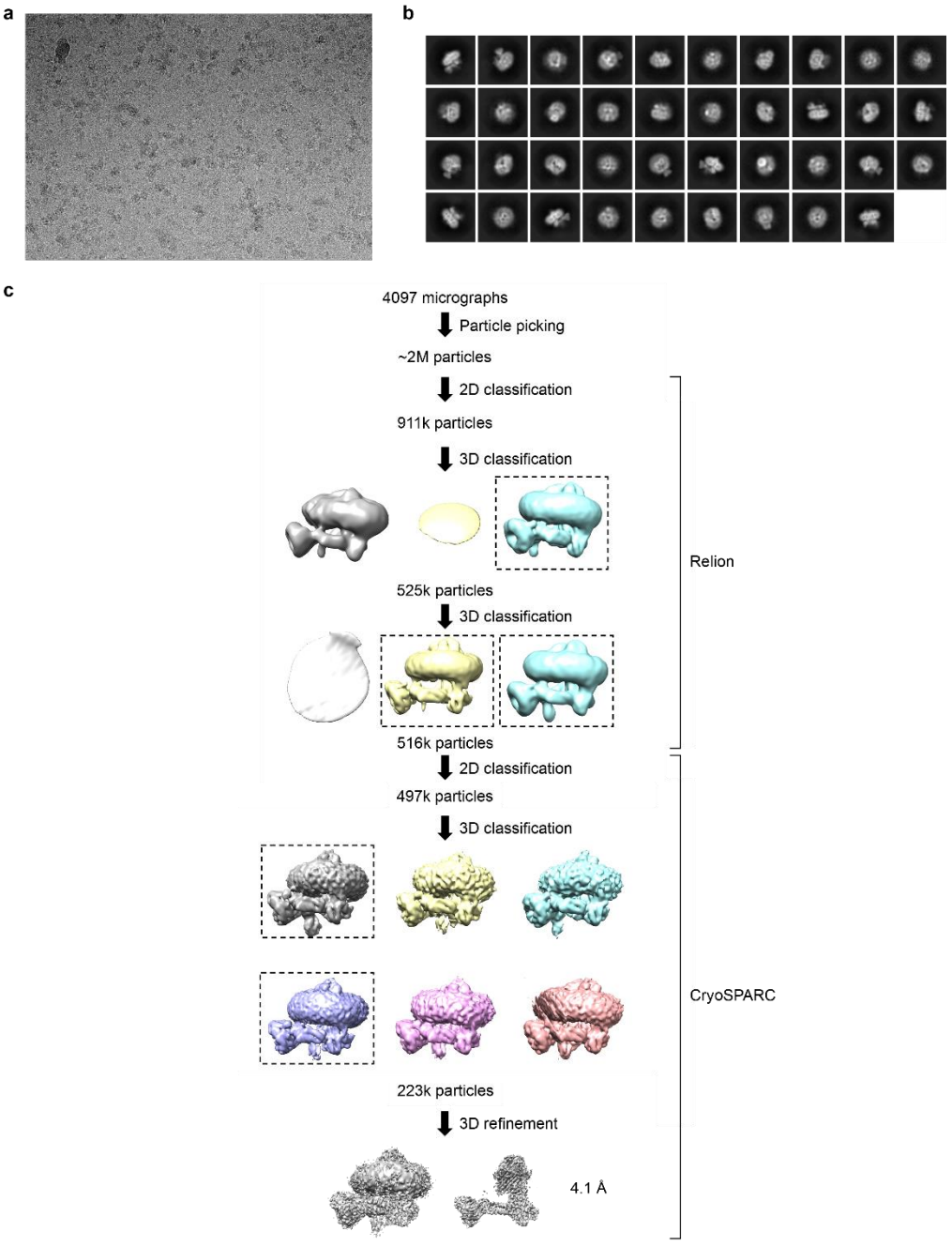

Figure 4.1. Cryo-EM data processing scheme
a, Representative cryo-EM micrograph of the substrate-bound Bam complex embedded in vitreous ice. **b**, Selected two-dimensional class averages of cryo-EM particle images. **c**, Scheme of three-dimensional classification and refinement of cryo-EM particle images.

Table 4.1. Cryo-EM data collection, refinement and validation statistics for 4.1 Å structure

	Substrate-bound Bam complex (EMD-20969)	
Data collection and processing		
Magnification	58,717	
Voltage (kV)	300	
Electron exposure (e^{-}/Å^2)	70	
Defocus range (μm)	1.1 to 2.8	
Pixel size (Å)	0.85	
Symmetry imposed	C1	
Initial particle images (no.)	2,054,956	
Final particle images (no.)	223,353	
Map resolution (Å)	4.1	
FSC threshold	0.143	
Refinement		
Initial model used (PDB code)	5D0O, 5D0Q, 2YH3	
Map sharpening *B* factor (Å^2)	-119	
	Without rigid body-docked components	With rigid body-docked components (PDB 6V05)
Model composition		
Non-hydrogen atoms	12198	13712
Protein residues	1555	1884
Ligands	0	0
B factors (Å^2)		
Protein	49.53	57.54
R.m.s. deviations		
Bond lengths (Å)	0.004	0.004
Bond angles (°)	0.937	0.953
Validation		
MolProbity score	1.61	1.74
Clashscore	7.74	9.22
Poor rotamers (%)	0.39	0.75
Ramachandran plot		
Favored (%)	96.86	96.34
Allowed (%)	3.14	3.61
Disallowed (%)	0.00	0.05
Map-model CC	0.80	0.79

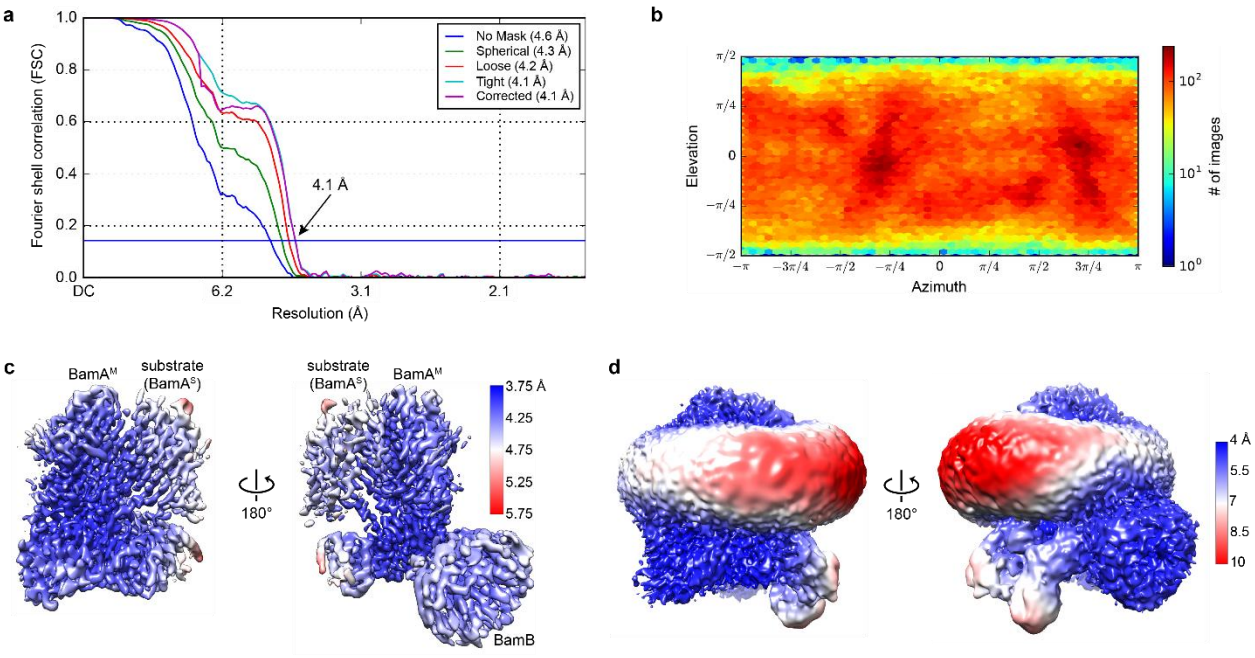

Figure 4.2. Cryo-EM data analysis
a, Gold-standard Fourier shell correlation (FSC) curves calculated with different masks in cryoSPARC. The resolutions were determined at FSC = 0.143 (horizontal blue line). The final corrected mask gave an overall resolution of 4.1 Å. **b**, Distribution of orientations over azimuth and elevation angles for particles included in the calculation of the final map. **c**, Cryo-EM map colored by local resolution. **d**, As in **c**, but presented at a lower contour level to allow observation of low-resolution features.

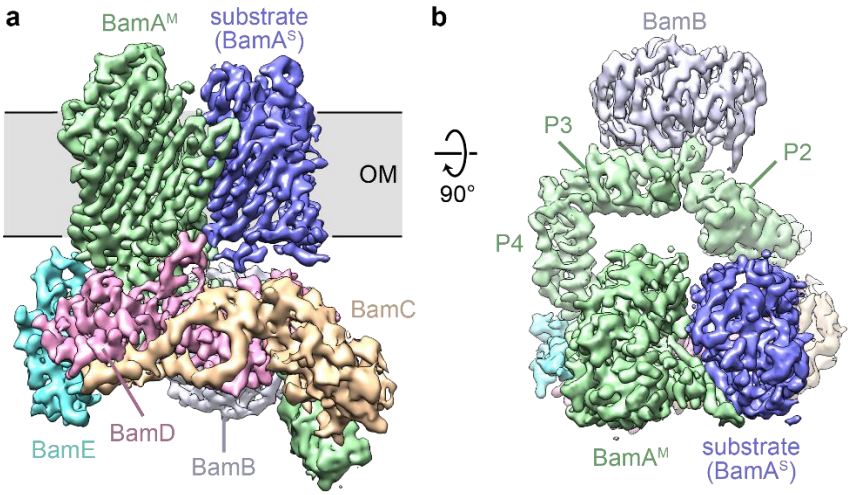

Figure 4.3. Cryo-EM map of substrate-engaged Bam complex
a, Side view of the cryo-EM reconstruction of the Bam complex bound to the substrate (BamA^S-ΔL1). BamA^M is colored in green, BamB in silver, BamC in wheat, BamD in pink, BamE in cyan, and the substrate in blue. The approximate boundaries of the outer membrane (OM) are shown. BamC and POTRA domains 1 and 2 of BamA^M are shown at a lower contour level than the rest of the structure. **b**, Top-down view turned 90 degrees from view in **a**. P2, P3, and P4 refer to the POTRA domains of BamA^M that are visible in this view.

4.2.2. Overall conformation of the complex

Using previously published X-ray crystallographic structures, we built an atomic model into our cryo-EM map (**Figure 4.4a**). For BamC and POTRA domains 1 and 2 of $BamA^{M}$, the density was poor; therefore, we docked existing structures into the corresponding density without additional refinement. The fit of the side chains into the map is reasonable for the 4.1 Å resolution obtained (**Figure 4.5**). The overall architecture of the Bam complex in its substrate-bound state is similar to that seen for previously reported substrate-free states, with the organization of the periplasmic components most closely resembling that of substrate-free structures containing a laterally-open form of BamA (**Figure 4.6**). Structures of the individual components are likewise similar to those already published (**Figure 4.7**).

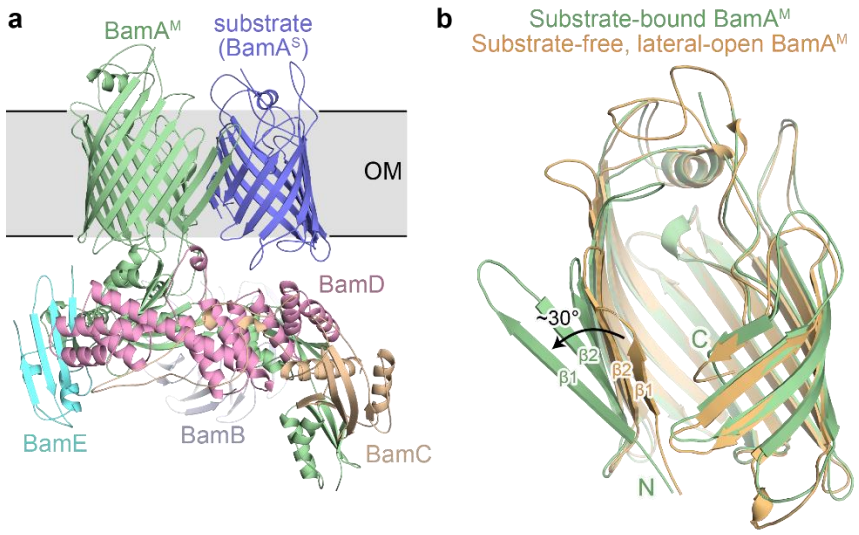

Figure 4.4. Atomic model of substrate-engaged Bam complex and interactions between $BamA^{M}$ and $BamA^{S}$
a, Side view of the atomic model. **b**, Overlay of $BamA^{M}$ from our substrate-bound complex (green) and a substrate-free, lateral-open complex (orange, PDB ID: 5EKQ).

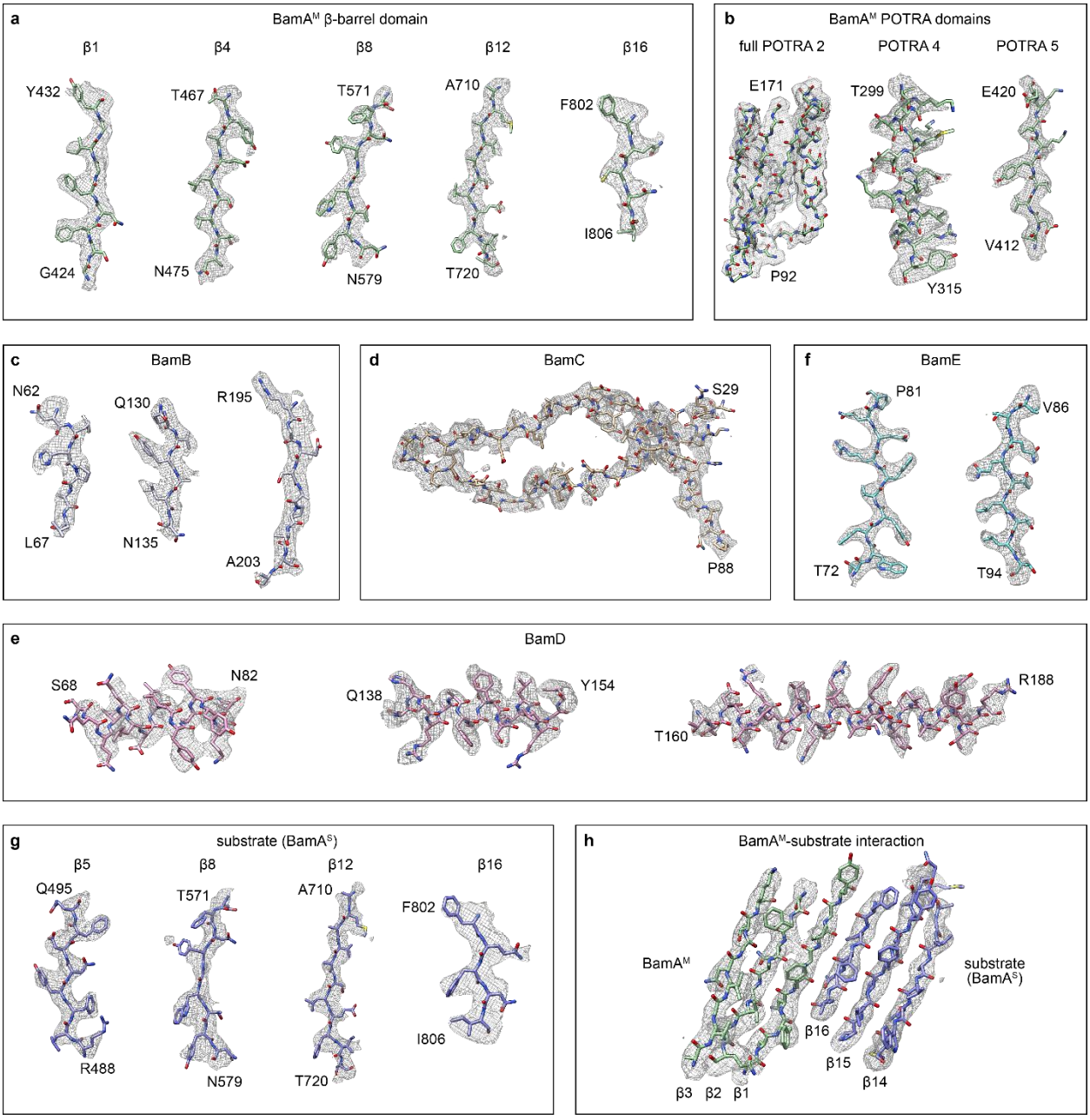

Figure 4.5. Fit of the atomic model into the cryo-EM map
The atomic model (in stick representation) shown with the corresponding portion of the cryo-EM map (shown in gray mesh) for selected regions in **a**, the BamA^{M} β-barrel domain; **b**, the BamA^{M} POTRA domains; **c**, BamB; **d**, BamC; **e**, BamD; **f**, BamE; **g**, the substrate; and **h**, the BamA^{M}-substrate interaction. Side-chain densities are visible, and individual β-strands can be resolved. Images were prepared in UCSF Chimera using a 2 Å carve radius.

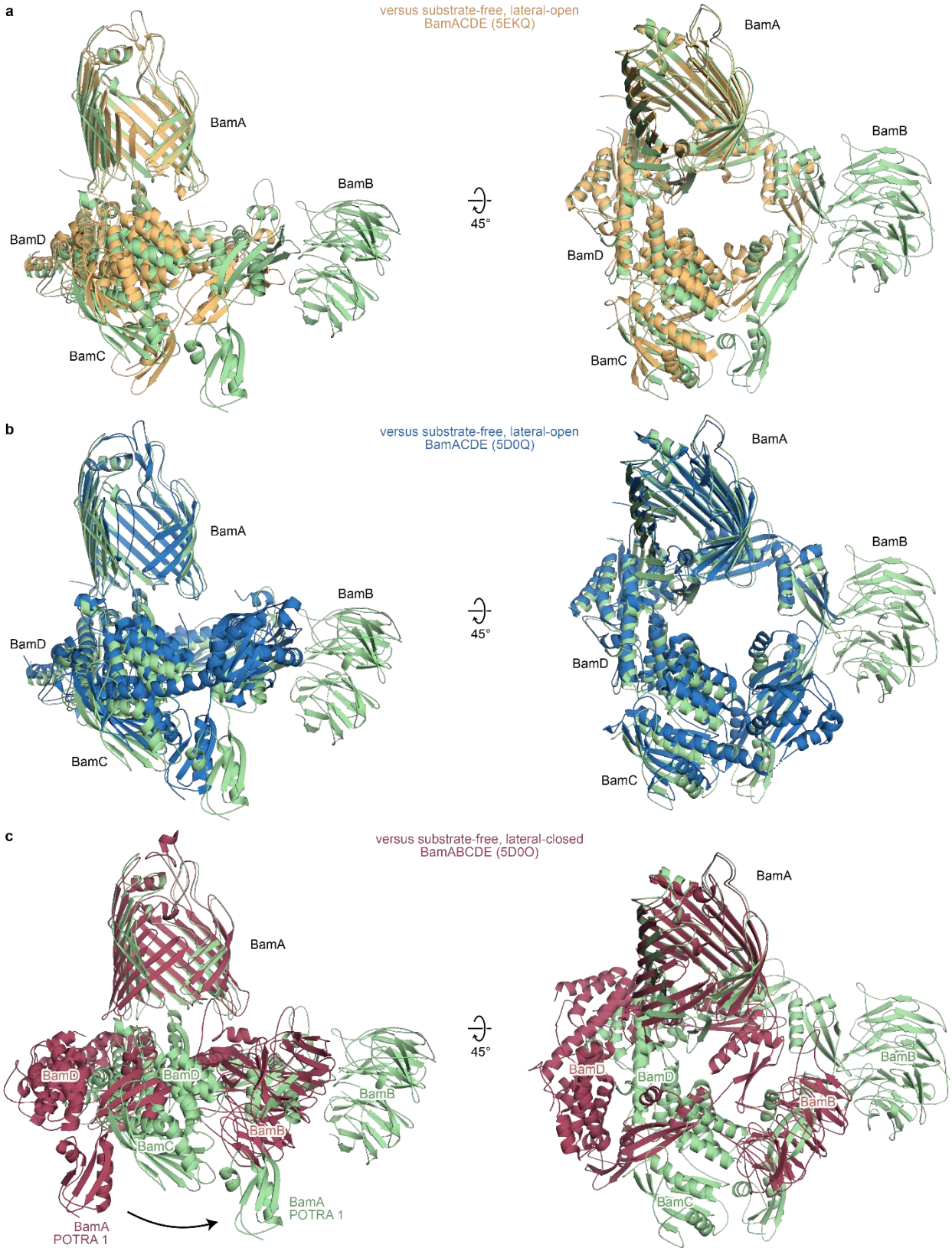
a
versus substrate-free, lateral-open
BamACDE (5EKQ)
BamA
BamB
BamD
BamC
45°
BamA
BamB
BamD
BamC
b
versus substrate-free, lateral-open
BamACDE (5D0Q)
BamA
BamB
BamD
BamC
45°
BamA
BamB
BamD
BamC
c
versus substrate-free, lateral-closed
BamABCDE (5D0O)
BamA
BamD
BamD
BamB
BamB
BamC
BamA
POTRA 1
BamA
POTRA 1
45°
BamA
BamD
BamD
BamB
BamB
BamC

Figure 4.6. Comparison of the architecture of the substrate-bound Bam complex to that of substrate-free Bam complexes

a-c, Alignments of substrate-bound Bam complex (all components shown in green) with published Bam complex structures (shown in different colors). Alignments are performed using the β-barrel domains of BamA.

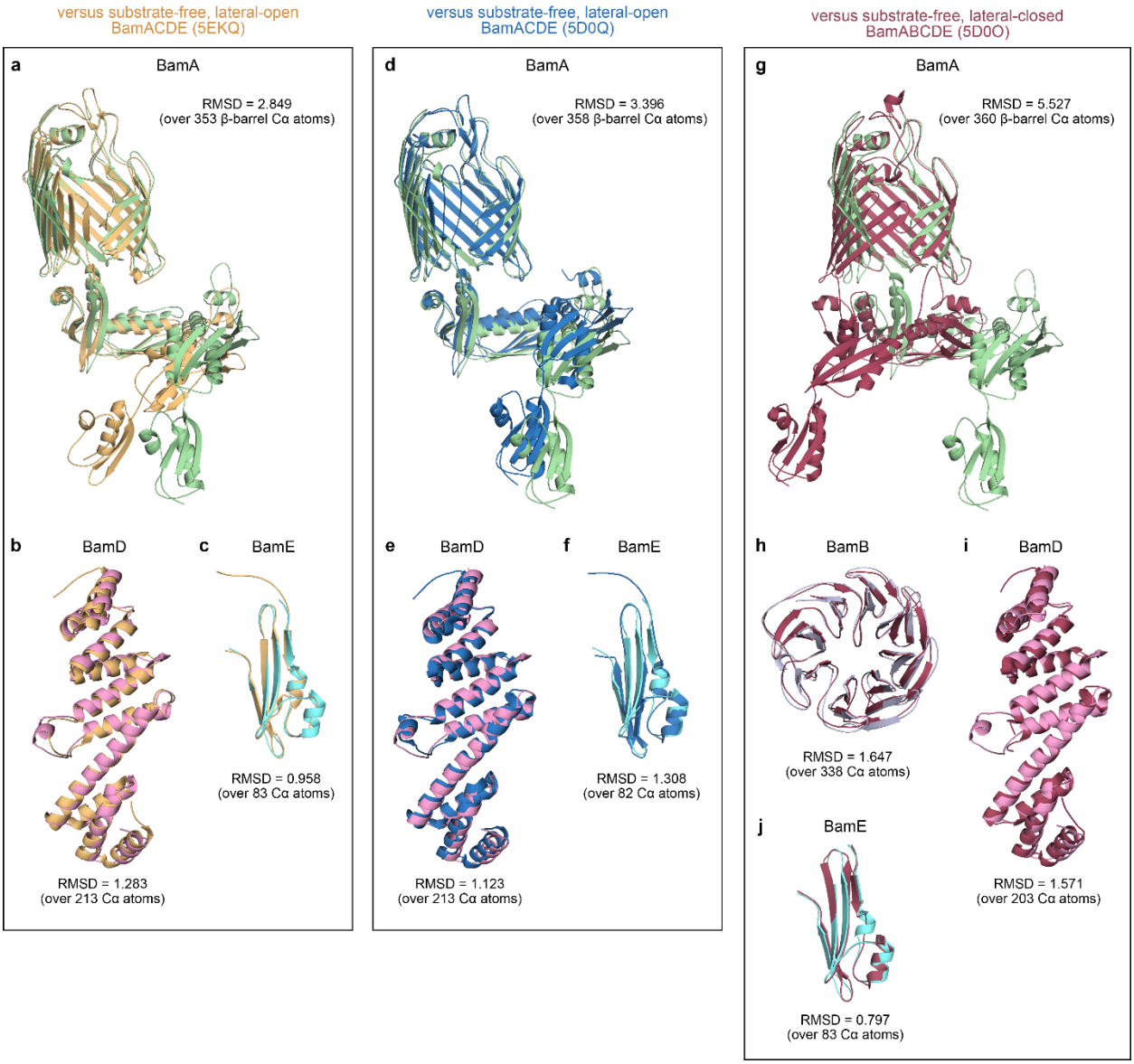

Figure 4.7. Comparison of conformations of Bam complex components from substrate-engaged and substrate-free complexes

Alignments of atomic models of **a**, BamA^{M}; **b**, BamD; and **c**, BamE from our substrate-bound complex with the corresponding components from a substrate-free, lateral-open Bam complex (light orange, PDB ID: 5EKQ). **d**-**f**, as in **a**-**c**, but using a different substrate-free, lateral-open Bam complex (blue, PDB ID: 5D0Q). Alignments of atomic models of **g**, BamA^{M}; **h**, BamB; **i**, BamD; and **j**, BamE from our substrate-bound complex with the corresponding components from a substrate-free, lateral-closed Bam complex (dark red, PDB ID: 5D0O). Alignments of BamA^{M} are performed using the β-barrel domains. Alignments of BamB are not shown for complexes in which BamB is absent. Alignments of BamC are not shown since BamC in our structure was obtained by docking BamC as a rigid body from PDB ID: 5D0Q. The RMSD value for each alignment, obtained with PyMOL, is indicated.

4.2.3. BamA undergoes further lateral opening in the substrate-engaged state

There are notable conformational differences between the N-terminal segment of the $BamA^{M}$ β-barrel in our substrate-engaged complex and the corresponding segment in any of the substrate-free, lateral-open structures (**Figure 4.4b**). In particular, the first and second β-strands of $BamA^{M}$ in the substrate-engaged β-barrel bend outward approximately 30 degrees, separating them more from the C-terminal strand than they are in the substrate-free β-barrel. This creates an extremely lateral-open β-barrel in the membrane. Smaller deviations in strands 3, 4, and 5 allow the larger changes in the first two strands, but all remaining strands superimposed well on previous structures of BamA (**Figure 4.7**).

4.2.4. The substrate accumulates on the Bam complex in a largely folded state

Like the $BamA^{M}$ β-barrel, the $BamA^{S}$ β-barrel is embedded in the membrane and has a lateral opening between its C-terminal β-strand and the most N-terminal resolvable β-strand (β3). Its location in the membrane is consistent with our findings that $BamA^{S}$-ΔL1 is insensitive to periplasmic proteases and urea extraction. β1 and β2 of the substrate could not be resolved, presumably because deletion of loop 1, which would normally connect these two β-strands (**Figure 4.8**), prevented proper formation of the β-hairpin. Additionally, the POTRA domains of the substrate could not be resolved, likely due to their flexibility in the absence of bound BamBCDE lipoproteins. However, the remainder of the substrate β-barrel resembles the folded form of BamA. The similarity of strands 3 through 16 between the trapped substrate and folded wild-type BamA suggests that the former represents an on-pathway folding intermediate.

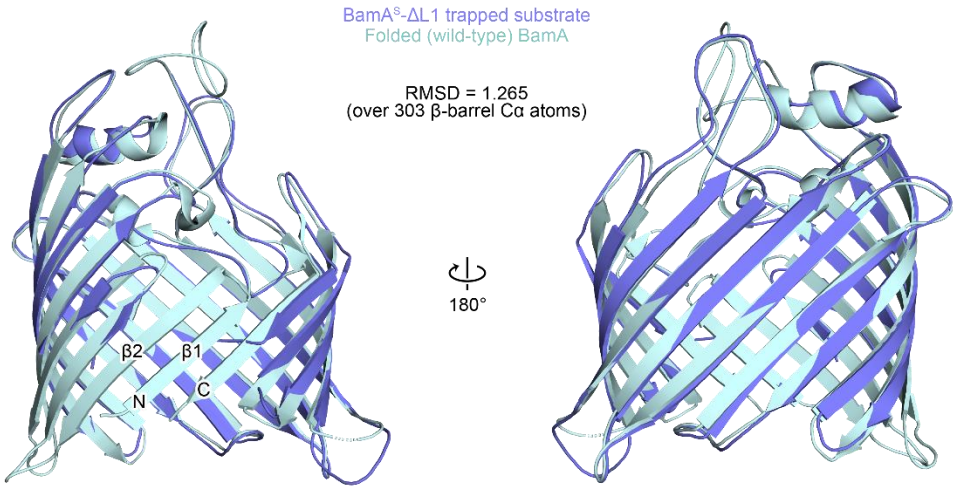

Figure 4.8. The BamA^S-ΔL1 substrate accumulates on the Bam complex in a largely folded state
Alignment of the atomic model of the BamA^S-ΔL1 substrate (blue) with that of folded wild-type BamA (cyan, PDB ID: 4N75). The first two β-strands of the substrate in our structure could not be resolved (β1 and β2, labelled on the structure of wild-type BamA).

4.2.5. The N-terminal β-strand of BamA interacts with the substrate

The N-terminal strand of BamA^M and the C-terminal strand of BamA^S are paired in an antiparallel interaction that involves six hydrogen bonds (**Figure 4.9**). The observed outward rotation of the N-terminal β-strands of BamA^M permits its interactions with the C-terminal β-strand of the substrate. This finding corroborates previous observations that this region of the machine interacts directly with substrates[143,144,174]. The hydrogen bonding network is the main contact between the Bam complex and the substrate at this late stage of assembly and explains how the machine can stably associate with the substrate during folding without releasing it prematurely.

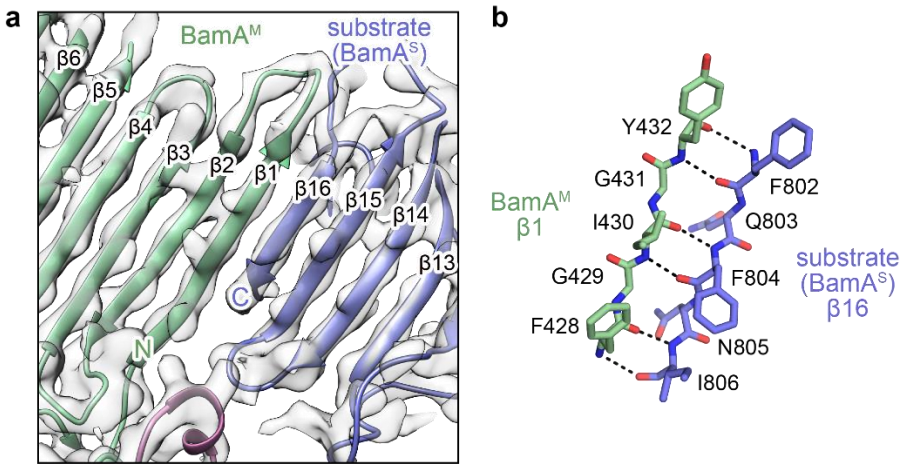

Figure 4.9. Interactions between the N-terminal β-strand of BamA^{M} and the C-terminal β-strand of the substrate
a, View of the atomic model from the same orientation as in Figure 4.4a. The corresponding portion of the cryo-EM map is shown in gray. **b**, Interactions between β1 of BamA^{M} and β16 of the substrate shown in stick representation. Hydrogen bonds between the β-strands are shown as dashed black lines.

4.2.6. Asymmetric interactions between BamA and the substrate

Whereas the C-terminal β-strand of BamA^{S} is tightly held by the N-terminal β-strand of BamA^{M} (**Figure 4.9**), the β-strands at the other ends of BamA^{S} and BamA^{M} are not hydrogen bonded (**Figure 4.10**). Instead, these edges of the proteins (the C-terminal edge of BamA^{M} and the N-terminal edge of BamA^{S}) curve inwards, so that the barrels make contact along their exterior surfaces that would typically contact membrane lipids. Although the result is an unpaired edge at one end of both BamA^{M} and BamA^{S}, the inward curvature ensures that these edges face into the lumen of the hybrid barrel, where they are solvated by water. The seal created by the hydrophobic interaction interface between the exterior surfaces of BamA^{M} and BamA^{S} could prevent entry of lipid molecules into the lumen. In Chapter 3, crosslinking experiments were reported using pBPA substitutions at outward-facing residues in the substrate (**Figure 3.6**). Those earlier results agree with the finding from our structure that most of the exterior surface of BamA^{S}-ΔL1 faces the membrane, while the portion of this substrate closest to the N-terminus interacts with BamA^{M}.

We did not observe any direct contacts between the substrate and the BamBCDE lipoproteins of the machine, although the final periplasmic loops of the substrate are in close

proximity to BamD. However, it is possible that substrates stalled at earlier stages of folding do contact these accessory components of the Bam complex, and this remains to be examined.

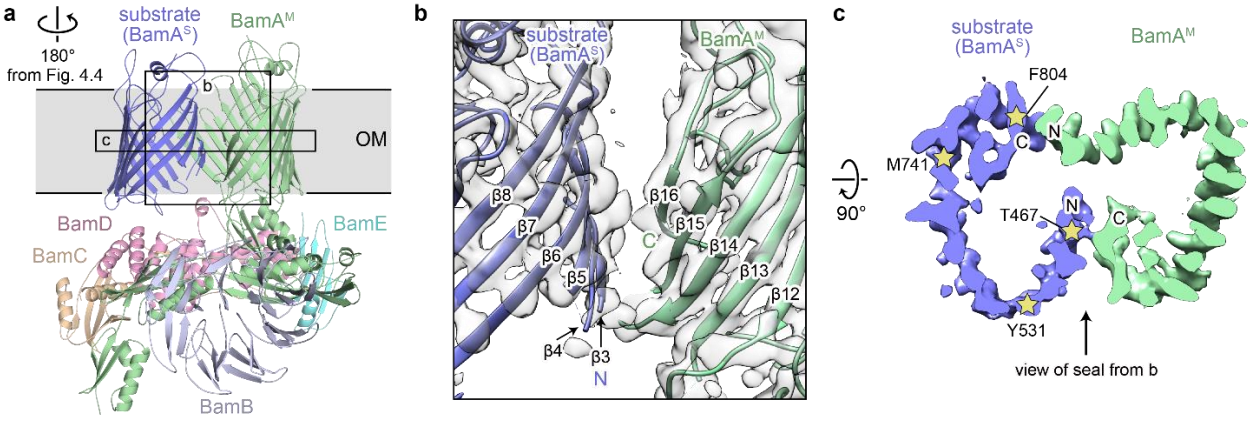

Figure 4.10. Interactions between BamA^M and BamA^S are asymmetric
a, Side view turned 180 degrees from view in Figure 4.4 to show interactions between the C-terminus of BamA^M and the N-terminus of the substrate. **b**, Enlarged view of boxed region from **a**. The corresponding portion of the cryo-EM map is shown in gray. This view shows the "seal" formed by the interaction of the C-terminal edge of BamA^M and the N-terminal edge of the substrate. **c**, Top-down view of a slice through the cryo-EM map, as indicated in **a**. The N- and C-termini of BamA^M and the substrate are indicated. The approximate positions of residues substituted with pBPA in Chapter 3 are indicated with stars.

4.2.7. Using the engineered disulfide bond does not affect the structure of the complex

Given that we used a disulfide bond to tether the C-terminal region BamA^M to the N-terminal region of BamA^S to stabilize the complex prior to purification, it was possible that the asymmetric interface we observe in our structure reflected a non-native conformation of the complex. To rule out this possibility, we sought to determine another structure by cryo-EM with the same substrate but using a cysteine-cysteine crosslink introduced between the other ends of BamA^M and BamA^S. We reasoned that, if the cysteine tether induced a non-native conformation, a tether on the opposite side of the complex should result in a structure with a different architecture of the BamA^M-BamA^S interface.

First, we first identified cysteines at the N-terminal end of BamA^M and the C-terminal end of BamA^S-ΔL1 that could form cysteine-cysteine crosslinks mediated by the chemical crosslinker 1,2-bis(maleimido)ethane (BMOE) (**Figure 4.11a-c**). Based on these results, E800C was selected for use within BamA^S-ΔL1 for cryo-EM, while S439C was introduced into BamA^M. Note

that this cysteine substitution still allows for proper, but slow, folding of the substrate (**Figure 4.11d**). We then purified the complex with this alternative cysteine pair (**Figure 4.11e-f)** and determined its structure by cryo-EM to 6.5 Å resolution (**Figure 4.11g-h and Table 4.2**).

This 6.5 Å structure is remarkably similar to our 4.1 Å structure in terms of the architecture of the complex (**Figure 4.12**). Therefore, the presence of the disulfide bond in the main (4.1 Å) structure did not introduce a non-native conformation into the complex. Our strategy to purify a substrate-bound Bam complex using cysteine tethers is thus valid and could be used in future studies in which other complexes are purified. For simplicity, all further discussions of the structure below will refer to the 4.1 Å structure, since the atomic model was generated using that map.

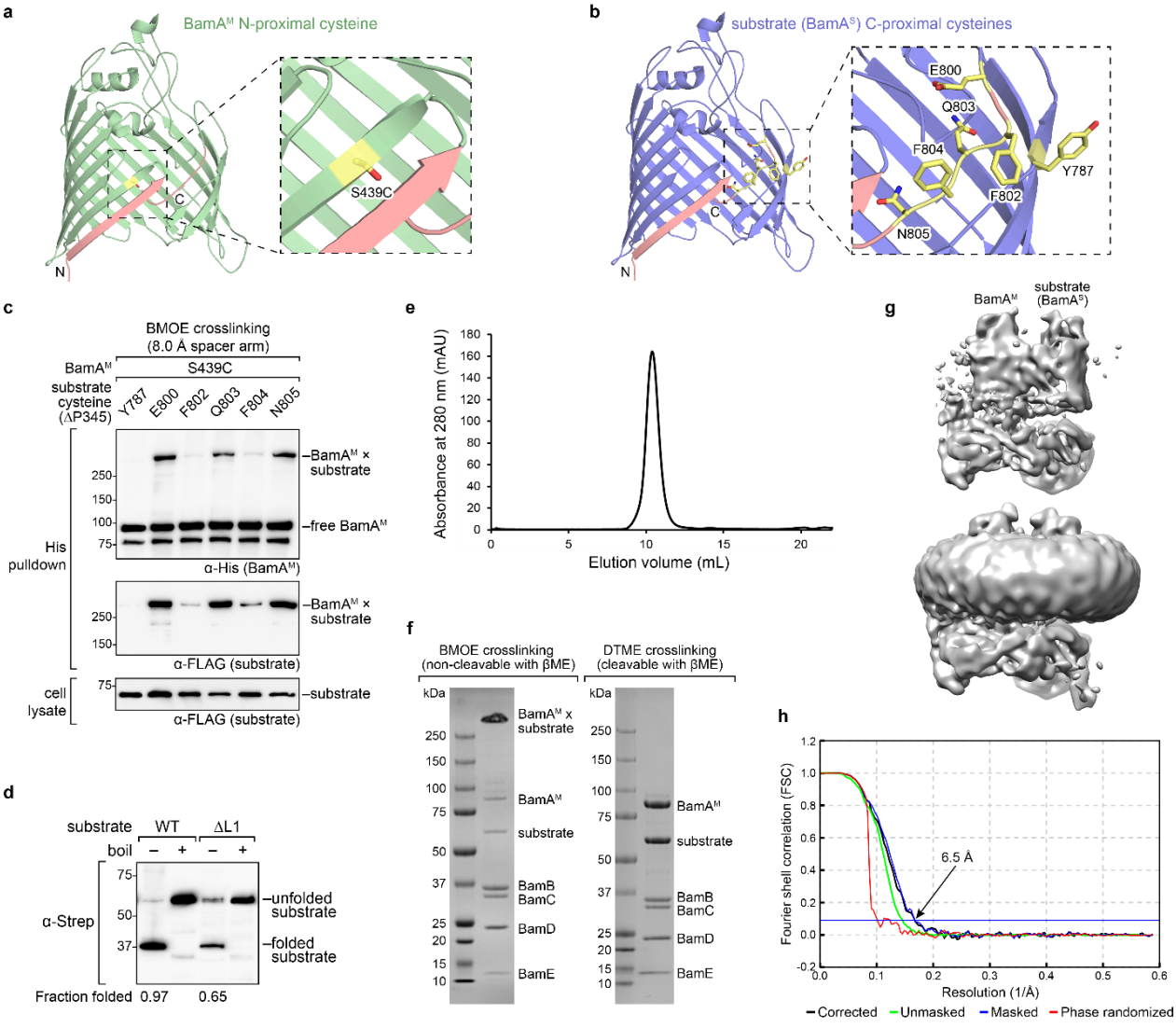

Figure 4.11. Low resolution cryo-EM structure of a substrate-bound Bam complex with an alternative cysteine crosslink

a, Residue near the N-terminus of the β-barrel domain of $BamA^{M}$ that was substituted with cysteine (yellow stick) shown on the structure of the β-barrel domain of BamA (green, PDB ID: 5D0O). The N- and C-terminal β-strands are labelled and shown in salmon. **b**, As in **a**, but showing residues near the C-terminus of the β-barrel domain of the substrate ($BamA^{S}$-ΔL1, blue) that were substituted with cysteine. **c**, Cysteine crosslinking between 6×His-tagged $BamA^{M}$ and 3×FLAG-tagged $BamA^{S}$-ΔP345-ΔL1 containing cysteine substitutions shown in **a** and **b**. The crosslinker 1,2-bis(maleimido)ethane (BMOE), which has an 8.0 Å spacer arm, was used. The presence of a Bam-substrate crosslink is detected as a high molecular weight adduct on α-His (top) and α-FLAG (middle) immunoblots. α-FLAG immunoblot of total cell lysates (bottom) shows that the substrates containing different cysteine substitutions are expressed at similar levels. **d**, Assessment of heat modifiability of substrate used for cryo-EM (ΔL1, right half of blot) that contains the E800C mutation and a substrate that contains loop 1 but is otherwise identical (WT, left half of blot). The samples in the unboiled lanes can be used to calculate the fraction of each substrate that is folded *in vivo* (indicated below the blot). The substrate containing the E800C mutation retains the ability to fold *in vivo*. **e**, Representative size-exclusion chromatogram of the substrate-bound Bam complex in which cysteine crosslink formation between the C-terminus of the substrate ($BamA^{S}$-ΔL1) and the N-terminus of $BamA^{M}$ was induced with BMOE. **f**, SDS-PAGE gels showing peak fractions from size-exclusion chromatography. The left gel shows the complex from **e** with addition of βME, which cannot break the crosslink formed by BMOE. The right gel shows the complex crosslinked with dithiobismaleimidoethane (DTME), which is similar to BMOE, but has a 13.3 Å spacer arm and is cleavable in the presence of βME to visualize the individual components of the complex. **g**, Cryo-EM map obtained using the $BamA^{M}$(S439C) and $BamA^{S}$-ΔL1(E800C) cysteine pair crosslinked with BMOE. The two images show the same view at different contour levels. $BamA^{M}$ and the substrate are labelled. **h**, Gold-standard Fourier shell correlation (FSC) curves for low resolution cryo-EM structure calculated in Relion. The final corrected mask gave an overall resolution of 6.5 Å at FSC = 0.143 (horizontal blue line).

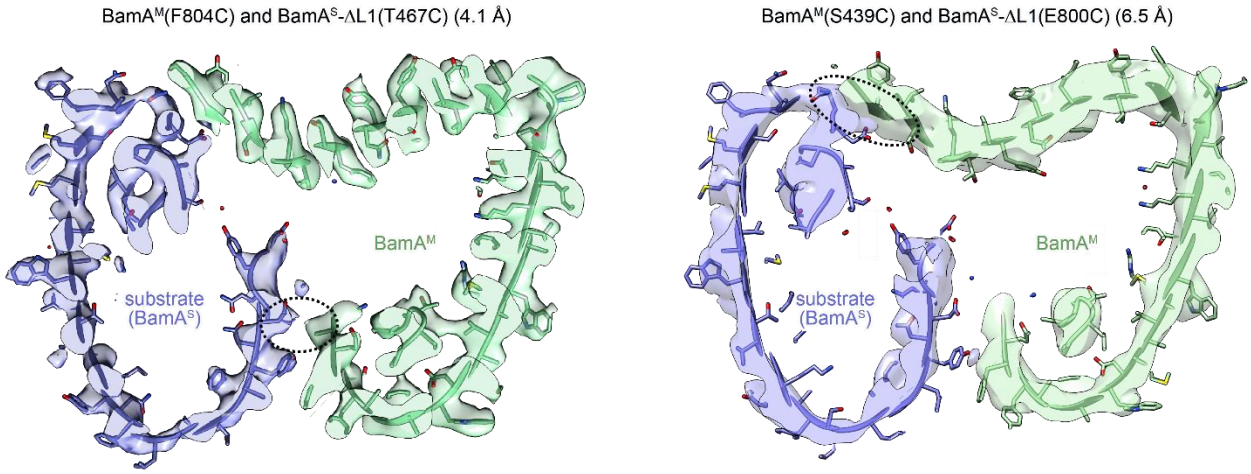

Figure 4.12. Comparison of the 4.1 Å structure to the 6.5 Å structure containing the alternative cysteine pair in $BamA^{M}$ and the substrate

In the left panel, a top-down slice through the 4.1 Å cryo-EM map is shown (similar to the slice in Fig. 3c) and is overlaid with the atomic model generated from that map. In the right panel, a similar slice through the 6.5 Å cryo-EM map is shown and is overlaid with the atomic model generated from the 4.1 Å map after fitting of the atomic model into the 6.5 Å map as a rigid body. The dashed oval represents the approximate location of the cysteine tether introduced into each complex.

Table 4.2. Cryo-EM data collection, refinement and validation statistics for 6.5 Å structure

	Substrate-bound Bam complex with alternative cysteine pair (EMD-21313)
Data collection and processing	
Magnification	58,717
Voltage (kV)	300
Electron exposure (e^{-}/Å^2)	55
Defocus range (μm)	1.1 to 3.2
Pixel size (Å)	0.85
Symmetry imposed	C1
Initial particle images (no.)	690,143
Final particle images (no.)	233,064
Map resolution (Å)	6.5
FSC threshold	0.143
Map sharpening *B* factor (Å^2)	-277

4.2.8. Intramolecular interactions between the ends of the substrate trigger release from the Bam complex

The six hydrogen bonds at the interface between the N-terminal β-strand of the BamA^M β-barrel and the C-terminal β-strand of the BamA^S β-barrel (**Figure 4.9**) prevent premature product release, but the strength of that interaction raises the question of how release occurs at all. In a membrane, disrupting a hydrogen bond costs approximately 4 kcal/mol[185,186]. Thus, the kinetic barrier to disrupting all six hydrogen bonds at once would be insurmountably high, even though the overall process is thermodynamically favorable.

Two unexpected features of the structure suggest a mechanism for substrate β-barrel closure and release. First, the edges of the two barrels at the folding end of the substrate do not engage in hydrogen bonding but instead turn inward toward the water-filled lumen of the hybrid barrel. Thus, no polar bonds to the machine need to be broken at that edge. Second, the register of the hydrogen bonds between the N-terminal β-strand of the BamA^M β-barrel and the C-terminal β-strand of the BamA^S β-barrel is not the same as it is in a fully closed BamA β-barrel[133,134]. The hydrogen-bonded residues on the BamA^M side are largely the same as those at the seam of a closed β-barrel, but they interact with different residues in the C-terminal β-strand of the substrate.

In the catalytic complex, BamA^{M} splays open at the extracellular end of its N-terminal β-strand (**Figure 4.4**), thereby facilitating pairing with residues in the C-terminal β-strand of BamA^{S} that precede those paired with the N-terminal β-strand in the fully folded structure (compare **Figure 4.9b and Figure 4.13a**). This pairing creates an overhang at the end of the BamA^{S} C-terminal β-strand by steric clash with the loop between β-strands 14 and 15 of BamA^{S} (**Figure 4.13b-c**). The overhang, together with spatial proximities shown by the structure, suggests a stepwise model for how release occurs (**Figure 4.14**). When the N-terminal β-strand of BamA^{S} adds to the substrate β-barrel, it approaches the C-terminal overhang. The N- and C-termini of the substrate begin to form hydrogen bonds, sequentially disrupting bonds between the substrate and the machine. Unlike concerted disruption, sequential displacement is energetically feasible. As hydrogens bonds form to close the substrate β-barrel, the C-terminal β-strand of the substrate peels away stepwise from its interaction with the N-terminal β-strand of the BamA^{M} β-barrel. The accompanying relaxation of strain in the splayed N-terminal hairpin of BamA^{M} might accelerate the strand exchange.

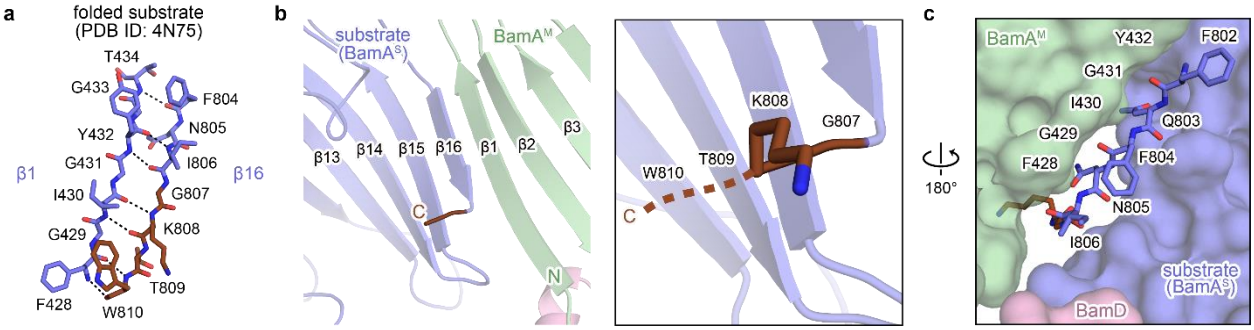

Figure 4.13. The C-terminal residues of the trapped substrate form an overhang
a, Interactions between β1 and β16 in the folded form of BamA shown in stick representation (PDB ID: 4N75). **b**, View of the interactions between the C-terminal end of the substrate and the N-terminal end of BamA^{M} (left). The view is from the interior of the two β-barrels. The residues belonging to the C-terminal overhang of the substrate are shown in brown. The final two residues of the substrate could not be resolved, and their expected location is indicated with a dashed brown line in the inset (right). **c**, View of BamA^{M}-substrate interactions turned 180 degrees from view in **b**. Proteins are shown in semi-transparent surface representation, except for the C-terminal β-strand and overhang of the substrate, which are shown in stick representation. Residues in the N-terminal β-strand of BamA^{M} and the C-terminal β-strand of the substrate are labelled.

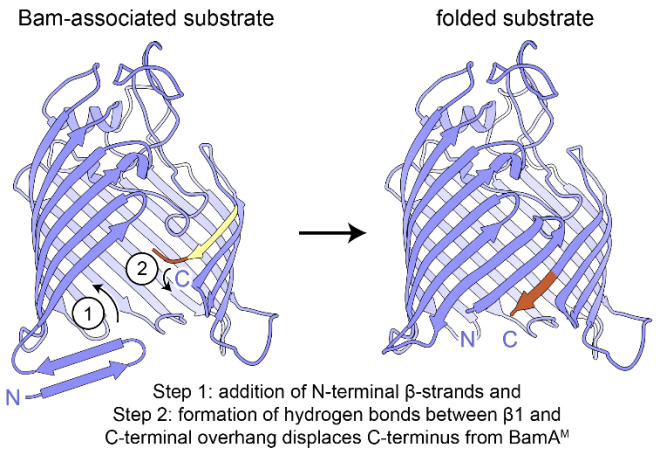

Figure 4.14. Model of the final step in substrate assembly
Release involves pairing between the C-terminal overhang and N-terminal β-strand of the substrate to effect release. The residues belonging to the C-terminal overhang of the substrate are shown in brown, and the residues that interact with the N-terminal β-strand of BamA^{M} in the Bam-associated substrate are shown in yellow.

We tested the implication of the strand-exchange mechanism as follows. First, we ruled out a special role for an invariant glycine present in BamA and other members of the Omp85 superfamily in mitochondria and chloroplasts (Gly807 in *E. coli*), just at the boundary of the C-terminal overhang in our structure (**Figure 4.15a**). Changing G807 to valine has been reported to cause outer membrane defects[187]. We made the G807V mutation to determine whether it prevents release of an otherwise wild-type BamA^{S} but found no defects in its assembly (**Figure 4.15b-d**), as judged by DegP sensitivity and by crosslinking. These results indicate that the G807V substitution alters BamA function and not its assembly.

Second, we found by testing other substitutions at position 807 that only proline had an effect. Proline distorts the conformation and hydrogen bonding in a β-sheet. We made substrate variants with proline substitutions at positions 807, 808 or 809 in the overhanging C-terminal segment to test if the assembly defect was position specific. DegP did not degrade these proline-substituted substrates (**Figure 4.15b-c**), indicating that the substrates are membrane integrated and protected from degradation. *In vivo* photocrosslinking yielded crosslinks from BamA^{M} to the proline-substituted variants (**Figure 4.15d**), showing that these substrates had accumulated on the machine in a membrane-integrated state. The crosslinks we found with proline mutations were

similar to those for substrates containing deletions of N-proximal loop sequences (**Figure 3.4**). Moreover, deleting the C-terminal residues (809-810) also caused late-stage assembly defects (**Figure 4.15b-d**). These results suggest that forming hydrogen bonds between the N-terminal β-strand and the C-terminal overhang facilitates release of $BamA^S$ from $BamA^M$.

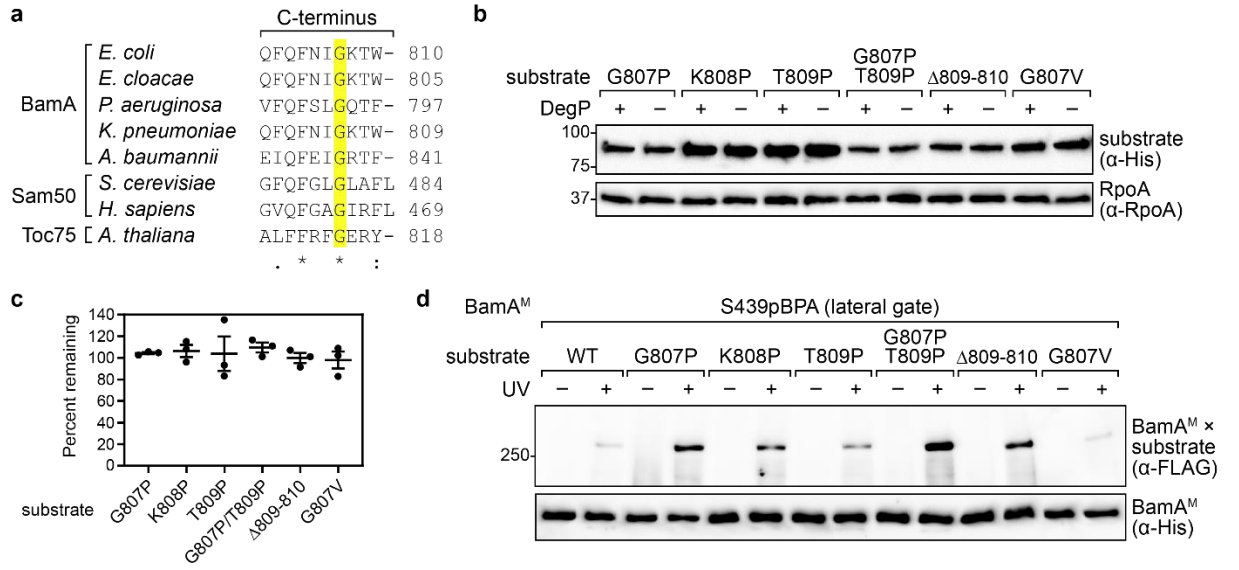

Figure 4.15. Mutations within the C-terminal overhang of the substrate lead to stalling at a late stage of assembly
a, Alignment of the C-terminal regions of BamA and other Omp85 superfamily homologs (Sam50 and Toc75) from various species. The invariant glycine is highlighted in yellow. **b**, Expression levels of 6×His-tagged substrates in strains with or without *degP* (MC4100 or MC4100 *degP::cam*, respectively). α-RpoA immunoblots are provided as loading controls. **c**, Quantification of the Western blot data shown in **b**, with expression levels for each substrate calculated as the percent that remains when *degP* is expressed. The plotted data represent mean ± SEM derived from quantification of immunoblotting data in **b** and additional independent replicates (n=3 for each substrate). **d**, *In vivo* photocrosslinking of 6×His-tagged full-length $BamA^M$ S439pBPA to 3×FLAG-tagged substrates with POTRA domains 3-5 deleted.

We believe that the β-barrel closure and release mechanism described here is general. Our structure shows that the N-terminal β-strand of $BamA^M$ contains a stretch of five residues (F428 to Y432) that interact with the C-terminal β-strand of the $BamA^S$-ΔL1 substrate. In contrast, the C-terminal β-strands of outer membrane proteins range in length from 7 to 25 residues (**Table 4.3**). Since the N-terminal β-strand of $BamA^M$ would be too short to accommodate the full length of the longer C-terminal β-strands of substrates, some portion of the latter could exist as an unpaired, overhanging sequence during folding that could mediate release in the mechanism we

describe for BamA^{S}. We predict that other Bam complex substrates have C-terminal overhangs that would protrude inside the hybrid barrel during folding to engage their own N-termini to trigger release.

Table 4.3. Summary of C-terminal β-strand lengths of bacterial outer membrane proteins

Number of strands	Protein name	Organism	PDB ID	Residues in C-terminal β-strand
8	Ail	*Yersinia pestis*	3QRA	14
	NspA	*Neisseria meningitidis*	1P4T	16
	OmpA	*Escherichia coli*	1QJP	9
	OmpA	*Klebsiella pneumoniae*	2K0L	7
	OmpW	*Escherichia coli*	2F1V	11
	OmpX	*Escherichia coli*	1QJ8	13
	OprG	*Pseudomonas aeruginosa*	2X27	12
	PagP	*Escherichia coli*	1MM4	16
	PagL	*Pseudomonas aeruginosa*	2ERV	11
	TtoA	*Thermus thermophilus*	3DZM	13
10	OmpT	*Escherichia coli*	1I78	25
	OpcA	*Neisseria meningitidis*	1K24	14
	Pla	*Yersinia pestis*	2X55	18
12	EspP	*Escherichia coli*	2QOM	14
	EstA	*Pseudomonas aeruginosa*	3KVN	13
	Hbp	*Escherichia coli*	3AEH	14
	Hia	*Haemophilus influenzae*	2GR8	10
	LpxR	*Salmonella typhimurium*	3FID	11
	NalP	*Neisseria meningitidis*	1UYN	14
	NanC	*Escherichia coli*	2WJQ	10
	OMPLA	*Escherichia coli*	1QD5	10
	Tsx	*Escherichia coli*	1TLW	11
14	FadL	*Escherichia coli*	1T16	19
	FadL	*Pseudomonas aeruginosa*	3DWO	19
	OmpG	*Escherichia coli*	2F1C	12
	TbuX	*Ralstonia pikettii*	3BRY	16
	TodX	*Pseudomonas putida*	3BS0	17
16	BamA	*Escherichia coli*	4N75	9
	BamA	*Haemophilus ducreyi*	4K3C	8
	FhaC	*Bordetella pertussis*	2QDZ	10
	Omp32	*Comamonas acidovorans*	1E54	13
	Omp32	*Delftia acidovorans*	2FGQ	13
	OmpK36	*Klebsiella pneumoniae*	1OSM	10
	OmpC	*Escherichia coli*	2J1N	10
	OmpF	*Escherichia coli*	2ZFG	10
	OprP	*Pseudomonas aeruginosa*	2O4V	13
	PhoE	*Escherichia coli*	1PHO	10
	Gdp	*Rhodobacter capsulatus*	2POR	10
	Gdp	*Rhodopseudomonas blastica*	1H6S	10

	PorB	*Neisseria meningitidis*	3A2R	14
18	BenF	*Pseudomonas fluorescens pf-5*	3JTY	13
	LamB	*Escherichia coli*	1MPM	14
	LamB	*Salmonella typhimurium*	2MPR	14
	OpdK	*Pseudomonas aeruginosa*	2QTK	14
	OprD	*Pseudomonas aeruginosa*	2ODJ	14
	ScrY	*Salmonella typhimurium*	1A0T	11
22	BtuB	*Escherichia coli*	1NQE	10
	Cir	*Escherichia coli*	2HDI	10
	FauA	*Bordetella pertussis*	3EFM	9
	FecA	*Escherichia coli*	1KMO	10
	FepA	*Escherichia coli*	1FEP	9
	FhuA	*Escherichia coli*	1BY3	9
	FptA	*Pseudomonas aeruginosa*	1XKW	10
	FpvA	*Pseudomonas aeruginosa*	2W75	9
	HasR	*Serratia marcescens*	3CSL	9
	ShuA	*Shigella dysenteriae*	3FHH	9
24	PapC	*Escherichia coli*	3FIP	12
26	LptD	*Escherichia coli*	4RHB	11
	LptD	*Klebsiella pneumoniae*	5IV8	11
	LptD	*Salmonella typhimurium*	4N4R	12
	LptD	*Shigella flexneri*	4Q35	14
36	SprA	*Flavobacterium johnsoniae*	6H3I	13

4.3. Discussion

The structure reported here supports the central idea of models in which substrates use the lateral opening of the BamA β-barrel as part of their assembly mechanism[133,135,156,158]. It is also consistent with crosslinking studies showing that the C-terminal β-strand of the folding substrate interacts strongly with the N-terminal edge of the $BamA^{M}$ β-barrel[143,144,174]. Moreover, its two unexpected features—the unpaired edges at one junction and the hydrogen-bond register and overhang at the other—resolve problems posed by models postulating hydrogen-bonded arrays at both edges of the machine and substrate. Sequential replacement of each substrate-machine hydrogen bond by a substrate-substrate hydrogen bond yields a succession of rapid steps rather than a single, very slow one. An understanding of how the Bam complex interacts with substrates may enable the design of antibiotics that bind to features of BamA to inhibit substrate assembly.

4.4. Materials and methods

4.4.1. SDS-PAGE and immunoblotting

Homemade Tris-HCl 4-20% polyacrylamide gradient gels or Mini-PROTEAN TGX 7.5% precast gels (BioRad) were used with Tris-glycine running buffer. 2× SDS sample buffer refers to a mixture containing 125 mM Tris (pH 6.8), 4% (wt/vol) SDS, 30% (vol/vol) glycerol, 0.005% bromophenol blue, and 5% (vol/vol) β-mercaptoethanol. SDS-PAGE analyses were performed at 200 V for 45 to 60 min. To analyze purified protein complexes for cryo-EM, SDS-PAGE was performed and followed by staining with Coomassie Brilliant Blue (Alfa Aesar). Coomassie-stained SDS-PAGE gels were imaged using the “Gel” feature of an Azure Biosystems C400 imager. For Western blotting, proteins were transferred onto Immun-Blot PVDF membranes (Bio-Rad) and then incubated with appropriate antibodies. All HRP conjugates were visualized with the Amersham ECL Prime Western blotting detection reagent (GE Healthcare). Western blots were imaged using the “lowest” sensitivity setting of the “Chemi” feature of an Azure Biosystems C400 imager.

4.4.2. Electron microscopy data collection

Purified substrate-bound Bam complex as described above was applied to glow-discharged Quantifoil R 2/1 holey carbon 400-mesh copper grids (Quantifoil). Grids were blotted for 4-5 s at 100% humidity with the blot force set to 16, and flash frozen by liquid nitrogen-cooled liquid ethane using a Thermo Fisher Scientific Vitrobot Mark IV (Thermo Fisher Scientific). The grid was then loaded onto a Titan Krios G3i electron cryo-microscope (Thermo Fisher) operated at 300 kV accelerating voltage. Image stacks (movies) were recorded on a Gatan Bioquantum K3 Imaging Filter (Gatan, USA) using the super-resolution counting mode and the calibrated magnification of 58717× using SerialEM[188]. The slit of the energy filter was set to 25 eV, with a defocus range between 1.1 and 2.8 µm. The electron dose rate was 17 e−/physical pixel per

second, and the subframe time was set to 0.06 second. A total exposure time of 3 second resulted in 50 subframes per image stack. The total electron dose was 70 e−/Å^2 (~1.4 e−/Å^2 per subframe). The multi-shot scheme in SerialEM was used for data collection, with settings of 4 holes per stage move, and 5 shots per hole, which greatly sped up the data collection.

4.4.3. Image processing and 3D reconstruction

The movie frames were motion-corrected and dose-weighted by MotionCor2[189] and CTF parameters were estimated by CTFFIND4[190]. Particle picking was carried out using crYOLO[191] giving 2,054,956 initial particles. Following successive rounds of 2D and 3D classification within Relion[192], 516,419 particles were then "polished" through the Bayesian polishing process[193]. These particles were then subjected to further 2D and 3D classification within CryoSPARC[194], after which 223,353 particles were selected (see classification flowchart in Extended Data Figure 3) which led to the final reconstruction at 4.1 Å resolution. Maps used for figures were filtered according to local resolution with b-factor sharpening within CryoSPARC. Structural biology applications used in this project (other than CryoSPARC) were compiled and configured by SBGrid[195].

4.4.4. Model building, refinement, and validation

The atomic model was generated using available structures of the Bam complex components. For BamA^{M}BDE, an initial model was generated by rigid-body fitting these components into the EM map using Chimera[196]. The BamA^M POTRA domains and BamE were obtained from PDB ID: 5D0O. The BamA^M β-barrel and BamD were obtained from PDB ID: 5D0Q. BamB was obtained from PDB ID: 2YH3. For the substrate, a homology model generated with I-TASSER was used[197–199].

The initial modeling was followed by manual adjustments using Coot[200]. All selenomethiones were replaced with methionines. The improved model was then refined in real

space against the cryo-EM map using real space refinement in PHENIX[201] with secondary structure restraints. Iterative rounds of manual and automated refinement in Coot and PHENIX, respectively, generated a model that included all components except for BamC and the POTRA domains 1 and 2 of $BamA^M$. Due to the lower resolution of these components, the atomic models were generated by docking previously determined structures into our map (from 5D0Q for BamC and 5D0O for the POTRA domains) without further refinement. For the entirety of POTRA domains 1 and 2 of $BamA^M$ (residues 1-171) and the C-terminal domain of BamC (residues 89-209), only the main chain atoms are included in the model. For the substrate, the two POTRA domains could not be resolved, so they are not included in the model. The final model was visually inspected for general fit to the map, and geometry was further evaluated using MolProbity[202]. Cryo-EM data collection, refinement, and validation statistics are summarized in Extended Data Table 1. Figures depicting the structure were prepared in Chimera or PyMOL (Schrödinger, https://www.pymol.org).

4.4.5. Bacterial strains

Bacterial strains used in the experiments reported here are provided in Table 4.4.

Table 4.4. Bacterial strains used in chapter 4

Strains	Genotype	Source
MC4100	F^- *araD139* Δ*(argF-lac)U169 rpsL150 relA1 flbB5301 deoC1 ptsF25 rbsR thi*	[173]
NR779	MC4100 *degP::cam*	[127]
NovaBlue	*endA1 hsd*R17 (r_K^-, m_K^+) *sup*E44 *thi*-1 *rec*A1 *gyr*A96 *rel*A1 *lac* F′[*pro*A^+ B^+ *lac*I^q ZΔM15::Tn*10*]	Novagen
BL21(DE3)	*fhuA2 [lon] ompT gal (λ DE3) [dcm]* Δ*hsdS*	Novagen

4.4.6. Bacterial plasmids

Bacterial plasmids used in the experiments here are provided in Table 4.5.

Table 4.5. Bacterial plasmids used in chapter 4

Plasmid	Description	Source
pDT209	pTrc99a/3×FLAG-BamA(Δ172-421)	This study
pDT340	pTrc99a/3×FLAG-BamA–BamB–BamC–BamD-8×His–BamE	Modified from [113]
pDT397	pTrc99a/BamA(S439C/C690S/C700S)–BamB–BamC–BamD–BamE-8×His)	Modified from [113]
pDT400	pZS21/6×His-BamA(S439C)	Lee et al.

pDT486	pTrc99a/3×FLAG-BamA(Δ172-421/Δ430-439/Y787C)	This study
pDT487	pTrc99a/3×FLAG-BamA(Δ172-421/Δ430-439/E800C)	This study
pDT481	pTrc99a/3×FLAG-BamA(Δ172-421/Δ430-439/F802C)	This study
pDT488	pTrc99a/3×FLAG-BamA(Δ172-421/Δ430-439/Q803C)	This study
pDT489	pTrc99a/3×FLAG-BamA(Δ172-421/Δ430-439/F804C)	This study
pDT490	pTrc99a/3×FLAG-BamA(Δ172-421/Δ430-439/N805C)	This study
pDT500	pBAD33/2×Strep-BamA(Δ172-421/Δ430-439/C690S/C700S/E800C)	This study
pDT504	pSup-BpaRS-6TRN(spec^{R})	Modified from [183]
pDT509	pTrc99a/2×Strep-BamA(Δ172-421/C690S/C700S/E800C)	This study
pDT510	pTrc99a/2×Strep-BamA(Δ172-421/Δ430-439/C690S/C700S/E800C)	This study
pDT520	pZS21/6×His-BamA(G807V)	This study
pDT521	pZS21/6×His-BamA(Δ809-810)	This study
pDT525	pTrc99a/3×FLAG-BamA(Δ172-421/G807V)	This study
pDT526	pTrc99a/3×FLAG-BamA(Δ172-421/Δ809-810)	This study
pDT534	pTrc99a/2×Strep-BamA(Δ172-421/C690S/C700S/T467C)	This study
pDT535	pTrc99a/2×Strep-BamA(Δ172-421/Δ430-439/C690S/C700S/T467C)	This study
pDT536	pZS21/6×His-BamA(G807P)	This study
pDT537	pZS21/6×His-BamA(K808P)	This study
pDT538	pZS21/6×His-BamA(T809P)	This study
pDT550	pZS21/6×His-BamA(G807P/T809P)	This study
pDT558	pTrc99a/3×FLAG-BamA(Δ172-421/G807P)	This study
pDT559	pTrc99a/3×FLAG-BamA(Δ172-421/K808P)	This study
pDT560	pTrc99a/3×FLAG-BamA(Δ172-421/T809P)	This study
pDT562	pTrc99a/3×FLAG-BamA(Δ172-421/G807P/T809P)	This study
pDT566	pTrc99a/3×FLAG-BamA(Δ430-439/G460C)	This study
pDT567	pTrc99a/3×FLAG-BamA(Δ430-439/T467C)	This study
pDT568	pTrc99a/3×FLAG-BamA(Δ430-439/A469C)	This study
pDT569	pCDF/2×Strep-BamA(Δ172-421/Δ430-439/F428pBPA)	This study
pDT571	pCDF/2×Strep-BamA(Δ172-421/Δ430-439/F440pBPA)	This study
pJL077	pZS21/6×His-BamA(S439pBPA)	This study
pSK476	pZS21/6×His-BamA	[73]

Chapter 5: Conclusion

5.1. Summary of work presented

A substantial amount of progress has been made over the past twenty years to identify the components of the Bam complex, elucidate its architecture, and understand its role in β-barrel assembly. Recently, inhibitors of this molecular machine have also been disclosed[203–206]. At the start of the work that constitutes this dissertation, it was unclear how the complex interacts with substrates in a manner that would promote their folding in the absence of an exogenous energy source.

To gain an understanding of the mechanism of substrate assembly, we used LptD as a substrate of the Bam complex to show that substrates pass into the interior of BamA, and that folding is catalyzed within that confined environment (Chapter 2). We then generated a panel of stalled substrate BamA variants to show directionality of folding and the location of substrates on the machine at different stages of folding (Chapter 3). Finally, we obtained a structure of a substrate-engaged Bam complex, which showed how BamA captures the substrate and how substrates can overcome the barrier to release from the machine after folding has finished (Chapter 4). With the experiments presented here, a more detailed picture is now available for how the Bam complex mediates the assembly of two substrates, BamA and LptD, by directly interacting with them in a way that likely templates folding.

5.2. Model of β-barrel assembly in Gram-negative bacteria

Our biochemical and structural data lead to a molecular model for β-barrel assembly. Initially, an unfolded substrate is recruited to the Bam complex (**Figure 5.1a**), likely by interactions with the lipoprotein BamD[139]. This recruitment triggers opening of the lateral gate of BamA, exposing the N-terminal β-strand of BamA[129]. This edge serves as a template for substrate folding by forming hydrogen bonds with the C-terminal β-strand of the substrate through β-sheet augmentation[137–139] (**Figure 5.1b**). The remainder of the substrate may then pass into the lumen of BamA. In this environment, which acts as a confined folding cage to promote substrate folding,

β-strands are added sequentially towards the N-terminus of the substrate (**Figure 5.1c**). The exposed edge of each added β-strand acts as a template for addition of the next by hydrogen bonding. As the growing substrate β-barrel is formed, it may exit out through the lateral gate of BamA (due to the limited space within the interior of BamA), possibly into the periplasmic cavity of the Bam complex. Once a substantial amount of the substrate has been folded, membrane integration occurs (**Figure 5.1d**). Upon addition of all the β-strands of the substrate to the nascent barrel, its N- and C-termini can interact, promoting release from the Bam complex (**Figure 5.1e and Figure 4.14**). This process would occur through the stepwise disruption of intermolecular machine-substrate hydrogen bonds and formation of intramolecular hydrogen bonds between the N- and C-terminal β-strands of the substrate. This mechanism allows the Bam complex to catalyze β-barrel assembly in the absence of an external energy source. Although we do not directly observe disruption of the membrane by the Bam complex as proposed in the BamA-assisted model for folding, it is likely that this occurs near the lateral gate of BamA to promote efficient membrane integration.

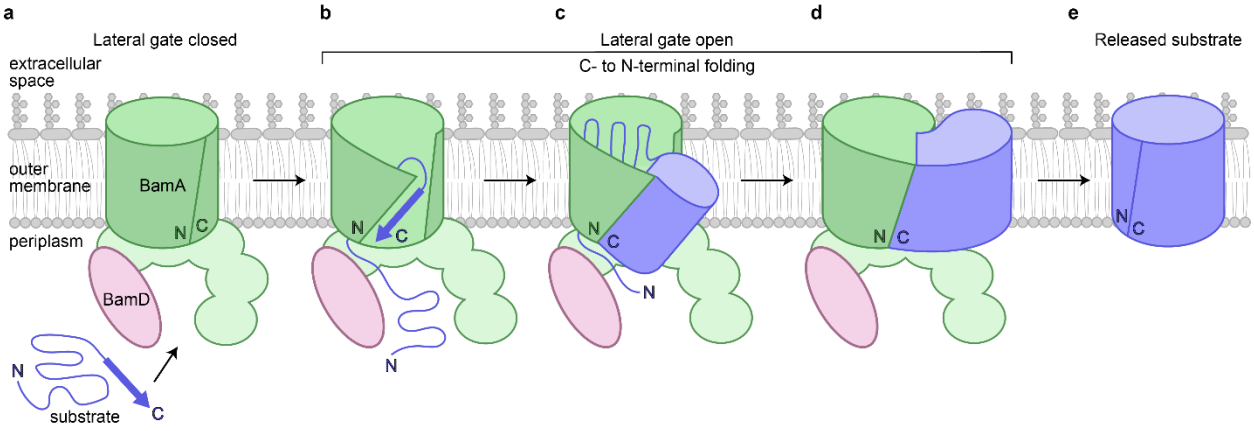

Figure 5.1. Proposed model of β-barrel assembly by the Bam complex
a, The substrate is recruited to the Bam complex. BamA^{M} is in a closed state prior to substrate-induced opening of its lateral gate. For simplicity, only BamA^{M}, the substrate, and BamD are shown. **b-d**, The C-terminus of the substrate interacts with the exposed N-terminal edge of BamA^{M}, and β-strands or β-hairpins of the substrate are added sequentially from the C- to N-terminus. Early folding may occur within the interior of the BamA^{M} β-barrel, and folded portions of the substrate may then be released outward. Full membrane integration could occur after a substantial amount of folding. Steps **b** and **c** represent intermediate stages to **d**, which corresponds to the cryo-EM structure of the substrate-bound Bam complex. **e**, The substrate is released into the membrane environment once its N- and C-terminal ends are joined.

Our results demonstrate how BamA accelerates β-barrel assembly. To initiate folding, proper orientation of a portion of the substrate, likely a β-strand or β-hairpin, is required. Once this nucleation event has occurred, rapid addition of the remaining β-strands would be possible. BamA appears to overcome this barrier by using its own β-sheet as a template that interacts with and orients the C-terminus of the substrate to initiate folding. Although there would be an entropic penalty to form the initial (C-terminal) β-strand of the substrate, further addition of β-strands would be accelerated using the preorganized template. With the C-terminal β-strand of the substrate held by stably by BamA, folding would proceed through the stepwise addition of β-strands in a C- to N-terminal direction, with the edge of each added β-strand acting as a template for addition of the next by hydrogen bonding. Thus, a barrier to folding is overcome by a mechanism in which a substrate assembles its own β-sheet by taking advantage of the existing β-sheet of BamA.

We provide a molecular explanation for how BamA mediates the directional folding of a substrate. The C-terminus of the substrate associates with BamA β1 via hydrogen bonding, a mode of interaction that could remain fixed throughout the folding process. Importantly, this would establish the proper topology of the substrate, with hydrophilic residues of the substrate oriented inward and hydrophobic residues oriented outward. We do not observe hydrogen bonding between the N-terminus of the substrate and the C-terminus of BamA. This makes sense in terms of the direction of folding, as the N-terminal end of the substrate would remain free for “growth” to occur at that end. Interactions between both ends of BamA with both ends of the substrate would not be feasible, as this would require breaking of the hydrogen bonding network at one end each time a new β-strand is added to the growing substrate β-barrel. Importantly, conserved sequence motifs, referred to as β-signals, are often found in the C-terminal β-strand of substrates[137–139]. We show that BamA recognizes substrates via the C-terminal β-strand, explaining how this machine can process substrates that are otherwise quite diverse. Through these means, BamA promotes the C- to N-terminal directional folding of a substrate by holding only the C-terminus, allowing the substrate to extend at its free N-terminus.

Our results provide an explanation for earlier studies reporting that opening of the lateral gate of BamA is required for survival of *E. coli* and full function in an *in vitro* reconstitution of folding[111,135]. In those experiments, BamA was fixed in a closed state by introducing a disulfide bond between the N- and C-terminal β-strands of the β-barrel of BamA. A disulfide bond would prevent the opening of the BamA and interaction with substrates by β-sheet augmentation. In contrast, it has also been reported that BamA can assemble OmpX, an 8-stranded β-barrel substrate, even when the lateral gate is fixed in a closed state[114]. This was rationalized via a mechanism in which dynamics of the N- and C-terminal edges of BamA, which exist even when the lateral gate is fixed in a closed state, would disrupt the lipid bilayer. OmpX could then bypass polar phospholipid head groups and insert into the membrane via its intrinsic folding ability[114,133,135,153–155]. Therefore, while smaller substrates such as OmpX may only require the local disruption of lipids by BamA in order to fold, based on our structure, larger substrates also appear to require that the Bam complex play a more active role in folding by likely templating β-sheet formation.

The location of substrate folding (i.e., within the periplasm or the membrane) has been debated. Here, we show that, for the assembly of BamA, early stages of folding occur outside the membrane environment within the periplasm and the interior of the BamA β-barrel, as early-stalled substrates are extractable by urea, susceptible to periplasmic proteases, and able to form crosslinks to pBPA installed in the interior wall of BamA. For substrate LptD, we show a similar localization within BamA and that the interior wall of BamA catalyzes folding. We presume that substrates at early stages interact, via their C-termini, with the N-terminus of BamA, and that this interaction remains constant throughout folding. At a later stage, the substrate undergoes membrane integration, directly interacting with membrane lipids while still bound to the N-terminal β-strand of BamA. This is likely possible given the extreme flexibility of BamA β1 that we observe. Therefore, a role of the Bam complex is to provide an environment, its interior, where a substrate can sample both hydrophilic and hydrophobic environments. Membrane integration could occur

spontaneously once the substrate has formed a sufficiently hydrophobic exterior surface that would favor localization within the membrane rather than within an aqueous environment. This would mirror the membrane integration process of substrates by the Sec translocon[207].

An issue with the previously proposed BamA-budding model is that, although it explains how the Bam complex could reduce the barrier to initiate folding, the formation of a BamA-substrate "super-barrel" structure raises the question of how substrates would be released once folding has completed. Although it is expected that the folded substrate would be more thermodynamically stable in the membrane where it could form a closed structure and its exterior surface could engage in additional interactions with membrane lipids, there would be a high kinetic barrier to release from the Bam complex due to the hydrogen bonding network between BamA and the substrate. We show, in contrast to the BamA-budding model, that only the C-terminal end of the substrate is held stably by the machine, reducing the number of interactions between BamA and the substrate that would need to be broken. Additionally, we show that features at the N- and C-terminal ends of the substrate are in proximity at the late stage of folding we observe, facilitating their joining through a stepwise exchange of hydrogen bonds from intermolecular machine-substrate bonds to intramolecular substrate-substrate bonds. This process would be feasible in the absence of an energy source.

5.3. β-barrel assembly in eukaryotes

Mitochondria and chloroplasts also contain β-barrel outer membrane proteins. Homologs of BamA exist in the outer membranes of mitochondria (Sam50, for sorting and assembly machinery)[208–210] and others exist in the outer membranes of chloroplasts (Toc75 and OEP80)[211,212]. Notably, there are not known homologs of the accessory lipoproteins of the Bam complex (BamBCDE) in mitochondria and chloroplasts.

It is believed that the Bam and Sam complexes share a similar mechanism of substrate folding, since the Bam complex can fold mitochondrial β-barrels and the Sam complex can fold

bacterial β-barrels[213–215]. Importantly, our findings pertaining to the topology of the stalled substrate on the Bam complex are in agreement with experiments performed with the mitochondrial protein Sam50[144]. Thus, similar principles of β-barrel assembly operate in Gram-negative bacteria and eukaryotes. The function of one protein of the Sam complex, Sam35, has been proposed to be recognition of substrates by binding C-terminal peptides, and a similar function for BamD has been proposed; however, there is no structural similarity between BamD and Sam35.

Less is understood about the homologs of BamA in chloroplasts, Toc75 and OEP80 (outer envelope protein). Toc75 has been identified as the import machinery for proteins into chloroplasts, while the OEP80 protein is believed to perform β-barrel assembly[211,212]. However, further evidence demonstrating the activity of OPE80, and whether its mechanism of function mirrors that of the Bam complex, is required.

5.4. Questions for future work

There are still unanswered questions regarding the mechanism of β-barrel assembly. Although BamD has a proposed role in substrate recruitment, the roles of the other lipoprotein components of the complex are unclear. It has been shown that multiple Bam complexes can co-localize in outer membrane “precincts” mediated by interactions by BamB[216]. However, the significance of this in the folding mechanism of the Bam complex remains to be examined in more detail. BamC has been shown to be localized to the cell surface, but it is unclear what the significance of this is, and this has yet to be observed by structural techniques. Also, how substrates are passed from chaperones to the Bam complex remains unclear. This passing of substrates is likely to the POTRA domains of BamA, and the role of these domains is also unclear. It has been proposed that they can interact with β-strands of substrates, but convincing evidence for this is not available. Also, it has been assumed that substrates traverse the periplasm through chaperones that deliver them to the outer membrane. Recently crosslinking experiments that

proposed that a super-complex consisting of the Sec translocon, the chaperone SurA, and the Bam complex exists to connect the inner and outer membrane for protein delivery[217]. Additional work will need to be done to confirm the relevance *in vivo* of these protein bridges.

It is currently unclear whether the principles described here for the assembly of LptD and BamA are applicable to that of other substrates. It is possible that the assembly of some substrates may mirror what we observe for LptD and BamA, while other substrates may simply require membrane disruption and less of a direct role of the lateral opening of BamA. We anticipate that additional structures of the Bam complex bound to other substrates, perhaps within a membrane environment, will allow a greater understanding of the assembly of β-barrel outer membrane proteins. In particular, studies of BamA earlier in its folding and of substrates either larger or smaller than BamA will reveal whether the principles of folding that we describe here are applicable.

5.5. Targeting the Bam complex as a strategy for antibiotic development

Antibiotic resistant infections, especially ones caused by Gram-negative bacteria, are an emerging problem. While the main goal of this research is to provide a basic understanding of how β-barrel membrane proteins are folded, the structural and functional information presented here may enable the design of new antibiotics since proper assembly of these proteins is critical for survival of Gram-negative bacteria. Indeed, studies have identified the Bam complex as a potential target for novel antimicrobial therapies[218]. Such therapies could interfere with proper function of the Bam complex by preventing interactions with substrates or by stalling substrates during their assembly. Importantly, since this machine is localized within the outer membrane, with extracellular loops exposed to the cell surface, it could be more straightforward to target compared to intracellular machines. As Bam complex is involved in production of virulence factors, targeting it could prevent a pathogen from penetrating a host's immune system. Recently, an antibody targeting BamA has been shown to perturb outer membrane integrity when added to

E. coli cells and can prevent proper folding of outer membrane proteins by binding extracellular loops of BamA[203]. More recently, small molecules that interact with the Bam complex have been identified that likely lead to cell death by preventing efficient function of this molecular machine[204–206].

References

1. Nikaido, H. & Vaara, M. Molecular basis of bacterial outer membrane permeability. *Microbiological Reviews* (1985). doi:10.1128/mmbr.49.1.1-32.1985
2. Nikaido, H. Molecular Basis of Bacterial Outer Membrane Permeability Revisited. *Microbiol. Mol. Biol. Rev.* **67**, 593–656 (2003).
3. Silhavy, T., Kahne, D. & Walker, S. The bacterial cell envelope. *Cold Spring Harbor perspectives in biology* **2**, 1–16 (2010).
4. Raetz, C. R. H. & Whitfield, C. Lipopolysaccharide endotoxins. *Annu. Rev. Biochem.* **71**, 635–700 (2002).
5. Raetz, C. R. H., Reynolds, C. M., Trent, M. S. & Bishop, R. E. Lipid A Modification Systems in Gram-Negative Bacteria. *Annu. Rev. Biochem.* **76**, 295–329 (2007).
6. Whitfield, C. & Trent, M. S. Biosynthesis and Export of Bacterial Lipopolysaccharides. *Annu. Rev. Biochem.* (2014). doi:10.1146/annurev-biochem-060713-035600
7. Karow, M. & Georgopoulos, C. The essential Escherichia coli msbA gene, a multicopy suppressor of null mutations in the htrB gene, is related to the universally conserved family of ATP-dependent translocators. *Mol. Microbiol.* (1993). doi:10.1111/j.1365-2958.1993.tb01098.x
8. Polissi, A. & Georgopoulos, C. Mutational analysis and properties of the msbA gene of Escherichia coli, coding for an essential ABC family transporter. *Mol. Microbiol.* (1996). doi:10.1111/j.1365-2958.1996.tb02642.x
9. Zhou, Z., White, K. A., Polissi, A., Georgopoulos, C. & Raetz, C. R. H. Function of Escherichia coli MsbA, an essential ABC family transporter, in lipid A and phospholipid biosynthesis. *J. Biol. Chem.* (1998). doi:10.1074/jbc.273.20.12466
10. Doerrler, W. T. & Raetz, C. R. H. ATPase activity of the MsbA lipid flippase of Escherichia coli. *J. Biol. Chem.* (2002). doi:10.1074/jbc.M205857200
11. Whitfield, C. Biosynthesis and Assembly of Capsular Polysaccharides in Escherichia coli . *Annu. Rev. Biochem.* (2006). doi:10.1146/annurev.biochem.75.103004.142545
12. Rietschel, E. T. *et al.* Bacterial endotoxin: molecular relationships of structure to activity and function. *FASEB J.* (1994). doi:10.1096/fasebj.8.2.8119492
13. Sperandeo, P. *et al.* Characterization of lptA and lptB, two essential genes implicated in lipopolysaccharide transport to the outer membrane of Escherichia coli. *J. Bacteriol.* **189**, 244–53 (2007).
14. Sperandeo, P. *et al.* Functional analysis of the protein machinery required for transport of lipopolysaccharide to the outer membrane of Escherichia coli. *J. Bacteriol.* **190**, 4460–9 (2008).
15. Ruiz, N., Gronenberg, L. S., Kahne, D. & Silhavy, T. J. Identification of two inner-membrane proteins required for the transport of lipopolysaccharide to the outer membrane of Escherichia coli. *Proc. Natl. Acad. Sci. U. S. A.* **105**, 5537–42 (2008).
16. Chng, S. S., Gronenberg, L. S. & Kahne, D. Proteins required for lipopolysaccharide assembly in escherichia coli form a transenvelope complex. *Biochemistry* (2010). doi:10.1021/bi100493e
17. Freinkman, E., Okuda, S., Ruiz, N. & Kahne, D. Regulated assembly of the transenvelope protein complex required for lipopolysaccharide export. *Biochemistry* **51**, 4800–4806 (2012).
18. Villa, R. *et al.* The Escherichia coli Lpt transenvelope protein complex for lipopolysaccharide export is assembled via conserved structurally homologous domains. *J. Bacteriol.* **195**, 1100–8 (2013).
19. Sherman, D. J. *et al.* Lipopolysaccharide is transported to the cell surface by a membrane-to-membrane protein bridge. *Science (80-.).* **359**, 798–801 (2018).
20. Okuda, S., Freinkman, E. & Kahne, D. Cytoplasmic ATP hydrolysis powers transport of lipopolysaccharide across the periplasm in E. coli. *Science* **338**, 1214–7 (2012).
21. Sherman, D. J. *et al.* Decoupling catalytic activity from biological function of the ATPase

that powers lipopolysaccharide transport. *Proc. Natl. Acad. Sci. U. S. A.* **111**, 4982–7 (2014).
22. Bos, M. P., Tefsen, B., Geurtsen, J. & Tommassen, J. Identification of an outer membrane protein required for the transport of lipopolysaccharide to the bacterial cell surface. *Proc. Natl. Acad. Sci. U. S. A.* **101**, 9417–22 (2004).
23. Wu, T. *et al.* Identification of a protein complex that assembles lipopolysaccharide in the outer membrane of Escherichia coli. *Proc Natl Acad Sci U S A* **103**, 11754–11759 (2006).
24. Tran, A. X., Dong, C. & Whitfield, C. Structure and functional analysis of LptC, a conserved membrane protein involved in the lipopolysaccharide export pathway in Escherichia coli. *J. Biol. Chem.* (2010). doi:10.1074/jbc.M110.144709
25. Suits, M. D. L., Sperandeo, P., Dehò, G., Polissi, A. & Jia, Z. Novel Structure of the Conserved Gram-Negative Lipopolysaccharide Transport Protein A and Mutagenesis Analysis. *J. Mol. Biol.* (2008). doi:10.1016/j.jmb.2008.04.045
26. Qiao, S., Luo, Q., Zhao, Y., Zhang, X. C. & Huang, Y. Structural basis for lipopolysaccharide insertion in the bacterial outer membrane. *Nature* **511**, 108–11 (2014).
27. Dong, H. *et al.* Structural basis for outer membrane lipopolysaccharide insertion. *Nature* **511**, 52–6 (2014).
28. Botos, I. *et al.* Structural and Functional Characterization of the LPS Transporter LptDE from Gram-Negative Pathogens. *Structure* (2016). doi:10.1016/j.str.2016.03.026
29. Freinkman, E., Chng, S.-S. & Kahne, D. The complex that inserts lipopolysaccharide into the bacterial outer membrane forms a two-protein plug-and-barrel. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 2486–91 (2011).
30. Malojčić, G. *et al.* LptE binds to and alters the physical state of LPS to catalyze its assembly at the cell surface. *Proc. Natl. Acad. Sci. U. S. A.* **111**, 9467–72 (2014).
31. Lee, J. *et al.* Characterization of a stalled complex on the β-barrel assembly machine. *Proc. Natl. Acad. Sci.* **113**, 8717–8722 (2016).
32. Ruiz, N., Chng, S.-S., Hiniker, A., Kahne, D. & Silhavy, T. J. Nonconsecutive disulfide bond formation in an essential integral outer membrane protein. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 12245–50 (2010).
33. Chng, S.-S. *et al.* Disulfide rearrangement triggered by translocon assembly controls lipopolysaccharide export. *Science* **337**, 1665–8 (2012).
34. Jones, N. C. & Osborn, M. J. Translocation of phospholipids between the outer and inner membranes of Salmonella typhimurium. *J. Biol. Chem.* (1977).
35. Langley, K. E., Hawrot, E. & Kennedy, E. P. Membrane assembly: Movement of phosphatidylserine between the cytoplasmic and outer membranes of Escherichia coli. *J. Bacteriol.* (1982).
36. Raetz, C. R. H. Enzymology, genetics, and regulation of membrane phospholipid synthesis in Escherichia coli. *Microbiol. Rev.* (1978). doi:10.1128/mmbr.42.3.614-659.1978
37. Sohlenkamp, C. & Geiger, O. Bacterial membrane lipids: Diversity in structures and pathways. *FEMS Microbiology Reviews* (2015). doi:10.1093/femsre/fuv008
38. Raetz, C. R. H. & Dowhan, W. Biosynthesis and function of phospholipids in Escherichia coli. *Journal of Biological Chemistry* (1990).
39. Gibellini, F. & Smith, T. K. The Kennedy pathway-de novo synthesis of phosphatidylethanolamine and phosphatidylcholine. *IUBMB Life* (2010). doi:10.1002/iub.337
40. Shrivastava, R. & Chng, S. S. Lipid trafficking across the Gram-negative cell envelope. *J. Biol. Chem.* (2019). doi:10.1074/jbc.AW119.008139
41. Donohue-Rolfe, A. M. & Schaechter, M. Translocation of phospholipids from the inner to the outer membrane of Escherichia coli. *Proc. Natl. Acad. Sci. U. S. A.* (1980). doi:10.1073/pnas.77.4.1867

42. Bayer, M. E. Areas of adhesion between wall and membrane of Escherichia coli. *J. Gen. Microbiol.* (1968). doi:10.1099/00221287-53-3-395
43. Kellenberger, E. The 'Bayer bridges' confronted with results from improved electron microscopy methods. *Mol. Microbiol.* (1990). doi:10.1111/j.1365-2958.1990.tb00640.x
44. Dalebroux, Z. D. *et al.* Delivery of cardiolipins to the salmonella outer membrane is necessary for survival within host tissues and virulence. *Cell Host Microbe* (2015). doi:10.1016/j.chom.2015.03.003
45. Rossi, R. M., Yum, L., Agaisse, H. & Payne, S. M. Cardiolipin synthesis and outer membrane localization are required for Shigella flexneri virulence. *MBio* (2017). doi:10.1128/mBio.01199-17
46. Malinverni, J. C. & Silhavy, T. J. An ABC transport system that maintains lipid asymmetry in the Gram-negative outer membrane. *Proc. Natl. Acad. Sci. U. S. A.* (2009). doi:10.1073/pnas.0903229106
47. Chong, Z. S., Woo, W. F. & Chng, S. S. Osmoporin OmpC forms a complex with MlaA to maintain outer membrane lipid asymmetry in Escherichia coli. *Mol. Microbiol.* (2015). doi:10.1111/mmi.13202
48. Yeow, J. *et al.* The architecture of the OmpC–MlaA complex sheds light on the maintenance of outer membrane lipid asymmetry in Escherichia coli. *J. Biol. Chem.* (2018). doi:10.1074/jbc.RA118.002441
49. Lloubès, R. *et al.* The Tol-Pal proteins of the Escherichia coli cell envelope: An energized system required for outer membrane integrity? *Research in Microbiology* (2001). doi:10.1016/S0923-2508(01)01226-8
50. Sturgis, J. N. Organisation and evolution of the tol-pal gene cluster. *J. Mol. Microbiol. Biotechnol.* (2001).
51. Shrivastava, R., Jiang, X. & Chng, S. S. Outer membrane lipid homeostasis via retrograde phospholipid transport in Escherichia coli. *Mol. Microbiol.* (2017). doi:10.1111/mmi.13772
52. Tokuda, H. & Matsuyama, S. I. Sorting of lipoproteins to the outer membrane in E. coli. *Biochimica et Biophysica Acta - Molecular Cell Research* (2004). doi:10.1016/j.bbamcr.2004.02.005
53. Okuda, S. & Tokuda, H. Lipoprotein sorting in bacteria. *Annu. Rev. Microbiol.* **65**, 239–59 (2011).
54. Kudva, R. *et al.* Protein translocation across the inner membrane of Gram-negative bacteria: The Sec and Tat dependent protein transport pathways. *Res. Microbiol.* (2013). doi:10.1016/j.resmic.2013.03.016
55. Hussain, M., Ichihara, S. & Mizushima, S. Mechanism of signal peptide cleavage in the biosynthesis of the major lipoprotein of the Escherichia coli outer membrane. *J. Biol. Chem.* (1982).
56. Tokunaga, M., Tokunaga, H. & Wu, H. C. Post-translational modification and processing of Escherichia coli prolipoprotein in vitro. *Proc. Natl. Acad. Sci. U. S. A.* (1982). doi:10.1073/pnas.79.7.2255
57. Sankaran, K. & Wu, H. C. Lipid modification of bacterial prolipoprotein. Transfer of diacylglyceryl moiety from phosphatidylglycerol. *J. Biol. Chem.* (1994).
58. Yamaguchi, K., Yu, F. & Inouye, M. A single amino acid determinant of the membrane localization of lipoproteins in E. coli. *Cell* (1988). doi:10.1016/0092-8674(88)90162-6
59. Yakushi, T., Yokota, N., Matsuyama, S. I. & Tokuda, H. LolA-dependent release of a lipid-modified protein from the inner membrane of Escherichia coli requires nucleoside triphosphate. *J. Biol. Chem.* (1998). doi:10.1074/jbc.273.49.32576
60. Yakushi, T., Masuda, K., Narita, S. I., Matsuyama, S. I. & Tokuda, H. A new ABC transporter mediating the detachment of lipid-modified proteins from membranes. *Nat. Cell Biol.* **2**, 212–218 (2000).

61. Ito, Y., Kanamaru, K., Taniguchi, N., Miyamoto, S. & Tokuda, H. A novel ligand bound ABC transporter, LolCDE, provides insights into the molecular mechanisms underlying membrane detachment of bacterial lipoproteins. *Mol. Microbiol.* (2006). doi:10.1111/j.1365-2958.2006.05378.x
62. Matsuyama, S., Tajima, T. & Tokuda, H. A novel periplasmic carrier protein involved in the sorting and transport of Escherichia coli lipoproteins destined for the outer membrane. *Embo J.* **14**, 3365–3372 (1995).
63. Yakushi, T., Masuda, K., Narita, S. I., Matsuyama, S. I. & Tokuda, H. A new ABC transporter mediating the detachment of lipid-modified proteins from membranes. *Nat. Cell Biol.* (2000). doi:10.1038/35008635
64. Matsuyama, S. i, Yokota, N. & Tokuda, H. A novel outer membrane lipoprotein, LolB (HemM), involved in the LolA (p20)-dependent localization of lipoproteins to the outer membrane of Escherichia coli. *EMBO J.* **16**, 6947–55 (1997).
65. Okuda, S. & Tokuda, H. Model of mouth-to-mouth transfer of bacterial lipoproteins through inner membrane LolC, periplasmic LolA, and outer membrane LolB. *Proc. Natl. Acad. Sci. U. S. A.* (2009). doi:10.1073/pnas.0900896106
66. Takeda, K. *et al.* Crystal structures of bacterial lipoprotein localization factors, LolA and LolB. *EMBO J.* (2003). doi:10.1093/emboj/cdg324
67. Tsukahara, J., Mukaiyama, K., Okuda, S., Narita, S. I. & Tokuda, H. Dissection of LolB function - Lipoprotein binding, membrane targeting and incorporation of lipoproteins into lipid bilayers. *FEBS J.* (2009). doi:10.1111/j.1742-4658.2009.07156.x
68. Konovalova, A. & Silhavy, T. J. Outer membrane lipoprotein biogenesis: Lol is not the end. *Philosophical Transactions of the Royal Society B: Biological Sciences* (2015). doi:10.1098/rstb.2015.0030
69. Koebnik, R., Locher, K. P. & Van Gelder, P. Structure and function of bacterial outer membrane proteins: Barrels in a nutshell. *Molecular Microbiology* (2000). doi:10.1046/j.1365-2958.2000.01983.x
70. Cho, S. H. *et al.* Detecting envelope stress by monitoring β-barrel assembly. *Cell* (2014). doi:10.1016/j.cell.2014.11.045
71. Ruiz, N., Falcone, B., Kahne, D. & Silhavy, T. J. Chemical conditionality: a genetic strategy to probe organelle assembly. *Cell* **121**, 307–17 (2005).
72. Voulhoux, R., Bos, M. P., Geurtsen, J., Mols, M. & Tommassen, J. Role of a highly conserved bacterial protein in outer membrane protein assembly. *Science* **299**, 262–5 (2003).
73. Wu, T. *et al.* Identification of a multicomponent complex required for outer membrane biogenesis in Escherichia coli. *Cell* **121**, 235–245 (2005).
74. von Heijne, G. The signal peptide. *The Journal of Membrane Biology* (1990). doi:10.1007/BF01868635
75. Driessen, A. J. M. & Nouwen, N. Protein Translocation Across the Bacterial Cytoplasmic Membrane. *Annu. Rev. Biochem.* (2008). doi:10.1146/annurev.biochem.77.061606.160747
76. Du Plessis, D. J. F., Nouwen, N. & Driessen, A. J. M. The Sec translocase. *Biochimica et Biophysica Acta - Biomembranes* (2011). doi:10.1016/j.bbamem.2010.08.016
77. Hegde, R. S. & Bernstein, H. D. The surprising complexity of signal sequences. *Trends in Biochemical Sciences* (2006). doi:10.1016/j.tibs.2006.08.004
78. Hagan, C. L., Silhavy, T. J. & Kahne, D. β-Barrel Membrane Protein Assembly by the Bam Complex. *Annu. Rev. Biochem.* (2011). doi:10.1146/annurev-biochem-061408-144611
79. Crooke, E. & Wickner, W. Trigger factor: a soluble protein that folds pro-OmpA into a membrane-assembly-competent form. *Proc. Natl. Acad. Sci. U. S. A.* (1987). doi:10.1073/pnas.84.15.5216

80. Valent, Q. A. *et al.* Early events in preprotein recognition in E. coli: interaction of SRP and trigger factor with nascent polypeptides. *EMBO J.* (1995). doi:10.1002/j.1460-2075.1995.tb00236.x
81. Valent, Q. A. *et al.* Nascent membrane and presecretory proteins synthesized in Escherichia coli associate with signal recognition particle and trigger factor . *Mol. Microbiol.* (1997). doi:10.1046/j.1365-2958.1997.4431808.x
82. Ferbitz, L. *et al.* Trigger factor in complex with the ribosome forms a molecular cradle for nascent proteins. *Nature* (2004). doi:10.1038/nature02899
83. Randall, L. L. & Hardy, S. J. S. SecB, one small chaperone in the complex milieu of the cell. *Cell. Mol. Life Sci.* (2002). doi:10.1007/PL00012488
84. Bechtluft, P., Nouwen, N., Tans, S. J. & Driessen, A. J. M. SecB - A chaperone dedicated to protein translocation. *Molecular BioSystems* (2010). doi:10.1039/b915435c
85. Zimmer, J., Nam, Y. & Rapoport, T. A. Structure of a complex of the ATPase SecA and the protein-translocation channel. *Nature* (2008). doi:10.1038/nature07335
86. Zwizinski, C. & Wickner, W. Purification and characterization of leader (signal) peptidase from Escherichia coli. *J. Biol. Chem.* (1980).
87. Sklar, J. G., Wu, T., Kahne, D. & Silhavy, T. J. Defining the roles of the periplasmic chaperones SurA, Skp, and DegP in Escherichia coli. *Genes Dev.* **21**, 2473–84 (2007).
88. Bitto, E. & McKay, D. B. The periplasmic molecular chaperone protein SurA binds a peptide motif that is characteristic of integral outer membrane proteins. *J. Biol. Chem.* (2003). doi:10.1074/jbc.M308853200
89. Hennecke, G., Nolte, J., Volkmer-Engert, R., Schneider-Mergener, J. & Behrens, S. The periplasmic chaperone SurA exploits two features characteristic of integral outer membrane proteins for selective substrate recognition. *J. Biol. Chem.* (2005). doi:10.1074/jbc.M413742200
90. Thoma, J., Burmann, B. M., Hiller, S. & Müller, D. J. Impact of holdase chaperones Skp and SurA on the folding of β-barrel outer-membrane proteins. *Nat. Struct. Mol. Biol.* (2015). doi:10.1038/nsmb.3087
91. Lazar, S. W. & Kolter, R. SurA assists the folding of Escherichia coli outer membrane proteins. *J. Bacteriol.* (1996). doi:10.1128/jb.178.6.1770-1773.1996
92. Rouvière, P. E. & Gross, C. A. SurA, a periplasmic protein with peptidyl-prolyl isomerase activity, participates in the assembly of outer membrane porins. *Genes Dev.* (1996). doi:10.1101/gad.10.24.3170
93. Vertommen, D., Ruiz, N., Leverrier, P., Silhavy, T. J. & Collet, J.-F. Characterization of the role of the Escherichia coli periplasmic chaperone SurA using differential proteomics. *Proteomics* **9**, 2432–43 (2009).
94. Bulieris, P. V., Behrens, S., Holst, O. & Kleinschmidt, J. H. Folding and insertion of the outer membrane protein OmpA is assisted by the chaperone Skp and by lipopolysaccharide. *J. Biol. Chem.* (2003). doi:10.1074/jbc.M211177200
95. Qu, J., Mayer, C., Behrens, S., Holst, O. & Kleinschmidt, J. H. The Trimeric Periplasmic Chaperone Skp of Escherichia coli Forms 1:1 Complexes with Outer Membrane Proteins via Hydrophobic and Electrostatic Interactions. *J. Mol. Biol.* (2007). doi:10.1016/j.jmb.2007.09.020
96. Burmann, B. M., Wang, C. & Hiller, S. Conformation and dynamics of the periplasmic membrane-protein-chaperone complexes OmpX-Skp and tOmpA-Skp. *Nat. Struct. Mol. Biol.* (2013). doi:10.1038/nsmb.2677
97. Callon, M., Burmann, B. M. & Hiller, S. Structural mapping of a chaperone-substrate interaction surface. *Angew. Chemie - Int. Ed.* (2014). doi:10.1002/anie.201310963
98. Walton, T. A. & Sousa, M. C. Crystal structure of Skp, a prefoldin-like chaperone that protects soluble and membrane proteins from aggregation. *Mol. Cell* (2004). doi:10.1016/j.molcel.2004.07.023

99. Lyu, Z. X., Shao, Q., Gao, Y. Q. & Zhao, X. S. Direct Observation of the Uptake of Outer Membrane Proteins by the Periplasmic Chaperone Skp. *PLoS One* (2012). doi:10.1371/journal.pone.0046068
100. Krojer, T., Garrido-Franco, M., Huber, R., Ehrmann, M. & Clausen, T. Crystal structure of DegP (HtrA) reveals a new protease-chaperone machine. *Nature* (2002). doi:10.1038/416455a
101. Krojer, T. *et al.* Structural basis for the regulated protease and chaperone function of DegP. *Nature* (2008). doi:10.1038/nature07004
102. Jiang, J. *et al.* Activation of DegP chaperone-protease via formation of large cage-like oligomers upon binding to substrate proteins. *Proc. Natl. Acad. Sci. U. S. A.* (2008). doi:10.1073/pnas.0805464105
103. Sklar, J. G. *et al.* 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 (2007).
104. Hagan, C. L., Kim, S. & Kahne, D. Reconstitution of outer membrane protein assembly from purified components. *Science* **328**, 890–892 (2010).
105. Gentle, I. E., Burri, L. & Lithgow, T. Molecular architecture and function of the Omp85 family of proteins. *Molecular Microbiology* (2005). doi:10.1111/j.1365-2958.2005.04906.x
106. Heinz, E. & Lithgow, T. A comprehensive analysis of the Omp85/TpsB protein superfamily structural diversity, taxonomic occurrence, and evolution. *Front. Microbiol.* (2014). doi:10.3389/fmicb.2014.00370
107. Webb, C. T., Heinz, E. & Lithgow, T. Evolution of the β-barrel assembly machinery. *Trends in Microbiology* (2012). doi:10.1016/j.tim.2012.08.006
108. Bakelar, J., Buchanan, S. K. & Noinaj, N. The structure of the β-barrel assembly machinery complex. *Science (80-.).* **351**, 180–186 (2016).
109. Gu, Y. *et al.* Structural basis of outer membrane protein insertion by the BAM complex. *Nature* (2016). doi:10.1038/nature17199
110. Han, L. *et al.* Structure of the BAM complex and its implications for biogenesis of outer-membrane proteins. *Nat. Struct. Mol. Biol.* (2016). doi:10.1038/nsmb.3181
111. Iadanza, M. G. *et al.* Lateral opening in the intact β-barrel assembly machinery captured by cryo-EM. *Nat. Commun.* **7**, 12865 (2016).
112. Malinverni, J. C. *et al.* YfiO stabilizes the YaeT complex and is essential for outer membrane protein assembly in Escherichia coli. *Mol. Microbiol.* **61**, 151–164 (2006).
113. Roman-Hernandez, G., Peterson, J. H. & Bernstein, H. D. Reconstitution of bacterial autotransporter assembly using purified components. *Elife* (2014). doi:10.7554/eLife.04234
114. Doerner, P. A. & Sousa, M. C. Extreme Dynamics in the BamA β-Barrel Seam. *Biochemistry* (2017). doi:10.1021/acs.biochem.7b00281
115. Hagan, C. L. & Kahne, D. The reconstituted Escherichia coli Bam complex catalyzes multiple rounds of β-barrel assembly. *Biochemistry* **50**, 7444–7446 (2011).
116. Plummer, A. M. & Fleming, K. G. BamA Alone Accelerates Outer Membrane Protein Folding in Vitro through a Catalytic Mechanism. *Biochemistry* (2015). doi:10.1021/acs.biochem.5b00950
117. Sampson, B. A., Misra, R. & Benson, S. A. Identification and characterization of a new gene of Escherichia coli K-12 involved in outer membrane permeability. *Genetics* **122**, 491–501 (1989).
118. Kim, S. *et al.* Structure and function of an essential component of the outer membrane protein assembly machine. *Science* **317**, 961–964 (2007).
119. Gatzeva-Topalova, P. Z., Walton, T. A. & Sousa, M. C. Crystal Structure of YaeT: Conformational Flexibility and Substrate Recognition. *Structure* (2008). doi:10.1016/j.str.2008.09.014

120. Gatzeva-Topalova, P. Z., Warner, L. R., Pardi, A. & Sousa, M. C. Structure and Flexibility of the Complete Periplasmic Domain of BamA: The Protein Insertion Machine of the Outer Membrane. *Structure* (2010). doi:10.1016/j.str.2010.08.012
121. Knowles, T. J. *et al.* Fold and function of polypeptide transport-associated domains responsible for delivering unfolded proteins to membranes. *Mol. Microbiol.* (2008). doi:10.1111/j.1365-2958.2008.06225.x
122. Harrison, S. C. Peptide-surface association: The case of PDZ and PTB domains. *Cell* (1996). doi:10.1016/S0092-8674(00)80105-1
123. Remaut, H. & Waksman, G. Protein-protein interaction through β-strand addition. *Trends in Biochemical Sciences* (2006). doi:10.1016/j.tibs.2006.06.007
124. Webb, C. T. *et al.* Dynamic association of BAM complex modules includes surface exposure of the lipoprotein BamC. *J. Mol. Biol.* (2012). doi:10.1016/j.jmb.2012.05.035
125. Ieva, R., Tian, P., Peterson, J. H. & Bernstein, H. D. Sequential and spatially restricted interactions of assembly factors with an autotransporter beta domain. *Proc. Natl. Acad. Sci. U. S. A.* **108**, E383-91 (2011).
126. Ricci, D. P., Hagan, C. L., Kahne, D. & Silhavy, T. J. Activation of the Escherichia coli β-barrel assembly machine (Bam) is required for essential components to interact properly with substrate. *Proc. Natl. Acad. Sci. U. S. A.* **109**, 3487–91 (2012).
127. Wzorek, J. S., Lee, J., Tomasek, D., Hagan, C. L. & Kahne, D. E. Membrane integration of an essential β-barrel protein prerequires burial of an extracellular loop. *Proc. Natl. Acad. Sci.* **114**, 2598–2603 (2017).
128. McCabe, A. L., Ricci, D., Adetunji, M. & Silhavy, T. J. Conformational changes that coordinate the activity of BamA and BamD allowing β-barrel assembly. *J. Bacteriol.* **199**, (2017).
129. Lee, J. *et al.* Substrate binding to BamD triggers a conformational change in BamA to control membrane insertion. *Proc. Natl. Acad. Sci. U. S. A.* **115**, 2359–2364 (2018).
130. Charlson, E. S., Werner, J. N. & Misra, R. Differential effects of yfgL mutation on Escherichia coli outer membrane proteins and lipopolysaccharide. *J. Bacteriol.* (2006). doi:10.1128/JB.00571-06
131. Konovalova, A., Mitchell, A. M. & Silhavy, T. J. A lipoprotein/b-barrel complex monitors lipopolysaccharide integrity transducing information across the outer membrane. *Elife* (2016). doi:10.7554/eLife.15276
132. Mahoney, T. F., Ricci, D. P. & Silhavy, T. J. Classifying β-barrel assembly substrates by manipulating essential bam complex members. *J. Bacteriol.* **198**, 1984–1992 (2016).
133. Noinaj, N. *et al.* Structural insight into the biogenesis of β-barrel membrane proteins. *Nature* (2013). doi:10.1038/nature12521
134. Ni, D. *et al.* Structural and functional analysis of the β-barrel domain of BamA from Escherichia coli. *FASEB J.* (2014). doi:10.1096/fj.13-248450
135. Noinaj, N., Kuszak, A. J., Balusek, C., Gumbart, J. C. & Buchanan, S. K. Lateral opening and exit pore formation are required for BamA function. *Structure* (2014). doi:10.1016/j.str.2014.05.008
136. Struyvé, M., Moons, M. & Tommassen, J. Carboxy-terminal phenylalanine is essential for the correct assembly of a bacterial outer membrane protein. *J. Mol. Biol.* (1991). doi:10.1016/0022-2836(91)90880-F
137. Robert, V. *et al.* Assembly factor Omp85 recognizes its outer membrane protein substrates by a species-specific C-terminal motif. *PLoS Biol.* (2006). doi:10.1371/journal.pbio.0040377
138. Kutik, S. *et al.* Dissecting Membrane Insertion of Mitochondrial β-Barrel Proteins. *Cell* (2008). doi:10.1016/j.cell.2008.01.028
139. Hagan, C. L., Wzorek, J. S. & Kahne, D. Inhibition of the β-barrel assembly machine by a peptide that binds BamD. *Proc. Natl. Acad. Sci.* **112**, 2011–2016 (2015).

140. Ieva, R. & Bernstein, H. D. Interaction of an autotransporter passenger domain with BamA during its translocation across the bacterial outer membrane. *Proc. Natl. Acad. Sci.* **106**, 19120–19125 (2009).
141. Pavlova, O., Peterson, J. H., Ieva, R. & Bernstein, H. D. Mechanistic link between β barrel assembly and the initiation of autotransporter secretion. *Proc. Natl. Acad. Sci. U. S. A.* **110**, E938-47 (2013).
142. Sikdar, R., Peterson, J. H., Anderson, D. E. & Bernstein, H. D. Folding of a bacterial integral outer membrane protein is initiated in the periplasm. *Nat. Commun.* (2017). doi:10.1038/s41467-017-01246-4
143. Doyle, M. T. & Bernstein, H. D. Bacterial outer membrane proteins assemble via asymmetric interactions with the BamA β-barrel. *Nat. Commun.* (2019). doi:10.1038/s41467-019-11230-9
144. Höhr, A. I. C. *et al.* Membrane protein insertion through a mitochondrial β-barrel gate. *Science (80-.).* **359**, eaah6834 (2018).
145. Kleinschmidt, J. H. & Tamm, L. K. Folding intermediates of a β-barrel membrane protein. Kinetic evidence for a multi-step membrane insertion mechanism. *Biochemistry* (1996). doi:10.1021/bi961478b
146. Kleinschmidt, J. H., Den Blaauwen, T., Driessen, A. J. M. & Tamm, L. K. Outer membrane protein A of Escherichia coli inserts and folds into lipid bilayers by a concerted mechanism. *Biochemistry* (1999). doi:10.1021/bi982465w
147. Huysmans, G. H. M., Baldwin, S. A., Brockwell, D. J. & Radford, S. E. The transition state for folding of an outer membrane protein. *Proc. Natl. Acad. Sci. U. S. A.* (2010). doi:10.1073/pnas.0911904107
148. Kleinschmidt, J. H., Bulieris, P. V, Qu, J., Dogterom, M. & den Blaauwen, T. Association of neighboring β-strands of outer membrane protein A in lipid bilayers revealed by site-directed fluorescence quenching. *J. Mol. Biol.* **407**, 316–32 (2011).
149. Kleinschmidt, J. H. Folding of β-barrel membrane proteins in lipid bilayers - Unassisted and assisted folding and insertion. *Biochimica et Biophysica Acta - Biomembranes* (2015). doi:10.1016/j.bbamem.2015.05.004
150. Danoff, E. J. & Fleming, K. G. Novel kinetic intermediates populated along the folding pathway of the transmembrane β-barrel OmpA. *Biochemistry* (2017). doi:10.1021/acs.biochem.6b00809
151. Ricci, D. P. & Silhavy, T. J. Outer Membrane Protein Insertion by the β-barrel Assembly Machine. *EcoSal Plus* (2019). doi:10.1128/ecosalplus.esp-0035-2018
152. Burgess, N. K., Dao, T. P., Stanley, A. M. & Fleming, K. G. β-Barrel proteins that reside in the Escherichia coli outer membrane in vivo demonstrate varied folding behavior in vitro. *J. Biol. Chem.* (2008). doi:10.1074/jbc.M802754200
153. Gessmann, D. *et al.* Outer membrane -barrel protein folding is physically controlled by periplasmic lipid head groups and BamA. *Proc. Natl. Acad. Sci.* (2014). doi:10.1073/pnas.1322473111
154. Danoff, E. J. & Fleming, K. G. Membrane defects accelerate outer membrane β-barrel protein folding. *Biochemistry* (2015). doi:10.1021/bi501443p
155. Schiffrin, B. *et al.* Effects of Periplasmic Chaperones and Membrane Thickness on BamA-Catalyzed Outer-Membrane Protein Folding. *J. Mol. Biol.* (2017). doi:10.1016/j.jmb.2017.09.008
156. Kim, K. H., Aulakh, S. & Paetzel, M. The bacterial outer membrane β-barrel assembly machinery. *Protein Science* (2012). doi:10.1002/pro.2069
157. Estrada Mallarino, L. *et al.* TtOmp85, a β-barrel assembly protein, functions by barrel augmentation. *Biochemistry* (2015). doi:10.1021/bi5011305
158. Schiffrin, B., Brockwell, D. J. & Radford, S. E. Outer membrane protein folding from an energy landscape perspective. *BMC Biology* (2017). doi:10.1186/s12915-017-0464-5

159. Braun, M. & Silhavy, T. J. Imp/OstA is required for cell envelope biogenesis in Escherichia coli. *Mol. Microbiol.* **45**, 1289–1302 (2002).
160. Chng, S. S., Ruiz, N., Chimalakonda, G., Silhavy, T. J. & Kahne, D. Characterization of the two-protein complex in Escherichia coli responsible for lipopolysaccharide assembly at the outer membrane. *Proc. Natl. Acad. Sci. U. S. A.* (2010). doi:10.1073/pnas.0912872107
161. Ureta, A. R., Endres, R. G., Wingreen, N. S. & Silhavy, T. J. Kinetic analysis of the assembly of the outer membrane protein LamB in Escherichia coli mutants each lacking a secretion or targeting factor in a different cellular compartment. *J. Bacteriol.* (2007). doi:10.1128/JB.01103-06
162. Bennion, D., Charlson, E. S., Coon, E. & Misra, R. Dissection of β-barrel outer membrane protein assembly pathways through characterizing BamA POTRA 1 mutants of Escherichia coli. *Mol. Microbiol.* (2010). doi:10.1111/j.1365-2958.2010.07280.x
163. Hagan, C. L., Westwood, D. B. & Kahne, D. Bam Lipoproteins Assemble BamA *in Vitro*. *Biochemistry* **52**, 6108–6113 (2013).
164. Schiffrin, B. *et al.* Effects of Periplasmic Chaperones and Membrane Thickness on BamA-Catalyzed Outer-Membrane Protein Folding. *J. Mol. Biol.* (2017). doi:10.1016/j.jmb.2017.09.008
165. Chin, J. W., Martin, A. B., King, D. S., Wang, L. & Schultz, P. G. Addition of a photocrosslinking amino acid to the genetic code of Escherichia coli. *Proc. Natl. Acad. Sci. U. S. A.* (2002). doi:10.1073/pnas.172226299
166. Bamert, R. S. *et al.* Structural basis for substrate selection by the translocation and assembly module of the β-barrel assembly machinery. *Mol. Microbiol.* (2017). doi:10.1111/mmi.13757
167. Leonard-Rivera, M. & Misra, R. Conserved Residues of the Putative L6 Loop of Escherichia coli BamA Play a Critical Role in the Assembly of β-Barrel Outer Membrane Proteins, Including That of BamA Itself. *J. Bacteriol.* **194**, 4662–4668 (2012).
168. Delattre, A. S. *et al.* Functional importance of a conserved sequence motif in FhaC, a prototypic member of the TpsB/Omp85 superfamily. *FEBS J.* **277**, 4755–4765 (2010).
169. Patel, G. J. & Kleinschmidt, J. H. The lipid bilayer-inserted membrane protein BamA of Escherichia coli facilitates insertion and folding of outer membrane protein A from its complex with Skp. *Biochemistry* (2013). doi:10.1021/bi400103t
170. Hayer-Hartl, M., Bracher, A. & Hartl, F. U. The GroEL-GroES Chaperonin Machine: A Nano-Cage for Protein Folding. *Trends in Biochemical Sciences* (2016). doi:10.1016/j.tibs.2015.07.009
171. Horwich, A. L. & Fenton, W. A. Chaperonin-mediated protein folding: Using a central cavity to kinetically assist polypeptide chain folding. *Quarterly Reviews of Biophysics* (2009). doi:10.1017/S0033583509004764
172. Weissman, J. S. *et al.* Mechanism of GroEL action: Productive release of polypeptide from a sequestered position under groes. *Cell* (1995). doi:10.1016/0092-8674(95)90098-5
173. Casadaban, M. J. Transposition and fusion of the lac genes to selected promoters in Escherichia coli using bacteriophage lambda and Mu. *J. Mol. Biol.* (1976). doi:10.1016/0022-2836(76)90119-4
174. Lee, J. *et al.* Formation of a β-barrel membrane protein is catalyzed by the interior surface of the assembly machine protein BamA. *Elife* **8**, (2019).
175. Metcalf, W. W., Jiang, W. & Wanner, B. L. Use of the rep technique for allele replacement to construct new Escherichia coli hosts for maintenance of R6Kλ origin plasmids at different copy numbers. *Gene* (1994). doi:10.1016/0378-1119(94)90776-5
176. Lutz, R. & Bujard, H. Independent and tight regulation of transcriptional units in Escherichia coli via the LacR/O, the TetR/O and AraC/I1-I2 regulatory elements. *Nucleic*

Acids Res. **25**, 1203–10 (1997).
177. Amann, E., Ochs, B. & Abel, K. J. Tightly regulated tac promoter vectors useful for the expression of unfused and fused proteins in Escherichia coli. *Gene* (1988). doi:10.1016/0378-1119(88)90440-4
178. Philippe, N., Alcaraz, J. P., Coursange, E., Geiselmann, J. & Schneider, D. Improvement of pCVD442, a suicide plasmid for gene allele exchange in bacteria. *Plasmid* (2004). doi:10.1016/j.plasmid.2004.02.003
179. Browning, D. F. *et al.* Mutational and topological analysis of the Escherichia coli BamA protein. *PLoS One* (2013). doi:10.1371/journal.pone.0084512
180. Gu, Y., Zeng, Y., Wang, Z. & Dong, C. BamA β16C strand and periplasmic turns are critical for outer membrane protein insertion and assembly. *Biochem. J.* (2017). doi:10.1042/BCJ20170636
181. Eppens, E. F., Nouwen, N. & Tommassen, J. Folding of a bacterial outer membrane protein during passage through the periplasm. *EMBO J.* (1997). doi:10.1093/emboj/16.14.4295
182. Leo, J. C., Oberhettinger, P. & Linke, D. Assessing the Outer Membrane Insertion and Folding of Multimeric Transmembrane β-Barrel Proteins. *Methods Mol. Biol.* **1329**, 157–67 (2015).
183. Ryu, Y. & Schultz, P. G. Efficient incorporation of unnatural amino acids into proteins in Escherichia coli. *Nat. Methods* (2006). doi:10.1038/nmeth864
184. Gibson, D. G. *et al.* Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat. Methods* **6**, 343–345 (2009).
185. White, S. H. & Wimley, W. C. Membrane protein folding and stability: Physical principles. *Annu. Rev. Biophys. Biomol. Struct.* (1999). doi:10.1146/annurev.biophys.28.1.319
186. Roseman, M. A. Hydrophobicity of the peptide CO···HN hydrogen-bonded group. *J. Mol. Biol.* (1988). doi:10.1016/0022-2836(88)90642-0
187. Lundquist, K., Bakelar, J., Noinaj, N. & Gumbart, J. C. C-terminal kink formation is required for lateral gating in BamA. *Proc. Natl. Acad. Sci.* (2018). doi:10.1073/pnas.1722530115
188. Mastronarde, D. N. Automated electron microscope tomography using robust prediction of specimen movements. *J. Struct. Biol.* (2005). doi:10.1016/j.jsb.2005.07.007
189. Zheng, S. Q. *et al.* MotionCor2: Anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nature Methods* (2017). doi:10.1038/nmeth.4193
190. Rohou, A. & Grigorieff, N. CTFFIND4: Fast and accurate defocus estimation from electron micrographs. *J. Struct. Biol.* (2015). doi:10.1016/j.jsb.2015.08.008
191. Wagner, T. *et al.* SPHIRE-crYOLO is a fast and accurate fully automated particle picker for cryo-EM. *Commun. Biol.* (2019). doi:10.1038/s42003-019-0437-z
192. Scheres, S. H. W. RELION: Implementation of a Bayesian approach to cryo-EM structure determination. *J. Struct. Biol.* (2012). doi:10.1016/j.jsb.2012.09.006
193. Zivanov, J., Nakane, T. & Scheres, S. H. W. A Bayesian approach to beam-induced motion correction in cryo-EM single-particle analysis. *IUCrJ* (2019). doi:10.1107/S205225251801463X
194. Punjani, A., Rubinstein, J. L., Fleet, D. J. & Brubaker, M. A. CryoSPARC: Algorithms for rapid unsupervised cryo-EM structure determination. *Nat. Methods* (2017). doi:10.1038/nmeth.4169
195. Morin, A. *et al.* Collaboration gets the most out of software. *Elife* (2013). doi:10.7554/eLife.01456
196. Pettersen, E. F. *et al.* UCSF Chimera - A visualization system for exploratory research and analysis. *J. Comput. Chem.* (2004). doi:10.1002/jcc.20084
197. Yang, J. *et al.* The I-TASSER suite: Protein structure and function prediction. *Nature Methods* (2014). doi:10.1038/nmeth.3213

198. Zhang, Y. I-TASSER server for protein 3D structure prediction. *BMC Bioinformatics* (2008). doi:10.1186/1471-2105-9-40
199. Roy, A., Kucukural, A. & Zhang, Y. I-TASSER: A unified platform for automated protein structure and function prediction. *Nat. Protoc.* (2010). doi:10.1038/nprot.2010.5
200. Emsley, P. & Cowtan, K. Coot: Model-building tools for molecular graphics. *Acta Crystallogr. Sect. D Biol. Crystallogr.* (2004). doi:10.1107/S0907444904019158
201. Adams, P. D. *et al.* PHENIX: A comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. Sect. D Biol. Crystallogr.* (2010). doi:10.1107/S0907444909052925
202. Chen, V. B. *et al.* MolProbity: All-atom structure validation for macromolecular crystallography. *Acta Crystallogr. Sect. D Biol. Crystallogr.* (2010). doi:10.1107/S0907444909042073
203. Storek, K. M. *et al.* Monoclonal antibody targeting the β-barrel assembly machine of Escherichia coli is bactericidal. *Proc. Natl. Acad. Sci. U. S. A.* (2018). doi:10.1073/pnas.1800043115
204. Hart, E. M. *et al.* A small-molecule inhibitor of BamA impervious to efflux and the outer membrane permeability barrier. *Proc. Natl. Acad. Sci. U. S. A.* (2019). doi:10.1073/pnas.1912345116
205. Luther, A. *et al.* Chimeric peptidomimetic antibiotics against Gram-negative bacteria. *Nature* (2019). doi:10.1038/s41586-019-1665-6
206. Imai, Y. *et al.* A new antibiotic selectively kills Gram-negative pathogens. *Nature* (2019). doi:10.1038/s41586-019-1791-1
207. Heinrich, S. U., Mothes, W., Brunner, J. & Rapoport, T. A. The Sec61p Complex Mediates the Integration of a Membrane Protein by Allowing Lipid Partitioning of the Transmembrane Domain. *Cell* (2000). doi:10.1016/S0092-8674(00)00028-3
208. Gentle, I., Gabriel, K., Beech, P., Waller, R. & Lithgow, T. The Omp85 family of proteins is essential for outer membrane biogenesis in mitochondria and bacteria. *J. Cell Biol.* (2004). doi:10.1083/jcb.200310092
209. Paschen, S. A. *et al.* Evolutionary conservation of biogenesis of beta-barrel membrane proteins. *Nature* **426**, 862–866 (2003).
210. Wiedemann, N. *et al.* Machinery for protein sorting and assembly in the mitochondrial outer membrane. *Nature* **424**, 565–571 (2003).
211. Töpel, M., Ling, Q. & Jarvis, P. Neofunctionalization within the Omp85 protein superfamily during chloroplast evolution. *Plant Signal. Behav.* (2012). doi:10.4161/psb.18677
212. Huang, W., Ling, Q., Bédard, J., Lilley, K. & Jarvis, P. In vivo analyses of the roles of essential Omp85-related proteins in the chloroplast outer envelope membrane. *Plant Physiol.* (2011). doi:10.1104/pp.111.181891
213. Walther, D. M., Papic, D., Bos, M. P., Tommassen, J. & Rapaport, D. Signals in bacterial β-barrel proteins are functional in eukaryotic cells for targeting to and assembly in mitochondria. *Proc. Natl. Acad. Sci. U. S. A.* (2009). doi:10.1073/pnas.0807830106
214. Walther, D. M., Bos, M. P., Rapaport, D. & Tommassen, J. The mitochondrial porin, VDAC, has retained the ability to be assembled in the bacterial outer membrane. *Mol. Biol. Evol.* (2010). doi:10.1093/molbev/msp294
215. Müller, J. E. N. *et al.* Mitochondria can recognize and assemble fragments of a β-barrel structure. *Mol. Biol. Cell* (2011). doi:10.1091/mbc.E10-12-0943
216. Gunasinghe, S. D. *et al.* The WD40 Protein BamB Mediates Coupling of BAM Complexes into Assembly Precincts in the Bacterial Outer Membrane. *Cell Rep.* (2018). doi:10.1016/j.celrep.2018.04.093
217. Wang, Y. *et al.* A supercomplex spanning the inner and outer membranes mediates the biogenesis of β-barrel outer membrane proteins in bacteria. *J. Biol. Chem.* (2016). doi:10.1074/jbc.M115.710715

218. Choi, U. & Lee, C. R. Antimicrobial agents that inhibit the outer membrane assembly machines of gram-negative bacteria. *Journal of Microbiology and Biotechnology* (2019). doi:10.4014/jmb.1804.03051