Roles of periplasmic chaperones and BamA in outer membrane protein folding

Bob Schiffrin

Submitted in accordance with the requirements for the degree of

Doctor of Philosophy

The University of Leeds

Faculty of Biological Sciences

September 2016

The candidate confirms that the work submitted is his own, except where work which has formed part of jointly-authored publications has been included. The contribution of the candidate and the other authors to this work has been explicitly indicated below. The candidate confirms that appropriate credit has been given within the thesis where reference has been made to the work of others.

Jointly authored publications

Throughout this thesis the work directly attributable to the candidate is as follows:

(i) Literature research and compilation of the manuscript stated above.

(ii) The candidate performed all the experimental work and data analysis unless otherwise stated.

Chapters 3 and 4 contain work from the following publication:

Schiffrin, B., Calabrese, A. N., Devine, P. W. A., Harris, S. A., Ashcroft, A. E., Brockwell, D. J. & Radford, S. E. Skp is a multivalent chaperone of outer-membrane proteins. *Nat Struct Mol Biol* **23,** 786–793 (2016).

In this publication B. Schiffrin designed and performed the kinetic experiments, computer modelling and molecular dynamics simulations. A. N. Calabrese designed and performed the mass-spectrometry and cross-linking experiments. P. W. A. Devine designed and performed initial *in vacuo* apo-Skp simulations. S. A. Harris assisted and supervised the molecular dynamics simulations. S. E. Radford, D. J. Brockwell, A. E. Ashcroft and S. A. Harris provided help with scientific discussions, data interpretation and manuscript preparation.

Acknowledgements

Firstly, I would like to thank my supervisors, Prof Sheena Radford and Dr David Brockwell for always generously sharing their expertise, and their unwavering support, encouragement and enthusiasm for this project. Everyone in the Radford/Brockwell lab has been kind, friendly and ready with advice on any technique. The expertise of Dr Sophia Goodchild in lipids and fluorescence was very helpful. There was great solidarity and encouragement in the early molecular biology stages of the project from Dr Chris Wilson and Rhys Thomas. I am also grateful to Rhys for letting me use his meticulously collected list of reagent manufacturers, saving me a time consuming, and not amazingly exciting, job. Many thanks also to Dr Theo Karamanos for all his encouragement and advice throughout my PhD, in particular on computational matters. Lab manager Nasir Khan has fed me bananas and curry to keep me going at just the right times, as well as sorting out a huge number of technical problems for me, and just generally being a model human being.

I'm also grateful to Dr Sarah Harris for all her assistance with the molecular dynamics simulations, and her infectious enthusiasm for the Skp:OMP project. Thank you also to Dr Emanuele Paci for all his training and help with molecular dynamics simulations. I would like to thank everyone in the High Performance Computing (HPC) team at the University of Leeds whose assistance was invaluable, particularly when our reviewers wanted additional (very computationally expensive!) simulations to be performed.

Part of the current project was performed in collaboration with mass spectrometrists at the University of Leeds. I am grateful to Prof Alison Ashcroft and Dr Anton Calabrese for a very fruitful collaboration. I have learned an enormous amount from Anton, and I thank him for always making time for helpful discussions and advice.

I am also grateful to those who have previously worked on outer membrane protein (OMP) folding in the lab. Dr Gerard Huysmans and Dr Alice Bartlett were very willing to help out with this current project. I owe a huge debt of gratitude to Lindsay McMorran who gave me a huge amount of her time and expertise; her guidance on all scientific matters in the early stages of this project was invaluable. Thanks to Tom Crosskey who made a great contribution to this project, and the lab, during his year as a master's student. Also, a big thanks to all the OMP team: Tom Watkinson,

Anton Calabrese, Julia Humes, Anna Higgins, Matt Iadanza, and Jim Horne. It's been a joy to see the team grow and strengthen as my project has progressed.

As part of my PhD training I had the opportunity to spend a great three months at the Research Complex at Harwell, Oxford. I am grateful to Dr Stephen Carr and Prof Simon Phillips for this opportunity, and all in the Carr group. In particular, a big thanks to Dr Avinash Punekar, who looked after me during my time in Oxford.

The Astbury Centre for Structural Molecular Biology at the University of Leeds is a fantastic research environment, and I am grateful to all members for creating such a supportive and inspiring place to learn and develop scientific ideas. I am also grateful to the BBSRC for funding my studentship, thereby making what follows possible. It's a real privilege to have the opportunity to try and put your brick in the wall of knowledge.

My parallel life as a musician has been a welcome distraction during my studies, so thank you to all my musical friends, for helping me to relax and have fun away from the lab. The joy of playing music with people has definitely contributed to sanity maintenance. In particular, a big thanks to members of Biscuithead and the Biscuit Badgers for being so much fun on our adventures around the country. Also, thanks to DJ Wonky Chris who provided weird analogue synth music to accompany molecular dynamics simulations of the Skp chaperone.

I have discovered that doing a PhD is a difficult thing, which I guess should not have come as a surprise. Given this, the love and support of those friends and family close to me has, as always, been invaluable. Thanks to my mum, dad and sister for their (lifelong!) love, support, and encouragement. The help and advice of my brilliant sister Mandy Schiffrin (Moo), in particular her wise words of encouragement to get writing up, meant that I, thankfully, didn't have to pull an all-nighter to get it done in the end! Finally, thanks for all the support from my amazing wife Joyce Lee Nga Ya, who, by making me happy every day, has contributed immeasurably to this work.

Abstract

A defining feature of living things is that they have an inside and an outside, and in order for all living cells to survive, whether they are part of a blue whale or a unicellular microscopic organism, they must have mechanisms to mediate exchange with their environment. Food and energy enters the cell, and material must also leave, such as the waste products of metabolism, or virulence factors from pathogenic organisms. Lipid membranes define these boundaries, but it is membrane proteins that mediate the exchange. Although lipid bilayers can self-assemble *in vitro*, the assembly of complex biological membranes containing proteins requires energy and careful coordination.

The work presented in this thesis examines the biogenesis pathway of β-barrel outer membrane proteins (OMPs) in Gram-negative bacteria. OMPs are synthesised on cytosolic ribosomes, translocated across the inner membrane, then chaperoned across the periplasm, before folding and insertion into the OM. While OMPs can fold spontaneously into lipid membranes, this process is too slow to be biologically relevant, so a dedicated folding catalyst, the β-barrel assembly machinery (BAM) complex, is required at the OM. Recent genetic, structural and biochemical investigations have increased our understanding of OMP assembly, but key questions remain, including: How do periplasmic chaperones bind and release OMP substrates? What are the roles and interactions of BAM subunits? What is the molecular mechanism by which BAM folds and inserts OMPs?

Here, an assay was developed to monitor OMP folding kinetics *in vitro* using intrinsic fluorescence in low concentrations of urea (0.24 M). This allowed comparison of the real-time folding behaviour of different OMPs under the same conditions for the first time (Chapter 3). The assay was then successfully extended to include OMP assembly factors, including the periplasmic chaperones Skp and SurA, and BamA, the principal component of the BAM complex, to obtain the following key results:

Investigations into the interactions between Skp and OMPs of varying size (tOmpA, PagP, OmpT, OmpF and tBamA) revealed that greater Skp:OMP ratios are required to prevent the folding of 16-stranded OMPs compared with smaller 8-stranded OMPs. Supported by ion mobility spectrometry-mass spectrometry (IMS-MS) data, computer modelling and molecular dynamics simulations, the results imply a new

mechanism for Skp chaperone activity involving the coordination of multiple copies of Skp to protect a single substrate from aggregation (Chapter 4).

Addition of further folding factors to the assay demonstrated that the model OMP tOmpA can be released and folded from its complex with Skp by BamA, possibly recapitulating an *in vivo* assembly pathway. BamA consists of a β-barrel membrane-embedded domain and soluble periplasmic domains, and while the release activity was shown to located in the membrane domain, the activity was greatest when full-length BamA was present. By contrast, SurA was not able to release tOmpA from Skp under the conditions employed, arguing against a sequential chaperone model (Chapter 5).

Next, kinetic studies were used to investigate the mechanism of OMP folding catalysis by the BAM complex. The effect of hydrophobic mismatch between the BamA β-barrel and the membrane was examined by monitoring the folding of tOmpA into liposomes containing lipids of different chain lengths in the presence or absence of BamA. The results showed that BamA has a greater catalytic effect in lipids with longer chain lengths, with the largest rate enhancement achieved in bilayers with a hydrophobic thickness close to that of the OM. The results establish the importance of hydrophobic mismatch in the mechanism by which OMPs are folded *in vivo*, which may be influenced by local thinning of the membrane and increases in lipid disorder in the vicinity of the BAM complex (Chapter 5).

Finally, based on the results obtained in this project, and consideration of the currently available literature, a new 'barrel elongation' model is proposed for the mechanism of OMP assembly by the BAM complex (Chapter 6).

The OMP assembly pathway is an attractive target for novel antibacterials given that it is surface located, highly conserved, and essential in clinically important pathogens. Understanding the molecular mechanisms of OMP biogenesis factors will facilitate the development of drugs targeting this pathway. The work described in this thesis provides new insights into the mechanisms of OMP assembly, using a wide range of biochemical and biophysical techniques, thereby contributing to the development of this fast-moving and fascinating field.

Table of Contents

List of Figures

List of Tables

Abbreviations

AAA	ATPases associated with diverse cellular activities
Apol	Amphipol
AFM	Atomic force microscopy
Au	Absorbance units
AUC	Analytical ultracentrifugation
BAM	β-barrel assembly machinery
BCA	Bicinchoninic acid
BME	β-mercaptoethanol
BN-PAGE	Blue native polyacrylamide gel electrophoresis
Bpa	p-benzoyl-L-phenylalanine
bR	Bacteriorhodopsin
BSA	Bovine serum albumin
CD	Circular dichroism
CL	Cardiolipin
cmc	critical micelle concentration
cOHB	Oligo-(R)-3-hydroxybutyrates
CCS	Collisional cross-section
cryoEM	Cryo-electron microscopy
dH_2O	Deionised water (greater than 15 MΩ resistance)
DDPC	Didecanoyl phosphatidylcholine
DDPE	Didecanoyl phosphatadylethanolamine
DHFR	Dihydrofolate reductase
DLPE	Dilauroyl phosphatadylethanolamine
DLPS	Dilauroyl phosphatadylserine
DMPC	Dimyristoyl phosphatidylcholine
DNA	Deoxyribonucleic acid
dNTP	Deoxyribonucleotide triphosphate
DOPC	Dioleoyl phosphatidylcholine
DPPC	Dipalmitoyl phosphatidylcholine
dsDNA	Double stranded deoxyribonucleic acid
DTPC	Ditridecanoyl phosphatidylcholine
DUPC	Diundecanoyl phosphatidylcholine
EDTA	Ethylenediaminetetraacetic acid
EM	Electron microscopy

ER	Endoplasmic reticulum
ESI	Electrospray ionisation
EtOH	Ethanol
FA	Facial amphipol
FlAsh	4'5'-bis(1,3,2-dithioarsolan-2-yl)fluorescein
FPOP	Fast photochemical oxidation of proteins
FRET	Förster resonance energy transfer
FTIR	Fourier transform infrared spectroscopy
GuHCl	Guanidine hydrochloride
HT	His-tagged
HDX	Hydrogen-deuterium exchange
HSQC	Heteronuclear single quantum coherence
IM	Inner membrane
IMS	Ion mobility spectrometry
ITC	Isothermal calorimetry
IPTG	Isopropyl-β-d-thiogalactoside
K_d	Equilibrium dissociation constant
KTSE	Kinetics of tertiary structure formation determined by electropheresis
LB	Lysogeny broth
LCFA	Long chain fatty acid
LDAO	Lauryldimethylamine-N-oxide
LMPG	Lyso-myristoylphosphatidylglycerol
Lol	Localisation of lipoproteins
LPR	Lipid to protein ratio
LPS	Lipopolysaccharide
Lpt	Lipopolysaccharide transport
LUV	Large unilamellar vesicle
MBP	Maltose-binding protein
MDR	Multidrug resistance
Mla	Maintanence of OM lipid asymmetry
MRE	Mean residue ellipticity
MRW	Mean residue weight
MS	Mass spectrometry
MST	Microscale thermopheresis
MWCO	Molecular weight cut-off

NAO	10-nonyl acridine orange
NCS	Nonconvential surfactant
NMR	Nuclear magnetic resonance
NOE	Nuclear Overhauser effect
OD	Optical density
OM	Outer membrane
OMP	Outer membrane protein
ORF	Open reading frame
PAGE	Polyacrylamide gel electrophoresis
PagP	PhoP/Q activated gene P
PC	Phosphocholine
PCR	Polymerase chain reaction
PDB	Protein data bank
PE	Phosphatidylethanolamine
PG	Phosphatidylglycerol
PMSF	Phenylmethanesulfonylfluoride
POTRA	Polypeptide transport-associated
PPIase	Peptidyl-prolyl isomerise
PPIs	Protein-protein interactions
PtdEtn	Phosphatidylethanolamine
RDG	Rayleigh-Gans-Debye
R_g	Radius of gyration
RMSD	Root mean square deviation
SANS	Small angle neutron scattering
SBDD	Structure based drug design
SDS	Sodium dodecyl sulphate
SDS-PAGE	Sodium dodecyl sulphate polyacrylamide gel electrophoresis
SEC	Size exclusion chromatography
SEC-MALS	Size exclusion chromatography multiangle light scattering
SER	Smooth endoplasmic reticulum
Skp	Seventeen kilodalton protein
SMFS	Single molecule force spectroscopy
SurA	Survival protein A
SUV	Small Unilamellar Vesicle
TAE	Tris-acetate-EDTA
TCA	Trichloroacetic acid

TDQF	Time-resolved distance determination by fluorescence quenching
TEMED	Tetramethylethylenediamine
TEV	Tobacco etch virus
TF	Trigger factor
ThT	Thioflavin T
TLR	Toll-Like Receptor
Tris	Tris (hydroxymethyl)-aminomethane
TROSY	Transverse-relaxation optimised spectroscopy
UV	Ultra violet
VDAC	Voltage-dependent anion channel
WT	Wild-type

1 Introduction

"One should really consider the sequence of a protein molecule about to fold into a precise geometric form as a line of melody written in a canon form and so designed by Nature to fold back into itself, creating harmonic chords of interaction consistent with biological function."

Chris Anfinsen, 1964

1.1 Life is defined by membranes

Despite the incredible diversity of forms and lifestyles that characterise life on Earth, at the molecular level living things are comprised of surprisingly few building blocks. In the ~350 years since the extraordinary discovery that living things are composed of microscopic cells which are invisible to the naked eye[1], we have learned that cells are mostly aqueous solutions of proteins, nucleic acids and carbohydrates, encapsulated by lipid membranes. In order to survive and replicate living things must exchange material, energy and information with their surroundings, and this exchange across lipid bilayers is mediated by membrane-embedded proteins. Thus, lipid membranes are fundamental to life, carrying out a vast array of functions. They separate organisms from the exterior world and allow the compartmentalisation of cellular processes in membrane-bound organelles within eukaryotic cells, such as lipid biosynthesis in the Smooth Endoplasmic Reticulum (SER), enzymatic digestion in lysosomes, and energy production in mitochondria. Membrane proteins (MPs) account for 20-30% of the Open Reading Frames (ORFs) within organisms[2] and make up over 50% of potential human drug targets[3].

For proteins to fold stably into the hydrophobic environment of a biological membrane their structure must fulfil two energetic requirements: (1) the hydrogen bonding potential of their polar backbone carbonyl and NH groups at their amide bonds must be mostly satisfied[4] to offset the energetic cost of peptide bond burial (~1.2 kcal/mol[5]); and (2) the amino acid side chain groups in contact with the acyl chains of the lipid bilayer must be predominantly hydrophobic. The secondary structures of α-helical and β-barrel membrane proteins allow them to solve these problems in different ways[6,7]. The residues within each helix in α-helical membrane proteins make backbone hydrogen bonds with each other, allowing the separate insertion of helices which can subsequently laterally associate with one another[8]. In contrast, β-barrel outer membrane proteins (OMPs) form cylindrical structures by

making hydrogen bonds between β-strand residues, potentially far from each other in sequence[9].

This project investigates the folding of β-barrel OMPs and, in particular, the role of the periplasmic chaperones Skp and SurA[10], and BamA[11], the main component of the β-barrel Assembly Machinery (BAM) complex which catalyses OMP folding and insertion into the outer membrane of Gram-negative bacteria[9,12-14]. There is considerable interest in the OMP assembly pathway as a potential target for novel antibiotics[9]. The BAM complex in particular is an attractive target given that it is surface located, highly conserved, and essential for viability[9,13,15-18]. This introduction will discuss the Gram-negative cell envelope and the OMP biogenesis pathway, as well as review recent genetic, structural and biochemical investigations which have dramatically increased our understanding of *in vivo* and *in vitro* OMP assembly over the last ~20 years.

1.2 The Gram-negative cell envelope

Gram-negative bacteria are surrounded by two membranes, the inner membrane (IM) which encloses the cytoplasm, and the outer membrane (OM) which is exposed to the extracellular space[19]. Between these is the periplasm, a viscous aqueous space (~0.14 μm^3)[20] containing soluble proteins and the peptidoglycan cell wall[21]. The distance between the IM and the OM can be inferred from the structures of machineries which span the periplasm which gives estimates in the range ~165-70 Å[17,19] to ~210 Å[22], the latter consistent with the dimensions observed by electron microscopy[23,24].

In contrast to the complexity of lipid species present in eukaryotic membranes[25], there are only three lipid types with an abundance >1% in *Escherichia coli* (*E. coli*)[26]: phosphatidylethanolamine (PE), phosphatidylglycerol (PG), and cardiolipin (CL) (Figure 1.1). Deletion of genes in lipid biosynthesis pathways have shown that *E. coli* are still viable in the absence of CL[27,28]. The absence of CL leads to an enrichment of the other anionic lipid PG[27,28], suggesting that CL may be functionally substituted by PG[26]. Strains lacking PE are also viable[29,30], a surprising result given the high content of PE in wild-type strains (70-80%), and that loss of PE is associated with misfolding and incorrect orientation of some membrane proteins, such as LacY[31]. Cells which lack PE have an absolute requirement for divalent cations, possibly to aid neutralisation of the additional charges on the membrane

due to the increased levels of anionic PG and CL[29,30]. Taken together, these results imply that it is the general physical properties of the lipids, rather than the specific types, that is important for both membrane and membrane protein biogenesis[26]. Accordingly, *E.* coli cannot survive when lacking both PE and CL[29]. PE and CL are both cone-shaped lipids, with the cross-sections of their headgroups smaller than that of their acyl chains[32] (Figure 1.1). As the shape of lipids within a bilayer is key to determining the physical properties of the membrane[33], it is likely that the lethality due to the removal of both PE and CL is due to a lack of lipids with these biophysical properties[26].

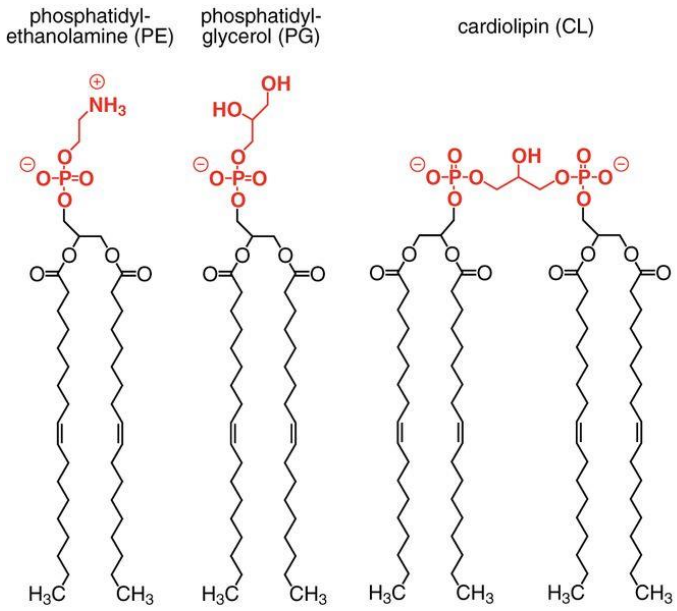


Figure 1.1: The three major lipid species in *E. coli.* Structures of PE, PG and CL. Each lipid is shown with dioleoyl acyl chains. Different chain lengths (C14-C18)[34] with different levels of saturation are found *in vivo* as discussed below. Image adapted from Oliver *et al.* (2014)[35].

The PE and PG components of bacterial membranes are thought to segregate into distinct domains[36], and the overall spatial distribution of phospholipids within *E. coli* is asymmetric[35]. Anionic lipids (PG and CL) are enriched at the poles, as determined by fluorescence microscopy experiments using the fluorescent dye 10-nonyl acridine orange (NAO), which specifically binds PG and CL[35], and quantitative liquid chromatography–mass spectrometry (LC-MS)[35]. This may have important consequences for the spatial control of OMP biogenesis, given that fluorescence microscopy experiments have recently demonstrated that nascent OMPs are inserted into the OM away from polar regions of the cell[37].

The IM is a classical bilayer composed of phospholipids (70% phosphatidylethanolamine (PE), 25% phosphatidylglycerol (PG), >5% cardiolipin[21,38] (Figure 1.1)) with acyl chain lengths of mostly C16 and C18 in both leaflets[34,39]. Inner membrane proteins (IMPs) are α-helical membrane proteins which constitute ~20% of the ~4000 ORFs in the *E. coli* genome[40]. IMPs are targeted to the SecYEG translocon by a hydrophobic N-terminal signal anchor sequence which is recognised by the Signal Recognition Particle (SRP) as the sequence emerges from the ribosome[40]. At the membrane, IMPs are assembled by the co-translational sequential insertion of α-helices into the IM by SecYEG using a lateral gating mechanism, with energy provided by the SecA ATPase[41]. A few lipoproteins are also associated with the IM, attached via an N-terminal lipid moiety[42].

The peptidoglycan layer (~6-7 nm in *E. coli* as measured by SANS[43], AFM[44], and cryoEM[24]) forms a mesh composed of saccharides cross-linked by short peptides[45] which is permeable to proteins under 100 kDa[46]. Consistent with its structural role in maintaining cell shape and integrity, the cell wall is resistant to extremes of pH and temperature, and is extremely strong (able to withstand ~3 atm of turgor pressure)[47]. Despite this, the peptidoglycan layer is highly elastic, able to reversibly expand or contract up to three-fold[48]. As well as being crucial for growth and division[48], this ability to deform parallel to the long axis of the cell[44] may also be important for allowing passage of components of the OM and their escorts across the periplasm.

The OM forms a highly impervious selective permeability barrier which protects against detergents and toxic substances[49]. Crucially, the OM must be assembled in an environment without an external energy source, as the periplasm is devoid of ATP[50]. In contrast to the IM, the OM is an asymmetric bilayer with the inner and outer leaflets composed of phospholipids and lipopolysaccharide (LPS), respectively. The maintenance of OM asymmetry is aided by the Mla (maintanence of OM lipid asymmetry) complex, which couples ATP hydrolysis at the IM to the removal of mislocalised phospholipids from the outer leaflet of the OM[51,52]. The phospholipid content in the OM is similar to that of the IM, but is slightly enriched in PE (~80%) and lipids with saturated acyl chains[34,39]. LPS is a glycolipid typically composed of three regions: Lipid A, a dissacharide linked to six saturated fatty acyl chains (mostly of C14 acyl chain length in *E. coli*[53,54]), a core oligosaccharide region, and a long polysaccharide termed the O-antigen (Figure 1.2). The O-antigen is highly variable between Gram-negative species[55], and also between organisms of the same species[56].

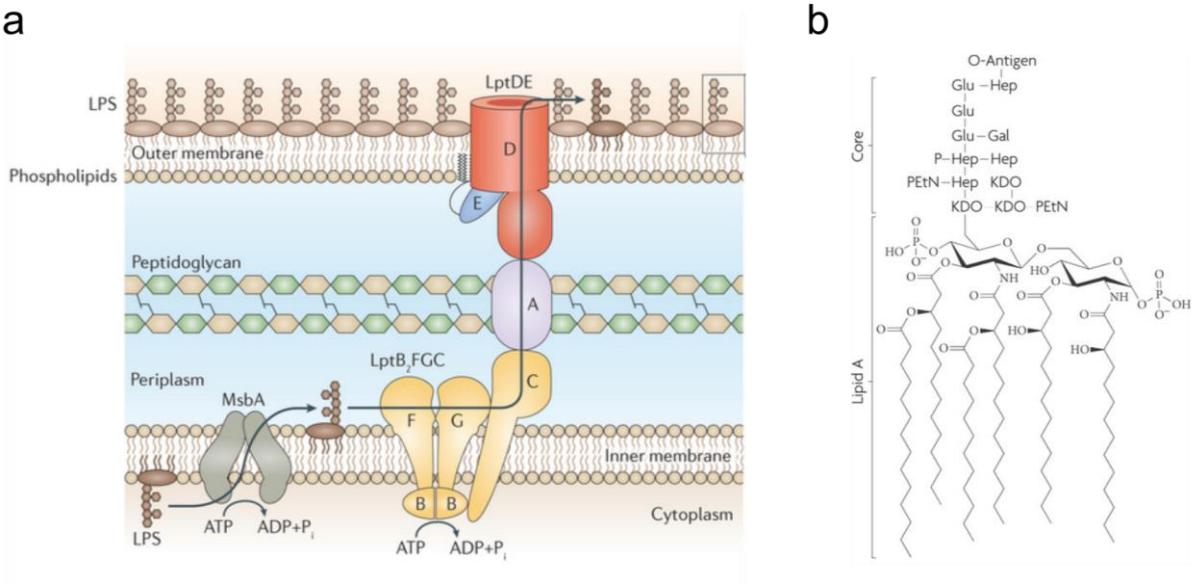


Figure 1.2: LPS biogenesis pathway in *E. coli.* (a) LPS is synthesised on the inner leaflet of the IM, flipped across the membrane by MsbA, assisted across the periplasm by LptA, B, C, and F, before insertion into the outer leaflet of the OM by the LptDE complex. Image from Okuda *et al*. (2016)[57] (b) Structure of LPS from *E. coli* K12. Kdo, 3-deoxy-D-*manno*-oct-2-ulosonic acid; Hep, L-*glycero*-D-*manno*-heptose; Glu, D-glucose; Gal, D-galactose; EtN, ethanolamine; P, phosphate. Image reproduced from Ruiz *et al.* (2009)[58].

Phosphate groups attached to the sugar moieties in LPS (Figure 1.2) are involved in coordinating divalent metal cations, such as Mg^{2+} and Ca^{2+}, allowing LPS molecules to pack closely together[57]. This packing of LPS saccharide units, which can number up to 600[56], protects Gram-negative bacteria against hydrophobic molecules, such as antibiotics[57]. LPS is essential in most Gram-negative bacteria, as are the genes responsible for its transport to the OM (LptA-G in *E. coli*)[57] (Figure 1.2), and stress responses are initiated when its integrity is compromised[59,60]. As well as performing a protective role, LPS also appears to be important in the assembly of at least some OMPs *in vivo*[61]. Consistent with this, *in vitro* work demonstrated that the 16-stranded porin PhoE folds with greater efficiency in phospholipid/LPS bilayers than in phospholipid/phospholipid bilayers[62]. The high density of negative charge on LPS molecules could explain why OMPs tend to have an excess of positive charge on their extracellular side (‘positive-outside’ rule)[63], in contrast to membrane proteins inserted into the Endoplasmic Reticulum (ER) and the IM which obey a ’positive-inside’ rule[64,65]. Further, a recent computational study predicted that LPS plays an important role in contributing to OMP thermodynamic stability[66].

In addition to the OMPs embedded within it (Section 1.2.1), the *E. coli* OM has ~100 associated lipoproteins, most of which are of unknown function[19,67]. Lipoproteins have their own dedicated transport system, the localisation of lipoproteins (Lol) system[68] (Figure 1.3). As lipoproteins emerge from the Sec translocon their signal peptides are cleaved by signal peptidase II (SPase II) adjacent to an N-terminal cysteine residue, which is then modified by the covalent attachment of membrane anchoring acyl chains[18]. If the residue C-terminal to the cysteine is not aspartate (the signal for IM retention[42]), the lipoprotein is released from the IM by the LolCDE complex, bound by LolA as it crosses the periplasm, and assisted into the OM by LolB[69]. The majority of OM lipoproteins are located in the periplasm, such as the major lipoprotein Lpp (Braun's lipoprotein) which covalently attaches to the peptidoglycan layer[70]. Lipoproteins can also be exposed on the surface of cells, as demonstrated in *E. coli* for RcsF[71], which senses defects in the cell envelope[60], the BAM complex member BamC[37,72], and the free-form of Lpp[73]. Some lipoproteins have highly specialised roles, such as LptE, which is important in the assembly, and possibly function, of the 26-stranded OMP LptD, which inserts LPS into the outer leaflet of the OM[57].

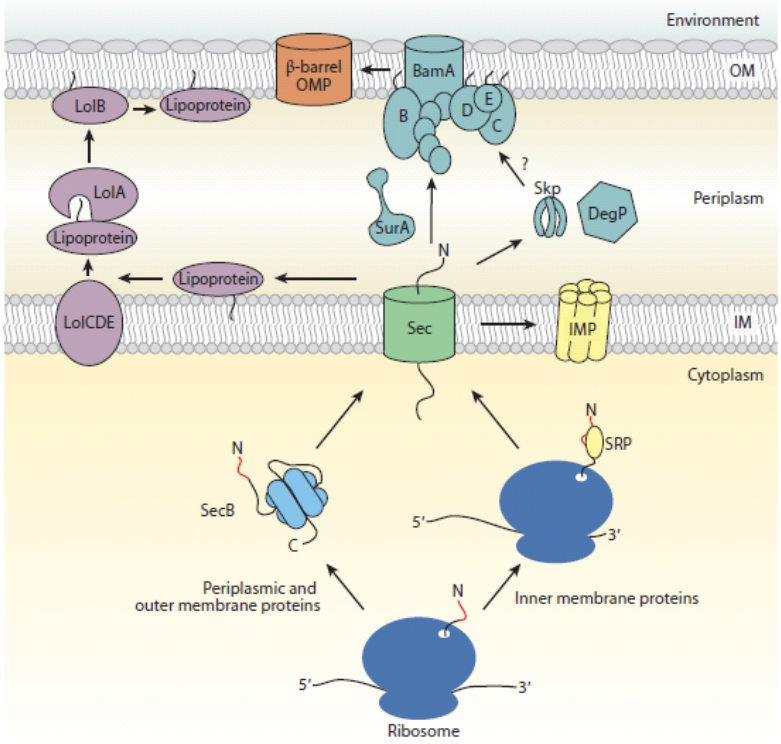


Figure 1.3: Pathways for the biogenesis of inner membrane proteins (IMPs), outer membrane proteins (OMPs) and lipoproteins. All three are synthesised in the cytoplasm and interact with the Sec translocon. IMPs are laterally released into the membrane. Lipoproteins are transported to the OM by the Lol machinery. OMPs are guided by periplasmic chaperones to the BAM complex which catalyses their folding and insertion into the OM. Image reproduced from Hagan *et al.* (2011)[9].

1.2.1 Outer membrane proteins (OMPs)

The OM is densely packed with proteins (protein:phophospholipid:LPS ratio of 5:1:1 (*w/w*))[74] which are involved in a myriad of functions including the uptake of nutrients, release of waste materials, secretion of virulence factors, and resistance to host defence systems[68]. OM proteins are almost excusively β-barrel membrane proteins (Figure 1.4), with the notable exception of the Wza octamer, which forms an α-helical barrel in the membrane[75]. The only other biological membranes which contain OMPs are the outer membranes of mitochondria and chloroplasts, consistent with the hypothesis that these organelles have an endosymbiotic origin[76].

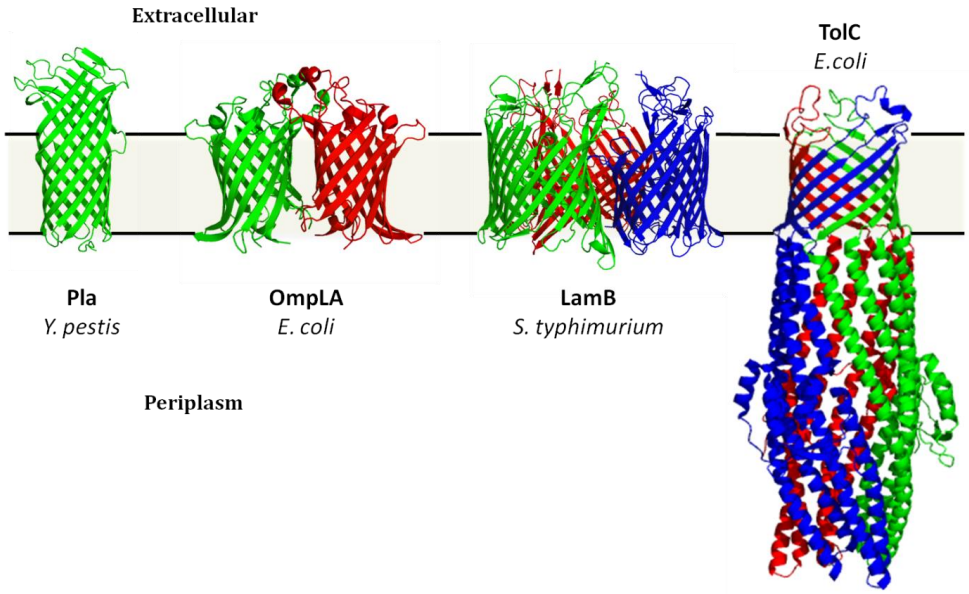


Figure 1.4: Examples of β-barrel outer membrane proteins from different Gram-negative organisms. OMPs can be found as monomers (Pla from *Yersinia pestis*; PDB:2X55[77]), multimers (dimeric OmpLA from *E. coli;* PDB: 1QD6[78], and trimeric LamB from *Salmonella typhimurium*; PDB: 2MPR[79]), or oligomers in which the β-barrel is made from several subunits (TolC from *E. coli*; PDB: 1EK9[80]). Image created with PyMOL, based on an image from Kim *et al.* (2012)[13] .

Approximately 2-3% of genes in Gram-negative bacteria are β-barrel OMPs[81] which are grouped into six families: (1) porins such as OmpC and PhoE; (2) passive transporters such as FadL; (3) active transporters such as the FhuA siderophore transporter and the BtuB vitamin B_{12} transporter; (4) enzymes such as the palmitoyl transferase PagP and the protease OmpT; (5) defensins such as OmpX; and (6) structural proteins such as OmpA[68]. Only two OMPs are known to be essential in *E. coli*, LptD (Figure 1.2) and BamA (Section 1.4.1). The trimeric porins are particularly abundant, covering ~70% of the membrane surface in *Roseobacter denitrificans*[82]. Non-specific 16-stranded porins such as OmpF and OmpC are present in ~10^5 copies per cell[83] and allow the passive diffusion of small hydrophilic molecules (up to ~600 Da)[68]. A consequence of the OM being porous to small ions, including protons, is that the pH of the periplasm is dictated by that of the surrounding medium[84]. Some passive transporters have substrate specificity such as PhoE (16-stranded), an anion selective porin[85], and LamB (18-stranded) which transports maltosaccharides across the OM[86]. Additionally, for essential molecules in low abundance in the external environment, active transport systems are used. For example, in the TonB system receptors, such as the 22-stranded OMPs BtuB, FepA

and FhuA, use energy from the IM via a complex energy-transducing system in order to import nutrients such as vitamin B_{12}, the enterabactin and the ferrichrome siderophores, respectively[87,88].

OMPs of known structure contain between 8-26 membrane-embedded β-strands, with almost all containing an even number of anti-parallel β-strands in their native state[89]. One exception is the mitochondrial OMP VDAC (Voltage-dependent anion channel), whose structure indicates a 19-stranded β-barrel[90-92]. The amphipathic β-strands within OMPs expose hydrophilic side chains to the inside of the channel (for those OMPs with an internal lumen), while those in contact with the bilayer are largely hydrophobic[89,93]. Two striking structural features of natively folded OMPs are likely to be important in their biogenesis: (1) both the N- and C-termini are located in the periplasm[94]; and (2) they have characteristically long loops on their extracellular side, while on their periplasmic sides the turns are short, consisting of only a few amino acids[89,94]. This latter feature may be important for the formation of intermediate β-hairpins; AFM experiments have shown that OMPs are able to unfold[95,96] and refold[97] via β-hairpin units.

While this project focuses on OMPs consisting of a single polypeptide chain, much larger β-barrel structures are formed in the outer membrane (OM) by the assembly of multiple monomeric subunits. These assemblies are usually involved in secretion, examples of which include CsgG, which is involved in curli biogenesis[98] forming a 36-stranded barrel in the OM from 9 subunits[99], and the dodecameric GspD[100-102], a member of the Type II secretion system (T2SS) family[103]. Available evidence suggests that these complexes assemble in pathways independent of the BAM complex[104,105], and are therefore not considered further in this work.

Another widely distributed class of proteins which form β-barrels in the OM are members of the autotransporter family[106,107]. Autotransporters consist of a large (often >100 kDa[108]) N-terminal 'passenger' domain, which is secreted into the extracellular space, and a 12-stranded β-barrel domain[107]. As crystal structures show that the diameter of this barrel (~10 Å) is too small to transport some of the folded passenger domains that autotransporters are able to secrete (10-20 kDa)[108], the name 'autotransporter' is perhaps misleading[109]. Indeed, assembly of autotransporters has been shown to involve the BAM complex[110-112], and/or the translocation and assembly module (TAM) complex[113] (Section 1.4) via mechanism(s) which remain unresolved[108].

1.2.2 The OMP biogenesis pathway

In common with other proteins with a high β-sheet propensity, OMPs are aggregation-prone so must be chaperoned at every stage of their biogenesis[14,81,114]. Following synthesis on cytosolic ribosomes OMPs are initially bound by Trigger Factor (TF) which prevents their interaction with the SRP as occurs for IMPs[115-117]. Binding of IMPs to SRP leads to their subsequent cotranslational membrane insertion by the SecYEG translocon using a lateral gating mechanism. IMPs have a highly hydrophobic signal anchor sequence at their N-terminus, whereas the N-terminal signal sequence (~20-30 residues) of proteins which are destined for secretion is less hydrophobic, and it is these differences in targeting signal characteristics which determine the pathway followed[41]. OMPs are escorted to the IM by the SecB chaperone[118,119] which binds to and transfers substrates to the SecYEG translocon via the SecA ATPase[120,121]. OMPs are then translocated across the IM by the SecYEG translocon using energy provided by the SecA ATPase (Figure 1.3)[122]. Once through the IM their leader sequence is cleaved by signal peptidase I (SPase I) and the mature OMPs are bound by periplasmic chaperones[13], such as SurA, Skp and DegP (Section 1.3). The time taken from synthesis to final folded state is in the order of minutes, with the lifetime of unfolded LamB in the periplasm reported to be ~2 min[123,124]. Recent kinetic modelling of the flux of unfolded OMPs across the periplasm suggests that, on average, OMPs make 100's of interactions with chaperones, which are present in concentrations such that there is always a free reservoir available for OMP binding[124].

OMPs finally reach the BAM complex, which catalyses the folding and insertion of β-barrel proteins into the OM[15,125,126]. Exactly how OMPs are targeted to the BAM complex is not known, but it has been proposed that, in addition to the recognition of non-specific β-sheet sequences, a C-terminal ~9 residue targeting motif is involved (Figure 1.5). Studies on the phosphate transporter PhoE demonstrated that a conserved phenylalanine residue at its C-terminus is essential for its efficient folding[127,128]. Robert and coworkers showed that the C-terminal motif is recognised by BamA in a species specific manner using *in vitro* binding experiments between BamA and unfolded OMPs or synthetic peptides derived from them[129]. The same study demonstrated that altering the C-terminal motif of neisserial PorA (a lysine to glutamine substitution at the -2 position) dramatically improved its assembly *in vivo* in the OM of *E. coli*[129]. However, a bioinformatics analysis of C-terminal targeting signals of all predicted OMPs from 437 organisms found little difference between

recognition motifs, such that OMPs from almost all Gram-negative bacteria should be able to be expressed in *E. coli*[130]. The targeting sequence is also important for the initiation of the σ^E-dependent stress response, which is triggered by the presence of misfolded OMPs in the periplasm[131], as peptides of the OMP signal sequence alone are able to trigger the σ^E response[132].

	10	9	8	7	6	5	4	3	2	1
PhoE	I	V	A	V	G	M	T	Y	Q	F
OmpF	T	V	A	V	G	I	V	Y	Q	F
PagP	V	Y	F	A	W	M	R	F	Q	F
OmpC	I	V	A	L	G	L	V	Y	Q	F
FadL	L	F	G	T	N	F	N	Y	A	F
OmpT	F	I	T	A	G	L	K	Y	T	F
BtuB	E	Y	T	L	S	G	S	Y	T	F
OmpLA	G	V	G	V	M	L	N	D	L	F
LamB	T	F	G	A	Q	M	E	I	W	W
BglH	V	G	G	A	Q	A	E	I	W	W

Figure 1.5: Sequence alignment of the C-terminal residues of ten OMPs from *E. coli*. The conserved residues indicated have been proposed to form a targeting sequence for BAM-mediated assembly of OMPs. The conserved sequence is ΦxΦxΦxYxF/W, where Φ is a hydrophobic residue and x is any amino acid. Conserved hydrophobic residues are highlighted in yellow; the consensus residues Y at position -3 and F/W at position -1 are highlighted in purple and blue, respectively. Numbers in green indicate distance from the C-terminus of the protein.

A targeting signal known as the β-signal is important for assembly of mitochondrial OMPs which has the consensus sequence ζxGxxΦxΦ, where ζ is a polar residue, Φ is a hydrophobic residue, and x is any amino acid[133]. This sequence is also present in some bacterial OMPs, such as BamA and OmpF, and at least some Gram-negative OMPs can be assembled into the mitochondrial OM[134] and *vice versa*[135]. Interestingly, a peptide which contained the β-signal sequence found in *E. coli* BamA was recently shown to inhibit BAM function *in vitro*, and cause growth defects and sensitivity to antibiotics *in vivo* when expressed in the periplasm[136].

1.3 Folding in the periplasm

A complex array of folding factors are involved in the biogenesis, homeostasis and protection of proteins in the periplasm including chaperones, proteases, peptidyl-proline isomerases (PPIases), and protein disulphide bond forming enzymes and isomerases[10]. Gram-negative bacteria are exposed to many environmental stresses, such as detergents, and extremes of pH and temperature. In such conditions many folding factors of the cell envelope are upregulated via one of five main stress responses (σ^e, Cpx, Rcs, Bae, and Psp)[18]. This is a large and fascinating area of study, however, the discussion here will focus on the folding factors for which experimental results have been obtained in the current project: the major OMP chaperones in *E. coli*, Skp and SurA.

1.3.1 Skp

The cellular location and role of Skp (Seventeen kilodalton protein) was initially subject to much controversy. Skp was reported to be a cell surface LPS-binding protein[137], a tetrameric DNA binding protein[138], an outer membrane protein[139] and a ribosome-binding protein[140]. Finally, Skp was demonstrated to be a soluble periplasmic protein[141] (consistent with the presence of an N-terminal cleavage signal sequence in its nucleotide sequence[142]), which selectively binds to OMPs[143] and is involved in maintaining OMP solubility in the periplasm[144]. Skp is a highly basic molecule (pI ~9.5) and its non-specific affinity for anionic molecules such as LPS, DNA and ribosomes most likely led to these initial artefactual results.

Several groups have shown that the Skp gene is not essential in *E. coli*[143-145]. An early study found that chromosomal inactivation of Skp led to reduced OM levels of the major porins, which could be reversed by the addition of a plasmid expressing wild-type Skp[143]. However, later gene deletion studies found little effect on OM composition[144,145], which was confirmed by a proteomics study which found that removal of Skp had no impact on the levels of 63 identified OMPs[146]. This indicates some redundancy in chaperone pathways, with other OMP chaperones, such as SurA and the chaperone/protease DegP able to functionally replace Skp[147]. Consistent with this, removal of additional chaperones in a Δ*skp* background have synergistic effects. One study found that Δ*skp*Δ*degP* cells accumulate protein aggegrates in the periplasm, have increased sensitivity to the hydrophobic anitbiotics rifampicin, vancomycin and novobiocin, and are non-viable at 37 °C[144].

Further, it was reported that a Skp/SurA double deletion is lethal[145], consistent with proteomic analysis of OMP levels in Δ*skp* cells in which SurA was depleted, which showed a reduction of nearly all OMPs[146].

Cross-linking studies in spheroplasts indicated that Skp operates early in the pathway, while OMPs are emerging from the SecYEG translocon[144,148]. By contrast, attempts to cross-link Skp to BamA *in vivo* have been unsuccessful, even in the absence of SurA[145].

The crystal structure of Skp revealed a trimeric structure consisting of a 'body' domain, containing a 9-stranded β-barrel which mediates trimerisation, with long coiled-coil α-helical 'tentacles' which extend away from the 'body' domain[149,150]. Skp has a very large electrical dipole (estimated at ~3,700 Debye[149]), with the region around the subunit tips particularly enriched in conserved lysine and arginine residues (Figure 1.6).

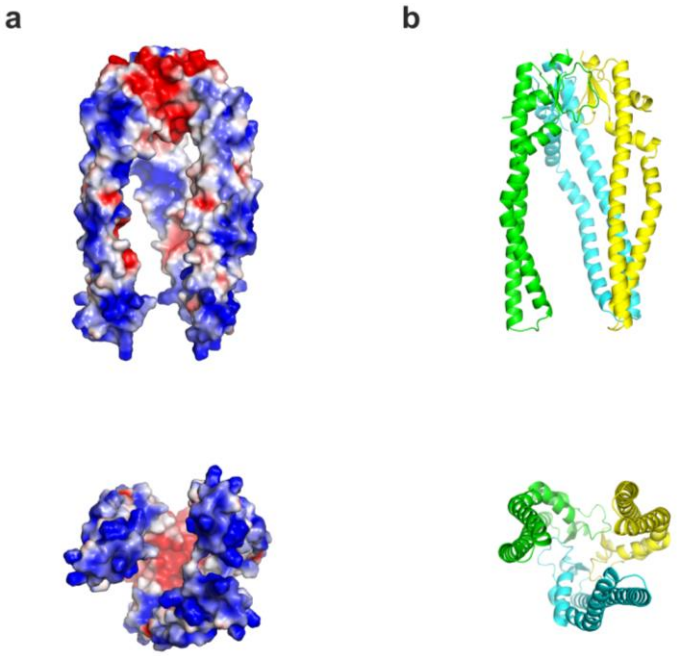


Figure 1.6: Electrostatic surface potential of the Skp chaperone. (a) Surface electrostatics of Skp (PDB: 1U2M[150]) viewed from the side (upper) and bottom (lower). Electropositive and electronegative surfaces are highlighted in blue and red, respectively. Electrostatic surface representations (-2 kT/e to +2 kT/e) were generated using the APBS plugin for PyMOL[151]. (b) Secondary structure of Skp, also viewed from the side (upper) and bottom (lower) (PDB: 1U2M[150]). Individual subunits are shown in yellow, green, and cyan. Missing residues in chains B and C were modelled from chain A using PyMOL.

The functional significance of this charge distribution may be that Skp acts as an electrostatic delivery vehicle[152], with positive residues in the tips interacting with either BAM components (all of which have a pI <7.0), or with the inner leaflet of the outer membrane which has an overall net negative charge due to the presence of the negatively charged lipids PG (~10%[39]) and CL (~3%[39]). Supporting the latter hypothesis, *in vitro* experiments have shown that the presence of Skp can increase the kinetics of folding into negatively charged synthetic liposomes for both OmpA[153] and PagP[152]. Further, a recent study suggested that Skp can fold and insert OMPs directly into native membranes *in vivo* in the absence of the BAM complex[154]. No direct *in vivo* cross-links to BAM components have been reported for Skp[145], however, site-specific photocrosslinking experiments showed that the β-barrel domain of EspP (β-domain) interacts sequentially with Skp and BAM components (BamA, BamB and BamD)[112].

Using an *in vivo* pull-down assay, Jarchow *et al.* found that Skp interacts with >30 OMPs, including a broad range of sizes from 8-stranded OMPs (OmpA, OmpX) to the 26-stranded LptD[155]. A Skp mutant which had a more hydrophobic interior was able to pull-down additional proteins to those pulled-down by wild-type Skp, providing evidence for client binding within the Skp cavity[155]. For one of the OMPs identified, OmpC, light-scattering assays have subsequently shown that Skp can prevent its aggregation *in vitro*[156]. The study by Jarchow *et al.* also demonstrated that, in addition to binding numerous OMPs, Skp also interacts with many soluble periplasmic proteins *in vivo*, including β-lactamase and maltose binding protein (MBP)[155]. The ability of Skp to effectively chaperone soluble proteins has been demonstrated by *in vitro* aggregation assays using lysozyme as the substrate[150,157], and co-expression of Skp has been shown to improve the yield of recombinant soluble proteins[158], including mammalian transcription factors[159] and antibody fragments[157]. *In vivo* cross-linking studies showed that Skp interacts with the soluble passenger domain of the autotransporter EspP[111].

While no specific Skp-recognition sequence(s) within OMPs have been identified, hydrophobic and electrostatic interactions have both been implicated in Skp binding[155,160,161]. Skp has been shown to have high affinity for its substrates with affinities measured by changes in tryptophan fluorescence (Skp binding to tOmpA (the transmembrane β-barrel domain of OmpA), NalP, OmpG, OmpA, and BamA)[160] and by Förster resonance energy transfer (FRET) experiments (Skp binding to OmpC)[162] in the nanomolar range. Qu *et al*. found that addition of 1 M NaCl reduced the apparent affinity of Skp for substrates, suggesting that electrostatics maybe important for Skp-substrate interactions. This effect was not due to effects on Skp trimerisation as sedimentation equilibrium analytical ultracentrifuge (AUC) experiments revealed Skp trimerisation is unaffected by salt[163]. However, salt has been shown to promote the aggregation of unfolded OMPs[164], so the apparent decrease in affinity of Skp for OMPs could be due to depletion of the unfolded OMPs due to aggregation[163].

A putative LPS-binding site was identified when the structure of Skp was first reported[150], based on the similarity in position and type between four residues in the Skp structure to those bound to LPS in the FhuA crystal structure[165]. The presence of LPS increases the observed kinetic folding rates of OmpA in the presence of Skp[153,166]. However, an extensive bioinformatics search found that some residues in the proposed binding site are not well conserved[167]. Further, nuclear magnetic

resonance (NMR) studies revealed that LPS acts in a similar manner to the detergent lauryldimethylamine-N-oxide (LDAO), in that it destabilises and denaturates Skp into monomeric subunits[167]. Therefore, it is likely that this non-specific denatuting effect of LPS on the Skp trimer, together with non-specific electrostatic effects between Skp and membrane-incorporated LPS[152,153], provides the explanation for the previous kinetic data[167].

NMR studies are ideally suited to the study of dynamic protein interactions such as those between chaperones and their clients[168,169]. NMR has shown that Skp maintains substrates in an unfolded state[170,171], with the high affinity between Skp and its substrates resulting from many weak and transient interactions[171]. Available evidence from NMR[161,170,171], *in vivo* pull-down experiments[155], *in vitro* cross-linking[170], and SANS[172], suggests that substrates are bound by Skp within its hydrophobic cavity. A selective isotope labelling strategy allowed the detection of intermolecular nuclear Overhauser effects (NOEs) between Skp and tOmpA[161]. The resulting residue-specific map of the Skp-tOmpA interface showed that tOmpA is bound in the centre of the Skp cavity[161]. The mass spectrometry results presented here in Chapter 4 also support the hypothesis that Skp sequesters its clients[173]. Skp bears a striking structural similarity to the eukaryotic chaperone prefoldin[174] (Figure 4.9f), which has also been shown to sequester substrates within its cavity[175].

For OMP substrates which are attached to a soluble domain, 2D NMR experiments have demonstrated that Skp can be bound to the transmembrane domain while allowing the periplasmic domain to fold independently[170]. $^{1}H^{15}N$ transverse-relaxation optimised spectroscopy (TROSY) heteronuclear single quantum coherence (HSQC) spectra were acquired for ^{15}N labelled constructs of OmpA and tOmpA (the transmembrane domain of OmpA) in complex with Skp, and the OmpA periplasmic domain alone. The tOmpA-Skp spectrum exhibited a narrow dispersion of peaks of varying intensities characteristic of an unfolded protein, while the spectrum for OmpA in complex with Skp contained additional well-dispersed peaks indicative of a folded protein. Many of these additional peaks overlapped with those from a $^{1}H^{15}N$ TROSY-HSQC spectrum for the periplasmic domain alone, showing that the periplasmic domain is folded outside the Skp cavity[170]. This was also confirmed by NMR experiments in which T_1 and T_2 relaxation parameters were obtained for residues in the OmpA periplasmic domain. The results showed that the average rotational correlation time (τ_c) was similar when the OmpA periplasmic

domain was free in solution to when it was covalently attached to the transmembrane domain in complex with Skp[170].

The potential of Skp as an antibiotic target may be limited as it is not an essential gene, at least in *E. coli* and *N. gonorrhoea*[176]. However, it does interact with the multidrug efflux pump TolC[80], which spans the periplasm and is known to mediate drug resistance in many pathogens, including *Vibrio cholerae*[177]. Further, genetic evidence has also shown that Skp has a specific role in the assembly of the essential OMP LptD which mediates LPS insertion into the OM[178]. Investigations into the effects of Skp deletion from *N. meningitides* found that levels of the porin PorB were affected, which could not be complemented by the addition of Skp from *E. coli*[179]. However, deletion of *N. meningitides* SurA had no effect on OMP biogenesis[179], in sharp contrast to the results obtained for *E. coli* SurA[145,180] (Section 1.3.2). This highlights the species-specificity of OMP chaperones, a factor which will need to be considered if drugs targeting OMP chaperones are to be developed.

1.3.2 SurA

SurA (Survival protein A) was first reported as a gene required for bacterial survival in the stationary phase[181,182]. Later SurA was identified as a member of the parvulin-like family of PPIases[183], which assists the correct folding of porins in the OM[183,184]. This was confirmed by genetic studies which found that depletion of SurA resulted in reduced OM density, in contrast to depletion of either Skp or DegP which had no effect[145]. Simultaneous removal of SurA and Skp is lethal[147,185], with similar synthetic lethality observed for the double deletion of SurA and DegP[147]. These results led to the suggestion that Skp and SurA function in partially redundant parallel pathways, with SurA as the main chaperone pathway for OMPs and the Skp/DegP pathway becoming more important in times of stress[145,147]. However, differential proteomic analysis of 23 OMPs from wild-type and *ΔsurA* strains found that only a subset of eight OMPs were substantially reduced in the absence of SurA[180]. These include the porins OmpA, OmpF and LamB, which make up the bulk of the protein mass in the OM, as well as FadL, OmpX, FecA, FhuA and LptD. RT-PCR analysis of mRNA levels in *ΔsurA* cells showed that FhuA and LptD are the only OMPs for which a reduction in protein expression levels could not be explained by reduced transcription of their genes due to σ^E induction, suggesting these are true SurA substrates[180]. Interestingly, a later proteomics study showed that BamA assembly, was not dependent on SurA or Skp[146].

SurA was found to be associated with the OM[186] and, by contrast to other periplasmic chaperones, has been directly cross-linked to the BAM complex *in vivo*[145,187]. An arginine residue (R64) in polypeptide transport-associated (POTRA) domain 1 of BamA is important for this interaction[187], and the P2 domain of SurA has been implicated in this interaction by *in vivo* photocross-linking[188].

SurA (45 kDa) consists of an N-terminal domain and a small C-terminal domain, interspersed with two PPIase domains (P1 and P2) (Figure 1.7). The SurA crystal structure showed an elongated structure[189] reminiscent of Trigger Factor, which also has a PPIase domain at one end[190].

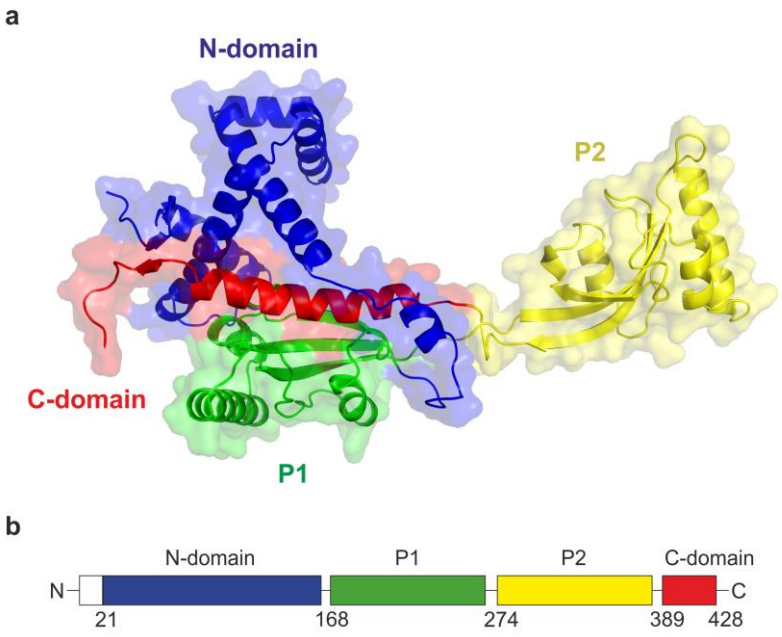


Figure 1.7: Crystal structure of SurA. (a) SurA consists of N- and C-terminal domains which form the core chaperone unit, interspaced by two PPIase domains. Image created in PyMOL (PDB: 1M5Y[189]), with missing residues added using MODELLER[191]. (b) Domain structure of SurA coloured as in (a).

Only the second PPIase domain (P2) of SurA exhibits significant PPIase activity, and general chaperone activity was still present in a mutant which lacked both the P1 and P2 domains (SurA-NCt)[185]. Interestingly, protection against aggregation of thermally-denatured citric synthase was enhanced for SurA-NCt compared with wild-type SurA[185]. Expression of the PPIase domains *in vivo* could not complement *ΔsurA* cells, by contrast to SurA-NCt which almost completely complements wild-

type SurA, suggesting that the main role of SurA is as a chaperone[185]. Consistent with this, SurA homologues in other species lack the P2 domain (e.g. *Haemophilus*)[185] or both P1 and P2 domains (e.g. *Rickettsia*)[192].

Xu *et al.* (2007) solved the structures of peptides in complex with the P1 domain of SurA and full-length SurA lacking P2 (SurA-ΔP2)[193]. Surprisingly, the structures showed peptides bound to P1, and in the case of SurA-ΔP2, a dimeric structure was solved involving a large-scale structural rearrangement. This suggests that SurA can adopt very different tertiary structures when binding substrates, and the authors propose that P1 may be involved in determining substrate specificity[193]. Recent genetic evidence highlighted the involvement of P1 in SurA function in *E. coli*; SurA mutations which suppress OMP assembly defects caused by BamA and BamB mutations, were found to map to the P1 domain[194].

The SurA C-terminal ~10 amino acids contain a single β-strand, which forms a three-stranded β-sheet with a β-hairpin from the N-terminal domain (Figure 1.8)[189]. Deletions of these C-terminal residues have shown that formation of this β-sheet is essential for SurA chaperone activity *in vivo*[195]. However whether this indicates a functionally relevant location, for example providing an interaction site for OMP substrates via β-augmentation[196], or merely disrupts the overall fold of the protein is unclear[195].

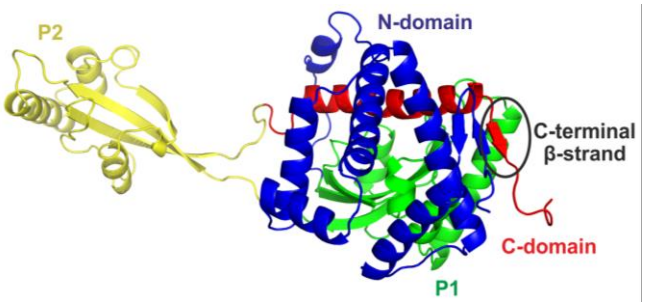


Figure 1.8: Structure of SurA highlighting the C-terminal β-strand essential for chaperone activity *in vivo*. Image created in PyMOL (PDB: 1M5Y[189]), with missing residues added using MODELLER[191].

SurA has been shown to bind unfolded OMPs[162,197] and have a specificity for substrates rich in the Ar-X-Ar motifs common in OMPs (where Ar is an aromatic residue and X is any amino acid)[198]. Measured affinities are in the low μM range (K_d)

for SurA binding to peptides from phage display libraries[186,197,198], and to the full-length OMPs OmpF and OmpG using a peptide competition assay[197]. Tighter binding (K_d of ~0.1 μM) was observed for SurA binding to the 16-stranded OmpC in fluorescence-based experiments[162].

It may transpire that rather than having one specific binding site, hydrophobic surfaces at multiple locations across SurA may interact with OMP substrates, as NMR studies have demonstrated for the chaperones Trigger Factor[199] and SecB[169] in complex with unfolded soluble proteins. A recent *in vivo* photocross-linking study showed that SurA interacts with OMPs in many regions across the surface of the N-terminal domain[188] (Figure 1.9). Specific regions within OMP substrates may also preferentially bind to SurA; recent *in vivo* photocross-linking of SurA to OMPs with genetically incorporated p-benzoyl-L-phenylalanine (Bpa)[200] (OmpF or LamB with Bpa at 17 or 19 different positions, respectively) showed that these OMPs interact with SurA preferentially at their N- and C-terminal regions[188]. Consistent with this, the Ar-X-Ar motif shown to be recognised by SurA[198] is very common in OMP sequences at their C-terminus (Figure 1.5)[130].

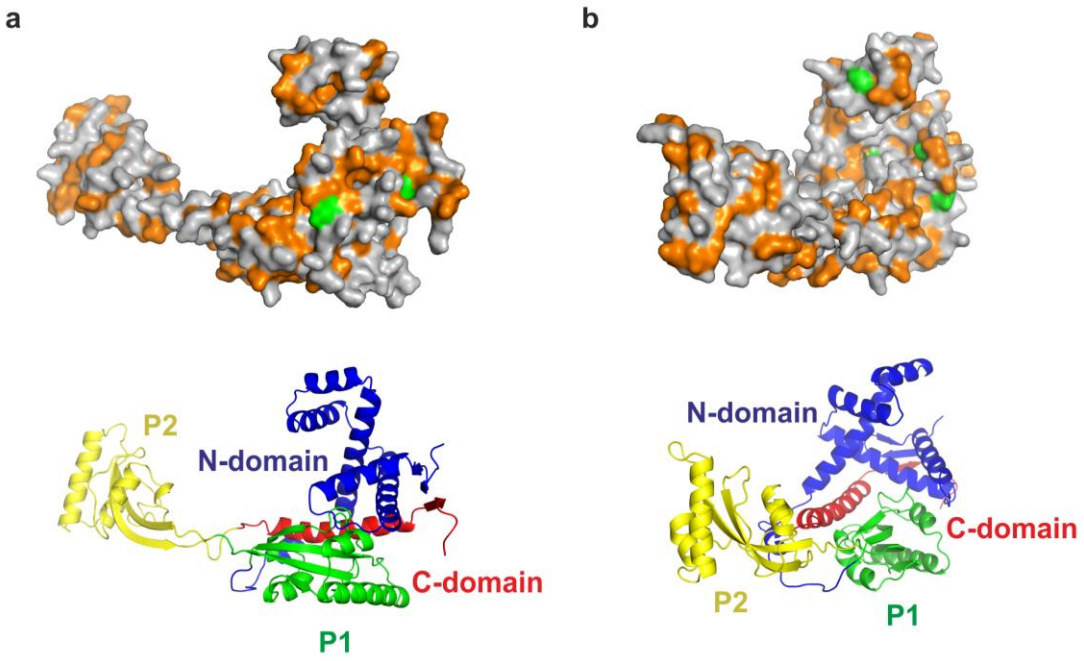


Figure 1.9: SurA has hydrophobic patches across the whole of its structure and *in vivo* cross-links map to the N-terminal domain. Views of SurA from (a) the side, and (b) the front – P2 facing out of the page). Hydrophobic residues (Gly, Ala, Val, Leu, Ile, Met, Phe and Trp) are shown in orange. Residues which, when replaced with Bpa cross-link to OMPs *in vivo*, map to different regions within the N-terminal domain[188] are highlighted in green (upper). Domain organisation is shown in cartoon representation (lower). Image created in PyMOL (PDB: 1M5Y[189]), with missing residues added using MODELLER[191].

It is also highly likely, given the size of SurA (45 kDa) relative to some of its known substrates (e.g. FhuA (79 kDa) and LptD (87 kDa)), that its chaperone activity involves multivalent binding. Indeed, recent analytical size exclusion chromatography (SEC) experiments for SurA in complex with FhuA were consistent with a 2:1 SurA:FhuA stoichiometry[97], and *in vivo* cross-linked products between SurA and OmpF suggested that 2-3 copies SurA may bind OmpF[188].

In vitro single molecule force spectroscopy (SMFS) experiments recently demonstrated that the presence of SurA allowed the refolding of FhuA in β-haipin units in a stepwise manner following mechanically induced unfolding[97]. FhuA was mostly misfolded in the absence of SurA, while FhuA remained unfolded in the presence of Skp, leading to the suggestion that SurA specifically biased the conformations of dynamic, unfolded states towards those which favour productive folding[97]. There is much left to discover about SurA, but its importance is highlighted by the link between SurA and virulence of a number of pathogens[201,202].

1.3.3 DegP

DegP (also known as HtrA[203]) is a dual function serine protease/chaperone which is upregulated in response to the accumulation of unfolded OMPs in the periplasm, and by envelope stresses such as heat shock, via both the σ^E and Cpx stress response systems[10]. The switch between DegP protease and chaperone function is governed by temperature, with protease activity dominating at high temperatures (42 ºC) and DegP functioning mostly as a chaperone below 28 ºC[204].

Structurally DegP (47 kDa) consists of three domains, a chymotrypsin-like protease domain and two PDZ domains involved in substrate binding[10]. The resting state is a hexamer which can undergo oligimerisation to form large 12-mer and 24-mer cages in the presence of substrate[205] (Figure 1.10). However, DegP protease activity is not dependent on the formation of these higher order structures[206]. A cryo-EM structure of the DegP 12-mer contained density which, surprisingly, was consistent with a folded monomer of the porin OmpC. This led the authors to speculate that, as well as forming a proteolytic chamber, DegP cages could provide a folding compartment for nascent OMPs[205].

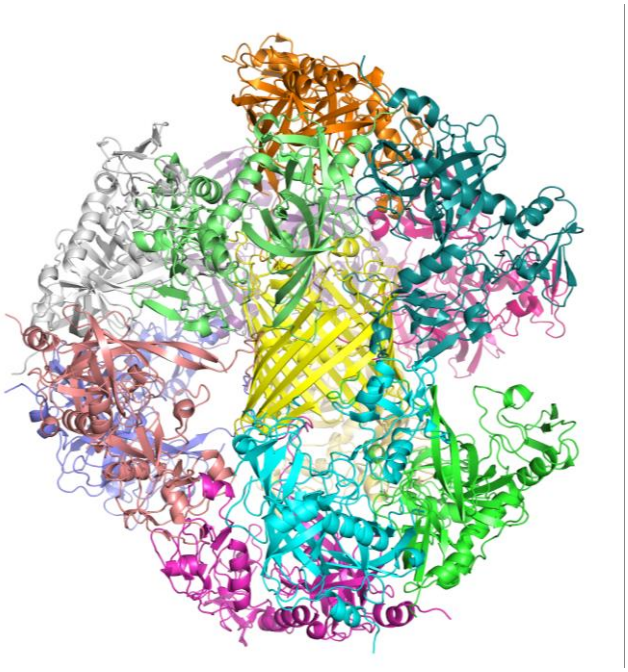


Figure 1.10: Model of a DegP 12-mer based on a cryo-EM reconstruction. Structure of a DegP 12-mer with folded 16-stranded porin (yellow) modelled in to account for the density in the centre of the cage. Image created with PyMOL (PDB: 2ZLE[205]).

Studies using fluorescently-labelled chaperones and OMPs showed that DegP can form ternary complexes with substrates and either Skp or SurA[162]. While DegP-OMP complexes are thermodynamically more stable than those between OMPs and either Skp or SurA, the association rate (k_{on}) between OMPs and DegP is ~1000-fold slower, suggesting a kinetic partitioning model for OMP transport across the periplasm[162]. This work has led to a proposal that DegP may operate late in the OMP biogenesis pathway[207], however, attempts to cross-link DegP to the BAM complex have been unsuccessful[145].

1.3.4 FkpA

FkpA (26 kDa), like SurA, is a dual function PPIase and chaperone implicated in OMP biogenesis. As for SurA, *in vivo* and *vitro* studies have indicated that these activities are independent from one another[10]. Deletion of FkpA is not lethal, but in one study led to an upregulation of periplasmic proteases and an increase in OM permeability[208]. FkpA appears to become important in OMP assembly at heat shock temperatures (44 ºC)[209]. The synthetic lethal phenotype of a *Δskp ΔsurA* strain could be suppressed at high temperatures by over-expression of FkpA. Further, *in vitro*

work showed FkpA is more effective at preventing the aggregation of OmpF at 44 ºC than at 37 ºC, and fluorescence-based experiments showed that both the association rate and affinity between FkpA and OmpF is increased at elevated temperatures[209].

Structurally FkpA consists of an N-terminal helical domain involved in chaperone activity and a C-terminal PPIase domain (Figure 1.11). The crystal structure showed a v-shaped dimer which suggested substrates may be cradled in the crevice between monomers[210], with this 'mother's arms' chaperone model later supported NMR studies[211].

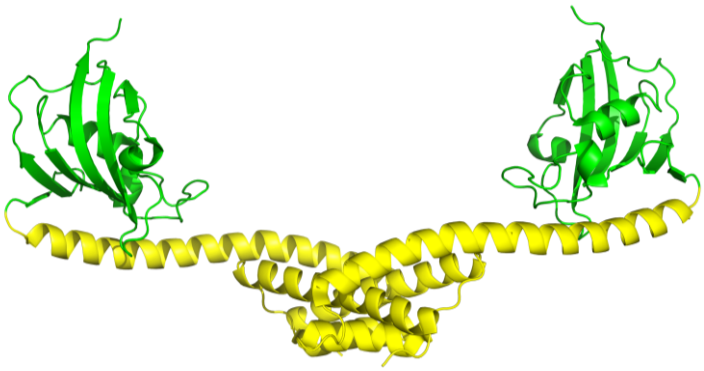


Figure 1.11: Crystal structure of the FkpA dimer. The N-terminal chaperone domain and C-terminal PPIase domain are shown in yellow and green, respectively. Image created with PyMOL (PDB: 1Q6U[210]).

Schwalm *et al*. (2013) showed that the assembly of LptD is compromised in *ΔfkpAΔskp* strains, providing further evidence that FkpA is involved in OMP biogenesis[178]. Interestingly, the defect in LptD assembly caused by loss of FkpA and Skp, could be suppressed by a mutation which leads to a 1000-fold increase in the expression level of the general chaperone Spy[178,212]. However, while this indicates Spy is able to chaperone OMPs, its upregulation was unable to complement for loss of SurA[178].

1.3.5 PPIases and disulphide isomerases

In addition to SurA and FkpA two other proteins in the periplasm are known to have PPIase activity: PpiA and PpiD[68]. PpiA is a highly active PPIase *in vitro*[213], but its

deletion had no effect on total OMP levels or their rate of assembly *in vivo*[214]. There is conflicting evidence on the role of PpiD in OMP folding. PpiD was first reported as a multicopy suppressor of SurA[215]. Deletion of the gene was reported to impair OMP folding and *surA* and *ppiD* double null mutants were found to be synthetically lethal[215]. However, these results have not been replicated and the Silhavy group later reported that PpiD deletion has no effect on OMP biogenesis[216]. Interestingly even quadruple *ΔppiDΔppiAΔfkpAΔsurA* null mutants were found to be phenotypically indistinguishable from a *ΔsurA* strain[216]. However, *in vivo* PpiD has been shown to have partially overlapping substrate specificity with SurA[217].

A recent *in vivo* cross-linking study found simultaneous cross-links between PpiD, SecYEG, SurA, and BAM complex members, leading the authors to speculate that PpiD is involved in a supercomplex that spans the periplasm[188]. However, if this were a major pathway for OMPs across the periplasm, OMP production in *ΔppiD* mutants should be severely comprised[216]. Further, kinetic modelling of OMP flux through the periplasm suggests that such a supercomplex is not needed to explain all known experimental data[124].

The correct formation of disulphide bonds in proteins can be crucial to their stability, and, together with the *cis-trans* isomerisation of peptidyl-prolyl bonds, is often the rate limiting step in protein folding. The disulphide isomerases DsbA and DsbB are involved in the catalysis of disulphide bond formation, with DsbC and DsbD operating later as 'proof-readers'. Their involvement in OMP biogenesis is limited as only a small subset of OMPs contain disulphide bonds[218]. LptD is a notable example, for which both DsbA and DsbC have been reported to assist in the correct ordering of disulphide bonds in the native state[219,220].

1.4 The BAM complex

The structure and function of the BAM complex has been the focus of intense research interest in the 13 years since the discovery of the role of BamA in OMP assesmbly[15]. BAM is a hetero-oligomeric complex which, in *E. coli*, consists of five subunits: BamA-E, formerly known as YaeT, YfgL, NlpB, YfiO and SmpA, respectively (Figure 1.12)[9,125,126]. This section will consider what is known about the roles of each subunit, and discuss the complete[221,222] and near-complete[221,223] BAM structures recently available, and their implications for BAM mechanism.

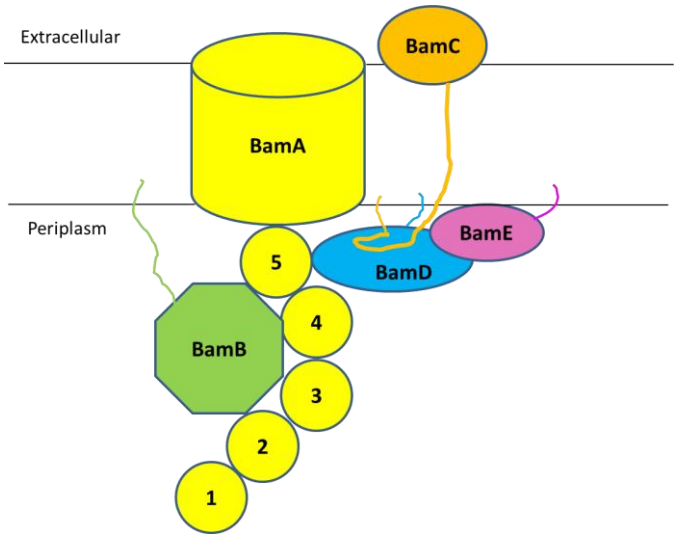


Figure 1.12: Schematic of the *E. coli* BAM complex. The major subunit, BamA, consists of a C-terminal membrane-embedded β-barrel domain and five N-terminal POTRA domains. The four lipoproteins BamB-E are scaffolded by the BamA POTRA domains. BamB binds BamA separately to BamCDE.

Most information on the BAM complex has come from studies in *E. coli*[9], but the subunits are widely conserved across Gram-negative species[224]. The largest subunit, BamA, is an integral membrane protein, and the only component present in all Gram-negative bacteria[224]. Its role in OMP assembly was first identified in *Neisseria meningitides*; depletion of the neisserial BamA homologue Omp85 led to growth and OMP assembly defects, and its deletion was lethal[15]. BamA homologues exist in mitochondria and chloroplasts[225-227], reflecting the endosymbiotic origin of these organelles[227]. In mitochondria, OMPs are assembled by the Sorting and Assembly Machinery (SAM) complex, whose BamA homologue is Sam50. Sam50 associates with two essential cytosolic proteins Sam35 and Sam37 which are unrelated to BamB-E[228]. Much less is known about OMP assembly in plants; the BamA homologue in chloroplasts (Toc75-V) is involved[229,230], but its accessory proteins remain to be identified[13,231].

BamB-E are lipoproteins, anchored to the OM via their N-terminal acyl chains, with molecular weights of 42, 37, 28 and 12 kDa, respectively in *E. coli*. The ratio between the BAM subunits in structures from BAM isolated from the *E. coli* OM is 1:1:1:1:1[221,222,232], making a core complex of ~203 kDa. This ratio was also suggested for the BAM complex made by reconstitution of purified BamAB and BamCDE subcomplexes, although it was difficult to resolve whether one or two copies of BamE were associated with the complex due to its low molecular

weight[233]. Native mass spectrometry (MS) of OM-purified BAM complex indicated the presence of a small population of complexes with an additional BamE subunit[232], and BamE dimerization has also been observed *in vitro*[234-236].

Co-immunoprecipitation experiments in *E. coli* showed that any BAM complex member expressed with a His-tag can pull-down the entire complex[125,126]. BamB binds BamA separately from BamD, and BamC and BamE both depend on BamD for their stable interaction with BamA[237]. Consistent with the idea of a modular BAM complex, sub-complexes of BamAB and BamCD, in addition to the BamABCDE complex, have been observed in the OM of wild-type cells, suggesting BAM sub-complexes may dynamically associate *in vivo*[72].

The oligomeric state of the BAM complex *in vivo* is not known, however in one study purified *E. coli* BamA was seen to form possible tetramers as evidenced by a 473 kDa peak in a size exclusion chromatograph (SEC)[129]. Monomers, dimers and trimers of BamA, as well as tetramers were detected by Blue Native Polyacrylamide Gel Electrophoresis (BN-PAGE). The authors accounted for the differences in the SEC and BN-PAGE data by suggesting that Coomassie dye binding to BamA shifts the equilibrium towards the formation of smaller oligomers and monomers[129]. The Omp85 family member HMW1B from *Haemophilus influenzae* also forms tetramers[238], however no tetramerisation of the whole BAM complex has yet been observed. Recent evidence indicates that a single copy of the whole BAM complex, in nanodiscs containing *E. coli* polar lipid extract, is sufficient for the assembly of the autotransporter EspP[239]. However, spatial clustering of BAM complexes may be functionally relevant, as BAM has been observed in 0.5 μm 'OMP islands' *in vivo*[37]. Further, recent genetic experiements suggested multiple copies of BAM may be involved in the assembly of trimeric porins[240].

Genetic deletion of BamA[15,125] or BamD[241] is lethal, while deletion of BamB, BamC or BamE leads to varying degrees of OMP assembly and growth defects[13]. The molecular details of how the BAM complex aids OMP insertion and folding have yet to be elucidated. However, a wealth of structural and biochemical information has become available in recent years for each of the BAM components which is summarised below.

1.4.1 BamA

BamA is a member of the TpsB/Omp85 superfamily which contains members with functions associated with protein translocation across, and assembly into, Gram-negative bacterial outer membranes[242]. BamA is the most conserved BAM component, and the only one with homologues in all known Gram-negative species[224]; the thermophilic bacterial species *Thermus thermophilus* and *Thermosynechococcus elongates* contain no homologues of the lipoproteins BamB-E[243,244]. Proteins in the Omp85 family consist of a C-terminal membrane-inserted 16-stranded β-barrel domain with up to seven tandemly-linked POTRA domains. The number of POTRA domains varies between one and seven according to the function of the Omp85 family member[226,244,245], and also between species. The Omp85 protein Mxan5763 of *Myxococcus xanthus* has seven POTRA domains[246]. Mitochondrial Sam50 has only one POTRA domain, and one study surprisingly reported that the Sam50 POTRA domain is non-essential for β-barrel assembly and cell growth in yeast[133]. BamA from cyanobacteria, such as *T. elongates,* have three POTRA domains[244] while BamA from proteobacteria, such as *N. meningitides* and *E. coli,* have five (Figure 1.13)[68,247].

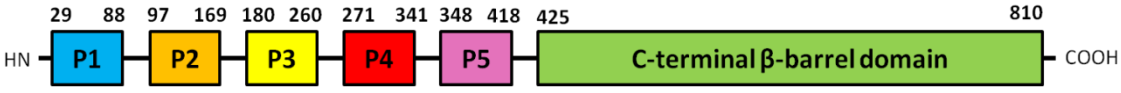


Figure 1.13: *E. coli* BamA domain organisation. BamA consists of five N-terminal tandemly-linked POTRA domains (P1-P5) of approximately 75 residues each, and a 385 residue C-terminal 16-stranded β-barrel domain.

The first and last POTRA domains in cyanobacteria (1 and 3) correspond, in terms of sequence and structure, to the first and last POTRA domains in proteobacteria (1 and 5)[244]. TamA, a BamA homologue involved in autotransporter assembly[113], has three POTRA domains[248]. The TAM complex has no other OM components, and instead is associated with an inner membrane protein, TamB[113,249]. The number of BamA POTRA domains that are essential appears to vary between bacterial species. In *E. coli*, it was initially observed by domain deletion experiments that only POTRAs 3-5 are essential[247]. However, a later study, using a different expression system, found that all five POTRA domains were required for normal growth and viability[250]. By contrast, in *N. meningitidis* it was found that only POTRA 5 (nearest

the C-terminal β-barrel) was required; only minor OMP defects resulted from deletion of POTRAs 1-4[251]. Replacement of BamA in *N. meningitides* and *E. coli* with BamA from other Gram-negative species was lethal, except in the case where species were closely related (*N. gonorrhoea* BamA could complement BamA from *N. meningitides*)[252]. The β-barrel domain of BamA was reported to be essential in *N. meningitides* as its substitution with an unrelated β-barrel was lethal[251]. However, in this study the OmpLA β-barrel was chosen. As OmpLA forms dimeric structures[253], the lack of viability of these mutants may due to the formation of dimers of POTRA domains-OmpLA barrel fusion proteins. Interestingly, it was found that chimeric *E. coli* BamA fusion proteins, carrying either the β-barrel or POTRA domains from various BamA orthologues, were viable[254]. While the BamA β-barrel domains were found to be functionally interchangeable between most species, compensatory mutations in the POTRA domains could improve OMP assembly, suggesting that the BamA membrane and soluble domains need to be precisely tuned for efficient OMP assembly[254].

BamA POTRA domains are proposed to have at least three roles: (1) to serve as docking sites for BamB-E, as evidenced by domain deletion experiments[247], and confirmed by recent structures[221-223]; (2) to interact with periplasmic OMP chaperones – POTRA 1 has been shown to interact with SurA[187]; and (3) to form a binding site for OMP substrates. Early evidence for BamA directly binding OMPs came from conductance experiments in planar lipid membranes which showed that BamA forms a pore whose conductance, and therefore presumably size, increases when bound to unfolded OMPs[255]. Robert *et al*. used this observation to demonstrate that the binding between OMPs and BamA is species specific[129]. NMR data have also demonstrated binding between β-strand sequences of PhoE and BamA POTRA domains[256]. However, this binding was very weak, as chemical shift changes in 50 μM POTRA domains 1-2 were not saturated on addition of up to 1.5 mM PhoE peptide[256]. There are no reports of POTRA domains binding directly to full-length OMP substrates *in vitro*.

In common with other Omp85 family members BamA POTRA domains have low sequence identity between each other[245]. However, the publication by the Kahne group in 2007 of a crystal structure of *E. coli* POTRA domains 1-4 (and a small section of POTRA 5) revealed domains which are highly superimposable, with a fold consisting of two α-helices and a 3-stranded β-sheet (β1-α1-α2-β2-β3) (Figure

1.14)[247]. A further crystal structure of POTRAs 1-4 by the Sousa group showed a structural difference to that of the Kahne group in the orientation of POTRAs 2 and 3[257]. In the Kahne structure the angle between the two domains is 30° ('bent' conformation), whereas in the Sousa structure the angle is 130° ('extended' conformation) (Figure 1.14).

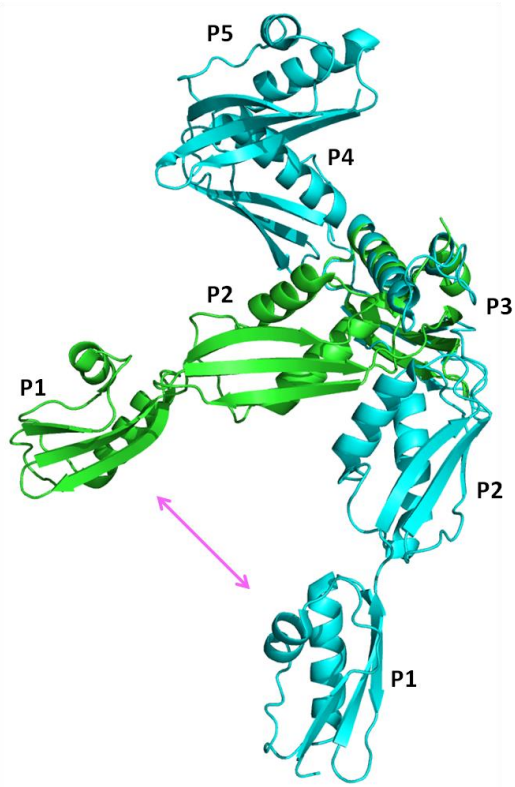


Figure 1.14: *E. coli* BamA POTRA domains in 'bent' and 'extended' conformations showing flexibility in the hinge between POTRAs 2 and 3. The two independent structures of POTRAs 1-4 are shown superimposed at POTRA 3 (Kahne structure in green (PDB: 2QCZ[247]), Sousa structure in blue (PDB: 3EFC[257]). The POTRA 5 domain has been modelled from a structure of POTRAs 4-5 (PDB: 3OG5[257]). The proposed functional movement of the POTRA domains, hinged at the loop between POTRAs 2 and 3[257], is indicated by a pink arrow. Structures aligned and images created with PyMOL.

NMR and SAXS data for POTRAs 1-5 confirmed that there is flexibility between these domains[256]. A subsequent crystal structure of POTRAs 4-5, published in 2010, revealed that the complete periplasmic domain of BamA forms a corkscrew-like structure[258]. The authors additionally performed Ensemble Optimisation Method (EOM) analysis on SAXS data for POTRAs 1-5, which suggested that the two crystal structures represent two preferred conformations of the complete periplasmic

domain, with some rigidity at all POTRA interfaces except POTRAs 2 and 3[258]. Solid state NMR studies of complete BamA also indicated flexibility within the POTRA domains[259]. Crystal structures of BamA POTRAs-BamB fusions demonstrated that the POTRA 2-3 hinge is the location bound by BamB[260,261], suggesting that BamB may be involved in the regulation of POTRA domain conformations[260]. Newly available X-ray and cryo-EM structures for full-length BamA[11], and BAM complexes[221-223,232] (Section 1.4.6) indicate flexibility between all POTRA domains, and between the POTRA domains and the BamA barrel. This is consistent with results from molecular dynamics (MD) simulations of full-length BamA embedded in a native OM[262]. However, the largest variation in the angles between domains in all structures containing BamA occurs at the POTRA 2-3 hinge[221-223,232,247,257].

A breakthrough came in 2013 when the first structures of BamA became available, from the organisms *Neisseria gonorrhoeae* and *Haemophilus ducreyi* (Figure 1.15)[11].

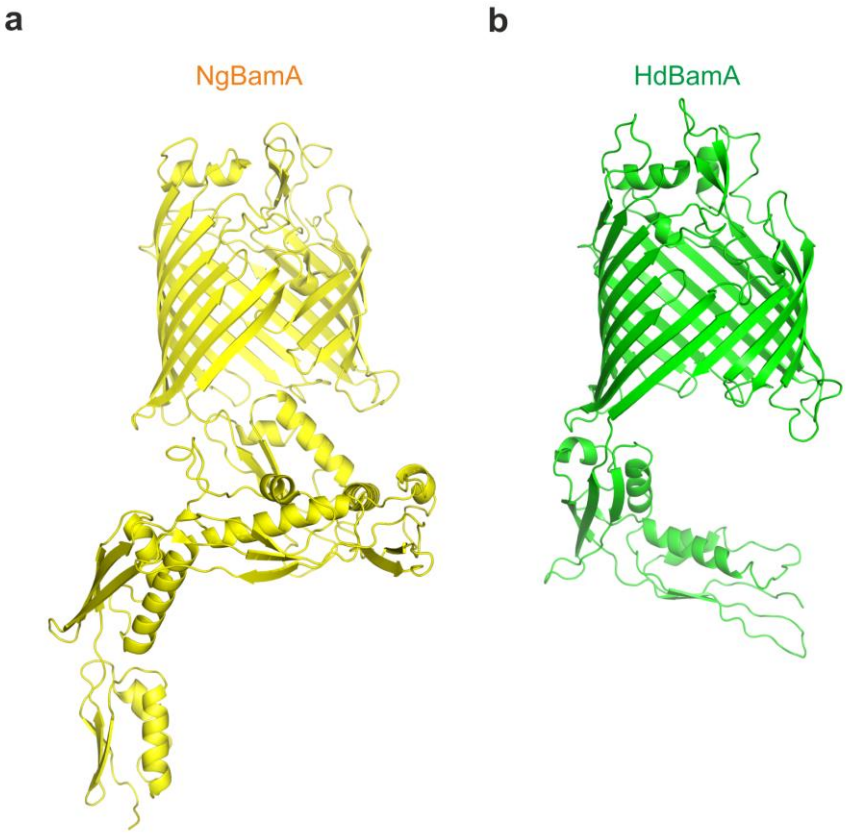


Figure 1.15: Crystal structures of BamA. BamA structure from (a) *N. gonorrhoea* (PDB: 4K3B[11]), and (b) *H. ducreyi* (PDB: 4K3C[11]).

The structures revealed three key features: (1) in the *N. gonorrhoeae* structure the lumen of the BamA β-barrel was occluded by POTRA 5, whereas in the *H. ducreyi* structure the cavity is accessible, suggesting a possible gating mechanism; (2) the side of the of the BamA β-barrel at the interface between β1 and β16 has a narrowing of the hydrophobic surface. This feature was proposed to be involved in local membrane thinning and increase in lipid disorder to aid OMP insertion, and was backed by evidence from MD simulations of the BamA barrel in DMPE bilayers[11]; and (3) in the *N. gonorrhoeae* structure only two hydrogen bonds are present between β1 and β16, with the C-terminus of β16 twisted into the barrel lumen, suggestive of a lateral gating mechanism. Lateral opening events at the β1/β16 seam were observed in MD simulations of the BamA barrel in DMPE bilayers at 340 K, which were not seen in control simulations of the BtuB and FhaC barrel domains[11]. Experiments in which cysteine mutations designed to cross-link the β1 and β16 strands were introduced into BamA *in vivo* were found to be lethal (Figure 1.16)[221,263].

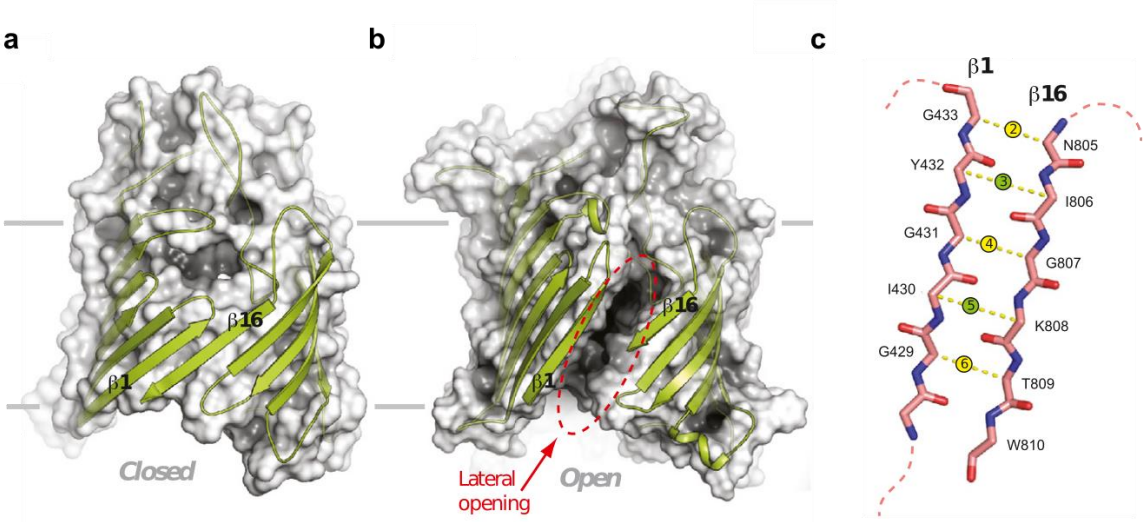


Figure 1.16: Lateral opening of the BamA barrel is essential *in vivo*. (a) BamA model in the 'closed' state, based on HdBamA crystal structure (PDB: 4K3C[11]). (b) BamA model in the 'open' state, from MD simulations of *E. coli* BamA. (c) Residues for which disulphide cross-links were introduced *in vivo*[263]. Yellow and green circles represent residues which face into the lumen of the barrel or into the membrane, respectively. All cross-link mutants were lethal in the absence of reductant, with the exception of the Y423C/I806C mutant which exhibited reduced colony formation. Image adapted from Noinaj *et al*. (2014)[263].

Recently, complementary experiments were performed *in vitro* which investigated the effect of cross-linking the BamA barrel of purified BAM complexes reconstituted in proteoliposomes of *E. coli* polar lipid extract[232]. Cross-linking of the BamA barrel substantially affected the folding of OmpT, which could be rescued to near wild-type levels in the presence of reductant[232]. Interestingly, BAM activity was not completely ablated in the oxidised state, suggesting that the lethal phenotype observed for BamA barrel cross-links *in vivo* may be due to kinetic effects on OMP insertion.

One key mechanistic problem in OMP assembly is how OMP extracellular loops and/or extracellular domains are translocated across the OM. MD simulations of the *E. coli* BamA barrel revealed an opening just above the lateral gate, between the loop connecting β1 and β2 (L1), and the L6 loop. This opening was proposed to function as a substrate exit pore[263]. Consistent with this, Cys mutants designed to prevent opening of the exit pore by disulphide cross-linking are lethal, but can be rescued by the presence of reductant (Figure 1.17)[221,263].

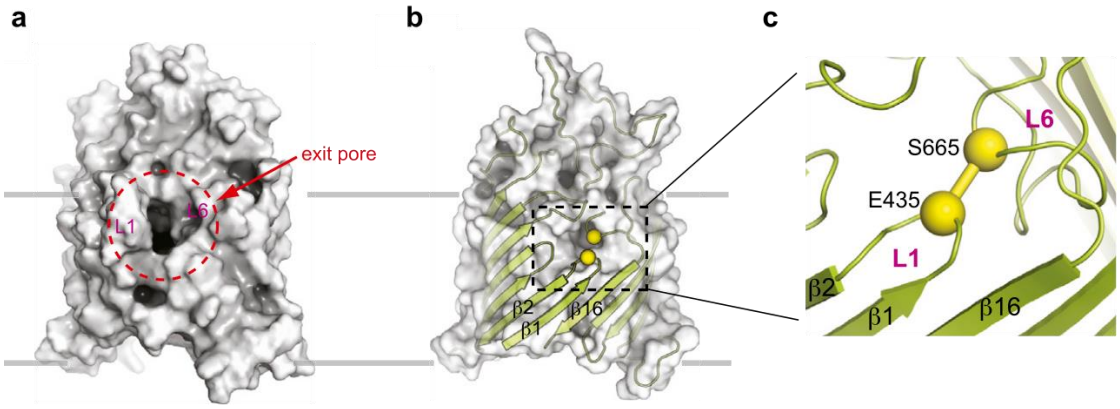


Figure 1.17: Putative substrate exit pore in the BamA β-barrel domain. (a) Opening of the proposed exit pore in MD simulations of the *E. coli* BamA barrel. (b) Model of the BamA barrel with the exit pore in the closed state. (c) Residues designed to form a cross-link between L1 and L6 which when mutated to Cys (E435 and S665) are shown as yellow spheres. The E435C/S665C mutant is lethal, and can be rescued by the presence of reductant[263]. Image adapted from Noinaj *et al*. (2014)[263].

1.4.2 BamB

Although deletion of BamB is non-lethal, strains lacking BamB have severe OMP assembly defects[125,264]. BamB deletions show the same phenotype as SurA mutants[123], and double deletion of these two proteins is lethal[264], suggesting that SurA and BamB operate in the same pathway[13]. The OMPs most affected in Δ*bamB* strains are large β-barrels (16-24 stranded[264-266]), leading to the suggestion that BamB may function by providing additional locations for OMP binding via β-augmentation for larger substrates[13,267]. However, not all larger barrels are equally affected in *bamB* null mutants; the levels of the 16-stranded OmpC are only slightly reduced in the absence of BamB[265]. Further, ITC experiments showed no evidence for binding of peptides derived from OMPs to BamB from *Pseudomonas aeruginosa*[268], and indeed no studies have reported an interaction between BamB and OMP substrates[269].

The crystal structures of BamB from *E. coli*[267,270-272], and *Pseudomonas aeruginosa*[268], reveal that it forms an eight-bladed β-propeller (Figure 1.18). BamB has a short, funnel-like shape with a height of 28 Å and a width of 48 Å, and an opening in the centre of ~10 Å at its narrowest point. This pore is a conserved structural feature in BamB not present in other β-propeller proteins. However, this

appears not to be functionally important as a BamB mutant designed to obstruct this pore (N62W/S329W) could fully complement a *ΔbamB* strain[268].

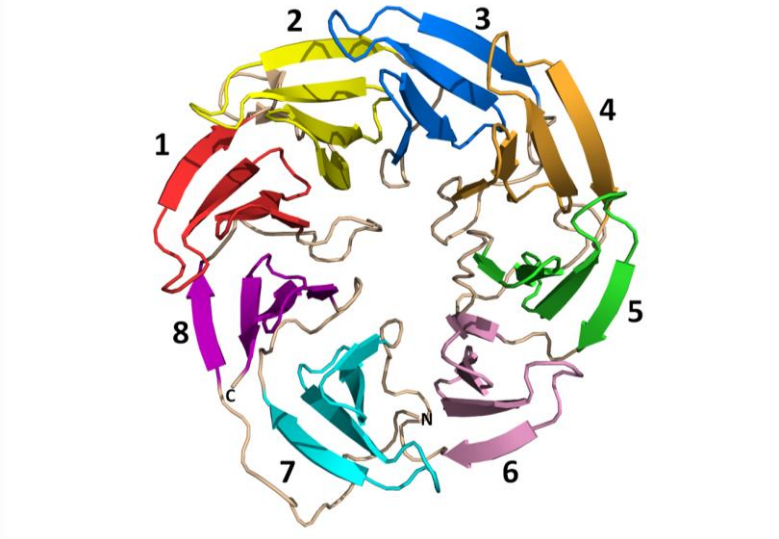


Figure 1.18: Crystal structure of BamB from *E. coli*. BamB forms an eight bladed β-propeller structure with a four-stranded β-sheet in each blade. Image created with PyMOL (PDB: 3PIL[273]).

In vitro experiments using a reconstituted BAM complex found that the absence of BamB reduced the efficiency of OMP assembly, consistent with *in vivo* results[233]. Thus, BamB is important in the BAM mechanism, but if it is not directly involved in binding OMPs, its main role may be to modulate the activity of BamA[268] (Section 1.4.1). Another possibility is that BamB interacts with SurA to facilitate the handover of nascent OMPs from SurA to BamA[272], but experimental evidence for such an interaction is lacking.

BamB, while not essential, is a good potential antibiotic target; useful drugs which act on proteins involved in OMP assembly need not be lethal. They may cause defects in the OM that allow the entry of other antibacterials, or may attenuate virulence mediated by OM factors. Recent evidence demonstrated that deletion of BamB from *Klebsiella pneumonia* attenuated virulence *in vivo*[274], highlighting the potential of BamB as a target, which may be less prone to selection pressure than essential bacterial proteins[275].

1.4.3 BamC

The role of BamC has also remained enigmatic. Genetic deletion of BamC results in reduced OMP levels and increased membrane permeability, but these defects are milder compared with deletion of BamB[125,276]. Structurally, BamC consists of an N-terminal unstructured region followed by two helix-grip domains connected by an α-helical linker (Figure 1.19).

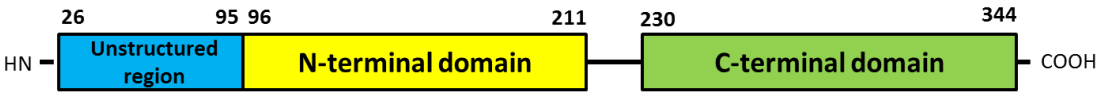


Figure 1.19 – Domain organisation of BamC. An N-terminal unstructured region (blue) is followed by two helix-grip fold domains separated by a short α-helical linker. The N-terminal domain (yellow) and the C-terminal domain (green) are both helix-grip domains.

The crystal structure of the C-terminal helix-grip domain revealed an antiparallel six-stranded β-sheet with one 3_{10} helix and two α-helices (Figure 1.20). The N-terminal BamC unstructured region is the most conserved part of the protein; the crystal structure of a BamCD complex demonstrated that it forms a 'lasso' which binds BamD over much of the BamD surface (Figure 1.21)[270].

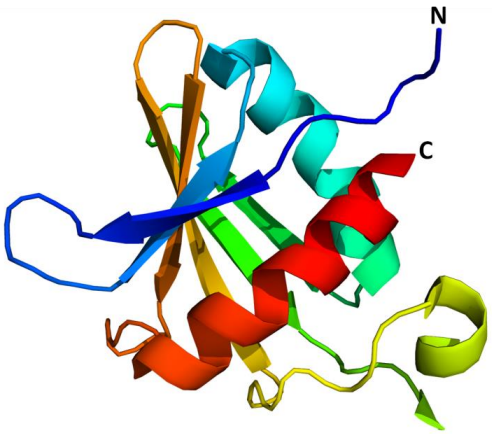


Figure 1.20 – Crystal structure of the *E. coli* BamC C-terminal domain. The domain forms a helix-grip fold consisting of a six-stranded β-sheet, two α-helices and a short 3_{10} helix. Image created with PyMOL (PDB: 3SNS[277]).

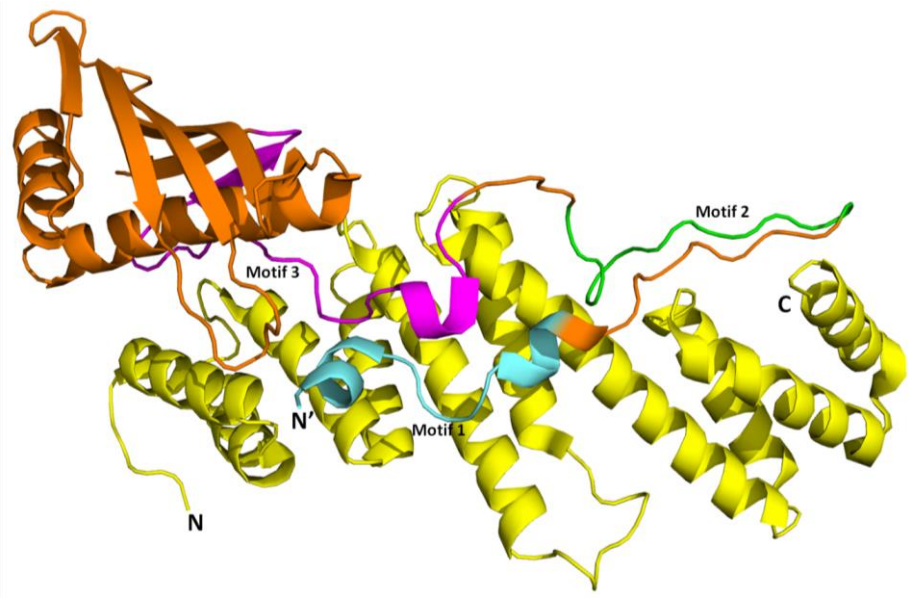


Figure 1.21: Crystal structure of the BamCD complex. The crystallised complex consists of full-length BamD and BamC (residues 26-217) lacking its C-terminal helix-grip domain. The unstructured N-terminal domain of BamC forms a lasso-like structure across the surface of BamD (yellow). Motifs 1, 2 and 3 in the lasso-like N-terminal region of BamC are highlighted in light blue, green and pink, respectively. The remaining residues in BamC, including the N-terminal helix-grip domain, are shown in orange. The N' denotes N-terminus of BamC. (PDB: 3TGO[270]).

Work by Webb and colleagues using immunofluorescence microscopy, protease shaving, and whole-cell ELISA experiments unexpectedly found that part of BamC is surface exposed on the extracellular face of the OM[72]. Further, a recent study was

able to use an anti-BamC antibody to track the spatial location of the BAM complex *in vivo*[37]. Bioinformatic analysis of the unstructured N-terminal domain of BamC revealed that it contains three conserved motifs (Motif 1, 2 and 3)[72]. Motifs 1, 2 and 3 correspond to the conserved regions $BamC_{28-42}$, $BamC_{54-68}$ and $BamC_{74-102}$, respectively (Figure 1.21). The crystal structure of the BamCD complex (Figure 1.21) reveals that a large majority of contacts with BamD are made by Motif 1 and Motif 2 of BamC[72,270]. Webb *et al.* further showed that only the first 25 residues of BamC are necessary for interaction with BamD, hence the residues in Motif 3 are not required for BamD interaction. The authors therefore suggest that both BamC helix-grip domains are extracellular with Motif 3 of the BamC N-terminal region spanning the OM and Motifs 1 and 2 interacting with BamD in the periplasm. The possible functional implications are not clear; the extracellular helix-grip domains of BamC may help to stabilise the interstrand loops of BamA[72].

1.4.4 BamD

BamD is the only essential BAM lipoprotein[241], and similarly to the other BAM lipoproteins, its exact functions are unclear. However, it has been suggested that BamD may play a role in initial substrate recognition[235,270]. Its structure consists of five tandem tetratricopeptide (TPR) motifs (Figure 1.22)[278,279]. TPR motifs are found in proteins which are involved in protein-protein interactions (PPIs), such as the Hop adaptor protein which mediates the interaction between the chaperones Hsp70 and Hsp90[280]. TPRs are also found in proteins which bind targeting sequences; PEX5, a cytosolic receptor for peroxisomal proteins, utilises TPR motifs to recognise and bind substrates via their C-terminal targeting sequences (Figure 1.23a)[281].

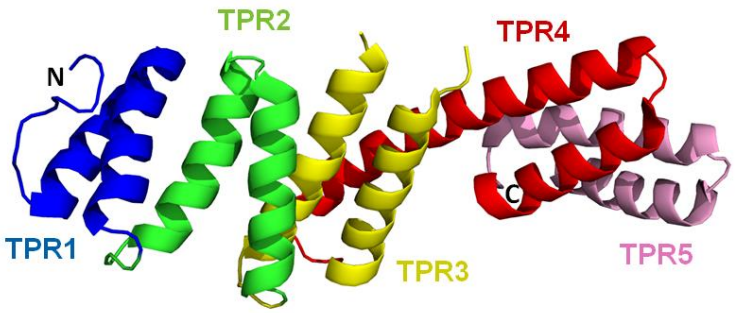


Figure 1.22: Crystal structure of BamD from *E. coli*. Five helix-turn-helix TPR motifs are separated by interconnecting loops. Image created with PyMOL (PDB: 3Q5M[278]).

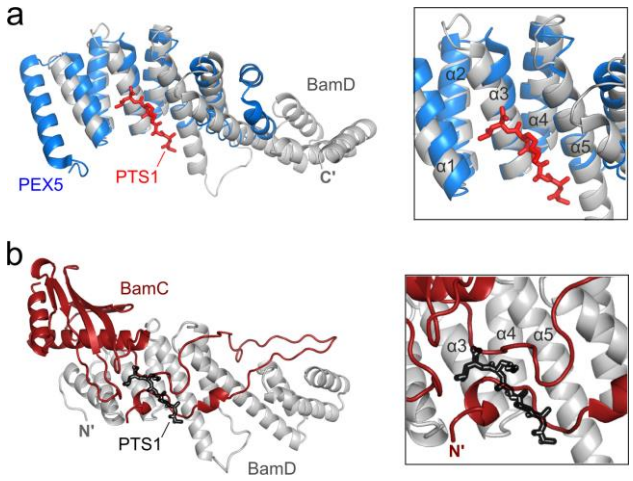


Figure 1.23: Proposed BamD binding pocket for OMP C-terminal targeting sequence. (a) The PEX5 receptor for peroxisomal proteins (blue) is shown aligned with BamD (grey). The pentapeptide containing the peroxisomal targeting sequence SKL (PTS1) which was crystallised with PEX5[281] is highlighted in red. (b) The N-terminal 'lasso' region of BamC (red) occludes the proposed OMP binding pocket in BamD (grey). The pentapeptide containing the PTS1 sequence, modelled in from the PEX5 crystal structure[281], is shown in black. Image reproduced from Kim *et al.* (2011)[270].

The structure of BamD is very similar to that of the TPR motifs of PEX5 (r.m.s.d. of 1.7 Å[270]), and peptides containing C-terminal OMP targeting sequences could be cross-linked to a truncated BamD construct of TPRs 1-3, as well as full length BamD[235]. Thus, it was proposed that the N-terminal binding pocket between TPRs 1 and 2 of BamD could possibly be the recognition site for OMP C-terminal targeting sequences (Figure 1.5). However, the structure of the BamCD complex shows that the N-terminal unstructured region of BamC also binds BamD in this region (Figure

1.21). This occludes the potential binding pocket, suggesting a possible mechanism for BamC to regulate the function of BamD (Figure 1.23b)[270]. However, the recent BAM complex structures show that the BAM lipoproteins and the BamA POTRA domains form a ring structure in the periplasm[221-223,232], and it seems likely that substrates enter this ring en route to the membrane. The N-terminal BamC 'lasso' which binds BamD faces out from, rather than into the periplasmic ring, suggesting that if BamD does bind OMP targeting sequences, it does so on the opposite face of the molecule to that proposed by Kim *et al*. (2011) (Figure 1.23)[270]. Recent *in vitro* evidence suggested that peptides based on OMP targeting sequences (β-signals) can bind BamD, and inhibit BAM function[136]. Further, *in vivo* expression of peptides containing β-signals caused growth defects and increased susceptibility to antibiotics. *In vivo* cross-linking studies showed that these peptides bound BamD, suggesting a promising starting point for the design of novel antibiotics which target the BAM complex[136].

Further evidence that BamD binds substrates separately to BamA comes from work by the Silhavy group which identified a temperature-sensitive lethal mutation (E373K) in POTRA domain 5 of BamA which prevents the association of the BamAB and BamCDE subcomplexes. Suppressor mutations in BamD (residue R197) could restore BAM function without re-establishing a stable interaction between the BamAB and BamCDE subcomplexes. The authors suggest this demonstrates that both BamD and BamA interact directly with OMPs, and that BamD is activated by the R197 mutations, a role normally performed by BamA[282].

One specialised role for BamD may be in the assembly of BamA. It was shown that BamD alone can assemble BamA into proteoliposomes composed of *E. coli* lipid extract, by contrast to OmpA, which requires additional BAM complex members[283].

1.4.5 BamE

BamE is the smallest protein in the complex and, similar to BamC, its deletion produces mild OMP assembly defects[126,276]. Structurally, BamE consists of two α-helices packed against a three-stranded antiparallel β-sheet (Figure 1.24)[234,236]. BamE can exist as a monomer or a domain-swapped dimer in solution[236], and there is conflicting evidence as to whether the dimer is relevant *in vivo.* One group found no BamE dimers in the periplasm[234], while another reported BamE dimers purified

from the OM[235]. Recently, native-MS of the intact BAM complex suggested a small population of complexes contain two BamE subunits[232].

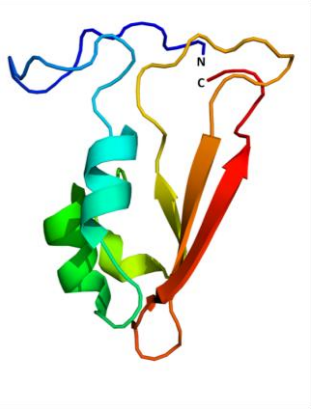


Figure 1.24: NMR structure of BamE from *E. coli*. The BamE fold has two α-helices at its N-terminal region and a three-stranded antiparallel β-sheet at its C-terminal region. Image created with PyMOL. (PDB: 2KXX[234]).

It was proposed that BamE may play a structural role as the BAM complex is destabilised in its absence, as judged by the difference in protein levels of BAM complex components that co-purify with His-tagged BamA between wild-type and Δ*bamE* strains[126]. NMR studies showed that BamE can bind simultaneously to BamD and PG, leading to the suggestion that BamE may be involved in anchoring the BAM complex to negatively charged PG-rich regions in the OM[284]. Native-MS experiments are needed to resolve if BAM selectively binds lipids in the OM[285].

BamE has been shown recently to play an essential role in the assembly of RcsF/OMP complexes, which are important in the Rcs stress response[60]. RcsF is an OM lipoprotein with a surface exposed region which senses envelope stress[286], and contains a hydrophilic transmembrane domain which is threaded through the lumen of OMPs[71]. In a *ΔbamE* strain RcsF/OmpA complexes were not detectable, demonstrating a key role for BamE in envelope stress signalling mechanisms[60].

1.4.6 Structural insights into the whole BAM complex

Despite the availability of structures for all individual components of the BAM complex, it was not until 2016 that the first insights into the overall architecture of BAM were obtained. First, a crystal structure of the BamACDE complex (lacking

BamB) was published[223], quickly followed by two crystal structures of the intact BamABCDE complex[221,222].

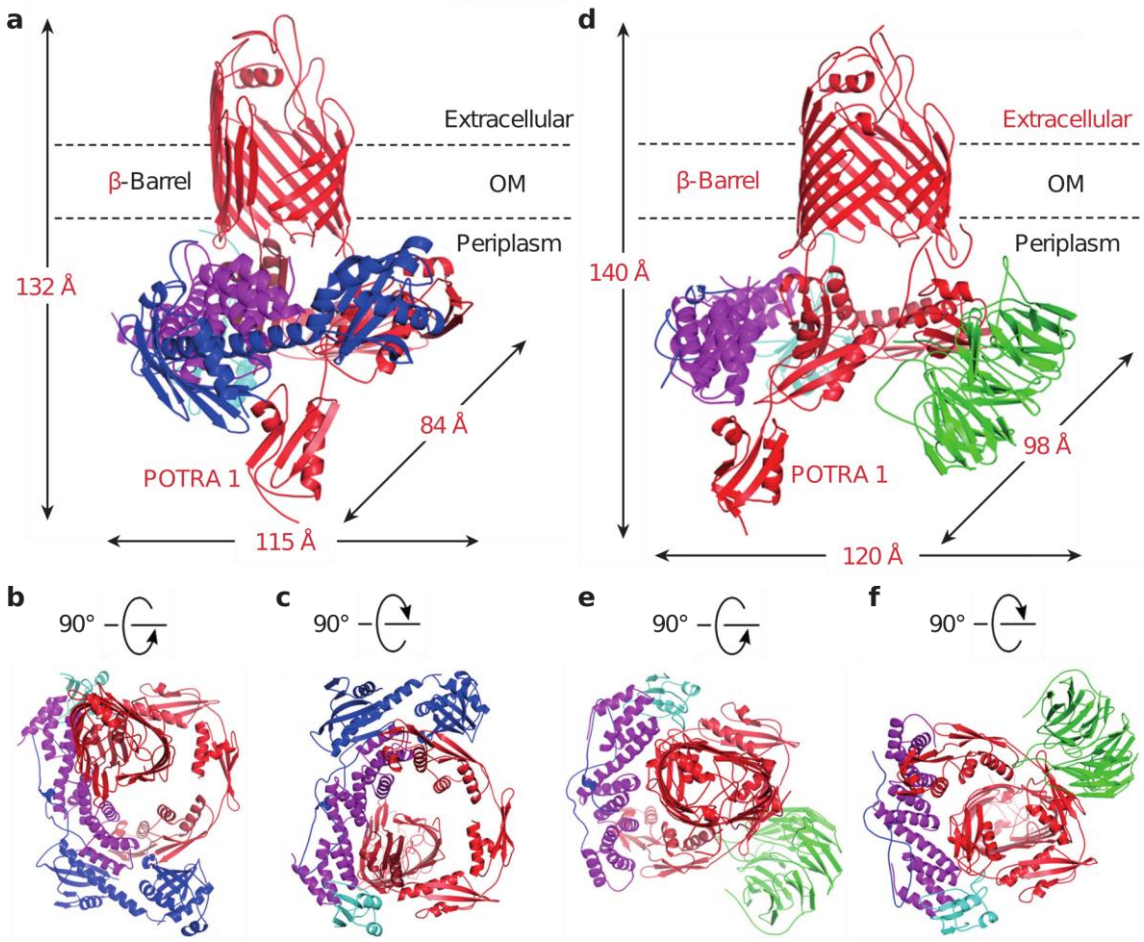


Figure 1.25: Crystal structures of BamACDE and BamABCDE complexes from *E. coli*. (a-c) Structure of BamACDE viewed from the (a) membrane plane, (b) extracelluar side, and (c) periplasm. BamD forms a ring structure by binding to POTRAs 1, 2 and 5. BamC binds to both BamD and POTRA 1 and 2 of BamA in this structure. BamE to both BamA and BamD. (d-f) Structure of BamABCDE viewed from the (d) membrane plane, (e) extracelluar side, and (f) periplasm. BamA-E are shown in red, green, blue, magenta, and cyan, respectively. Image reproduced from Gu *et al*. (2016)[221].

The BAM complex forms a 'hat-like' shape, with the BamA barrel forming the crown of the hat in the OM, and the BamA POTRA domains and BAM lipoproteins forming a ring structure (the brim of the hat) in the periplasm (Figure 1.25)[221,222]. Substrates presumably pass through this ring en route to the membrane. In the BamACDE structures the extracellular loops L1-3 which help form a dome over the barrel in previously published BamA structures are displaced upwards, and the first six β-strands (β1-β6) of the BamA barrel are rotated away ~65º from β16 to form an

opening in the lateral gate with a maximum distance of ~15 Å (Figure 1.25a)[221]. The lumen of the BamA barrel in the BamACDE structures is occluded by POTRA 5 and periplasmic turns T1-T4 of BamA (Figure 1.25b,c)[221,223]. This BAM conformation was termed 'lateral open'. By contrast, in the BamABCDE structures (Figure 1.25d)[221,223] the BamA barrel is in a closed conformation and sealed by extracellular loops, similar to that observed in previously published BamA structures[263,287]. In the BamABCDE structures the whole periplasmic ring structure in the periplasm is rotated by 30º compared with 'lateral open' structures, with the BamA barrel lumen open to the periplasm as it is no longer occluded by POTRA 5 (Figure 1.25e,f)[221,222]. Conformational changes also occur throughout the POTRA domains, with the largest changes occurring at the hinge region between POTRAs 2-3[232,257]. This BAM conformation has been termed 'inward open'[221], but is also referred to as 'lateral closed'[232].

The presence of two possible BAM conformations is consistent with bacterial genetic studies by the Silhavy group which identified two conformations of BamA *in vivo*, a protease sensitive conformation which could be labelled in the L6 loop by PEG-maleimide, and a protease resistant conformation which could not be labelled. These may correspond to the 'lateral open' and 'inward open' conformations, respectively, observed in crystal structures[221-223].

In the BAM complex crystal structures the 'lateral open' state was only observed in the absence of BamB, allowing the possibility that lateral opening in BamA may be driven by BamB dissociation. However, the recent publication of a cryo-EM structure of the intact BamABCDE complex in a 'lateral open' conformation excludes this possibility (Figure 1.26)[232].

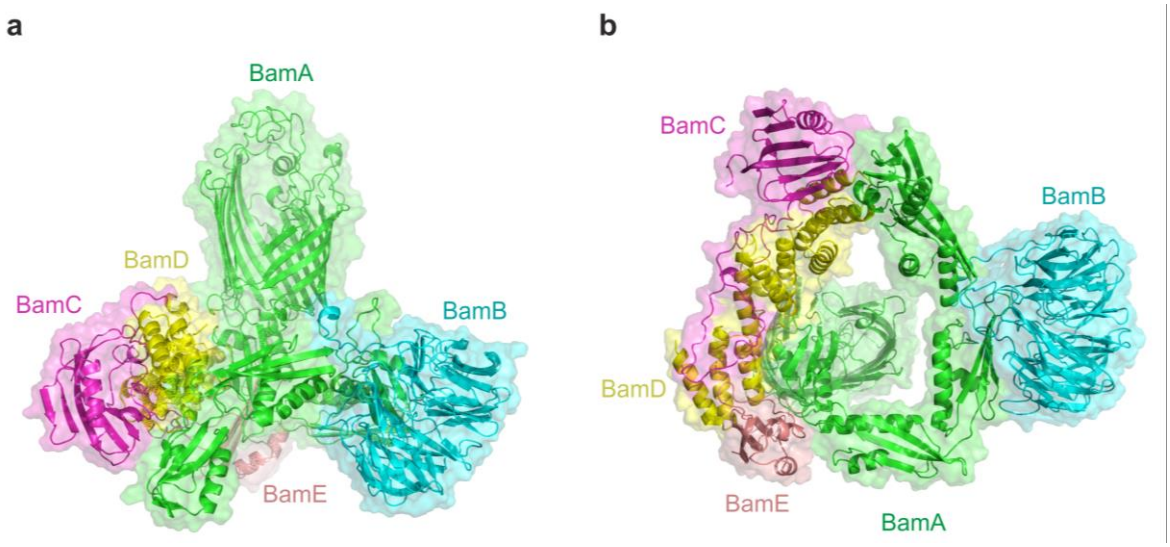


Figure 1.26: Cryo-EM structure of the *E. coli* BAM complex. BamA binds BamB via POTRAs 2 and 3. BamC binds BamD via its unstructured N-terminal region; no density for the BamC C-terminal domain is visible in the structure. BamD interacts with BamA separately mainly via POTRA 5, but also makes contacts with POTRAs 1 and 2. BamE binds BamD and also interacts with BamA via POTRA 5. The BAM complex structure viewed from (a) membrane plane, and (b) periplasm. Image created with PyMOL (PDB: 5LJO[232]).

The solution structure of the BAM complex, in contrast to the BAM complex crystal structures, also contained density for the detergent micelle. N-terminal residues of BamB, and loops containing hydrophobic residues in BamA POTRA 3 and BamD are observed making contact with the detergent micelle. This suggests that regions of the BAM complex other than the BamA barrel may be involved in contacting the membrane, and possibly modulating the lipid environment of the OM to facilitate OMP folding[232] (Figure 1.27). Consistent with this, the two tryptophan residues in POTRA 3 were observed inserting into the OM in MD simulations of BamA in a native membrane[262].

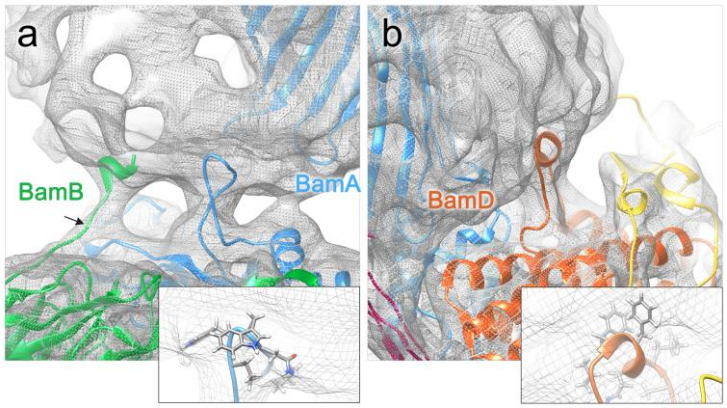


Figure 1.27: Interactions between BamA POTRA 3 and BamD and the detergent micelle in the cryo-EM structure of the BAM complex. (a) A hydrophobic loop within the body of BamA POTRA 3 (blue, white arrow) is buried within the micelle (grey) (inset). The N-terminus of BamB (green, black arrow), also makes contact with the micelle. (b) A hydrophobic 3_{10} helix in BamD (orange) is inserted into the micelle. Hydrophobic residues are buried in the acyl chains of the detergent (inset). Image reproduced from Iadanza *et al*. (2016)[232].

The two different conformations of the BAM complex in crystal structures has led to speculation that these represent 'resting' and 'active' states of the BAM reaction cycle, with these corresponding to the 'inward open' ('lateral closed') and 'lateral open', respectively. However, in the cryo-EM solution structure in DDM micelles, BAM is in the 'lateral-open' state, and there was no evidence of alternative conformations in the data set[232]. Therefore, the 'lateral open' conformation could be the true BAM resting state, and intriguingly, this leaves the β1 strand available for potential interactions with nascent OMPs via β-augmentation (Section 1.4.7). However, this conformation could possibly be due to the detergent used, and further structures in more native-like membranes will be needed to resolve this.

1.4.7 Models of β-barrel folding and insertion by the BAM complex

The least well understood part of BAM function in OMP biogenesis is the final folding and insertion step. In 2012, Kim *et al.* summarised the four models that have been proposed (Figure 1.28)[13]:

1) 'Export model': OMPs are translocated through the BamA pore before entry into the OM (Figure 1.28a). The BamA barrel lumen is ~25 Å in diameter[11,129], therefore sufficiently large to accommodate unfolded OMPs during translocation. However, it is not clear how OMPs could fold into the membrane from the

extracellular environment. If the C-termimal domains of BamC are indeed extracellular[72] and are involved, deletion of BamC would be expected to produce more severe OMP assembly defects than observed experimentally[125,276].

2) 'BamA-assisted model': OMPs are folded and inserted at the lipid-BamA interface with the BamA β-barrel domain acting as a scaffold (Figure 1.28b).
3) 'Oligomeric BamA-assisted model': Tetramerisation of the BAM complex allows the folding and insertion of OMPs in the space between the four complexes, again using the BamA β-barrel domain as a scaffold (Figure 1.28c).
4) 'Budding model': BamA β-barrel opens up between its N- and C-terminal β-strands which serve as templates for the newly folding OMP (Figure 1.28d). As the substrate completes folding and forms a β-barrel the terminal β-strands of BamA are released to return the BamA barrel to the closed state. This model has the merit that the energetic cost of opening up the BamA barrel is offset by the formation of new hydrogen bonds with the incoming substrate[13].

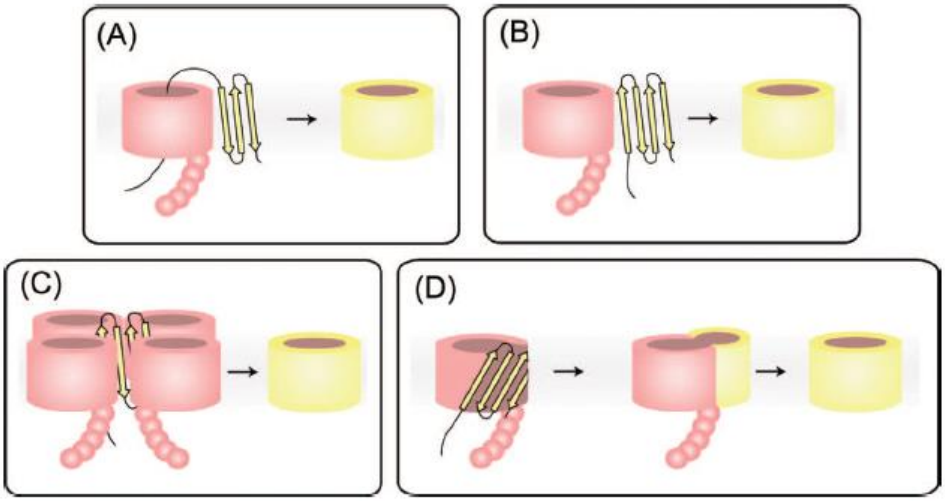


Figure 1.28: Four models of β-barrel assembly by the BAM complex. (a) OMPs are first translocated through the BamA channel. (b) The BamA β-barrel acts as a scaffold for folding. (c) OMPs fold between a BAM complex tetramer. (d) The BamA β-barrel domain opens up to template OMP folding, forming a hybrid barrel with the nascent OMP. Image from Kim *et al.* (2012)[13].

Reconstituted whole BAM complex in nanodiscs ~10 Å in diameter[288,289], a size unable to accommodate more than one BAM complex copy, allowed the assembly of the autotransporter EspP[239], arguing strongly against the need for BAM to oligomerise in order to be functional (Figure 1.28c). However, the finding that BAM is located in clusters with other OMPs in the OM ('OMP islands')[37,290] suggests that

the spatial closeness of individual BAM complexes may play a role in the assembly of some OMPs. These may include the trimeric porins which have been shown recently to be specifically affected by reduction in the expression levels of BamA and BamD[240].

In all four models the BamA POTRA domains could bind OMPs by β-augmentation prior to OMP insertion. These models are not necessarily mutually exclusive; the Buchannan group favour a hybrid model in which the BamA barrel destabilises the membrane ('BamA-assisted'), but is also involved in templating folding of nascent OMPs and the formation of a hybrid BamA:OMP barrel[263] (Figure 1.29). The molecular mechanisms underlying BAM activity may also vary depending on the substrate[240]; a single mechanism may be insufficient to explain all aspects of BAM activity. Much more work is needed to elucidate the molecular mechanisms of how the BAM complex catalyses folding as there is little evidence to support any of these models.

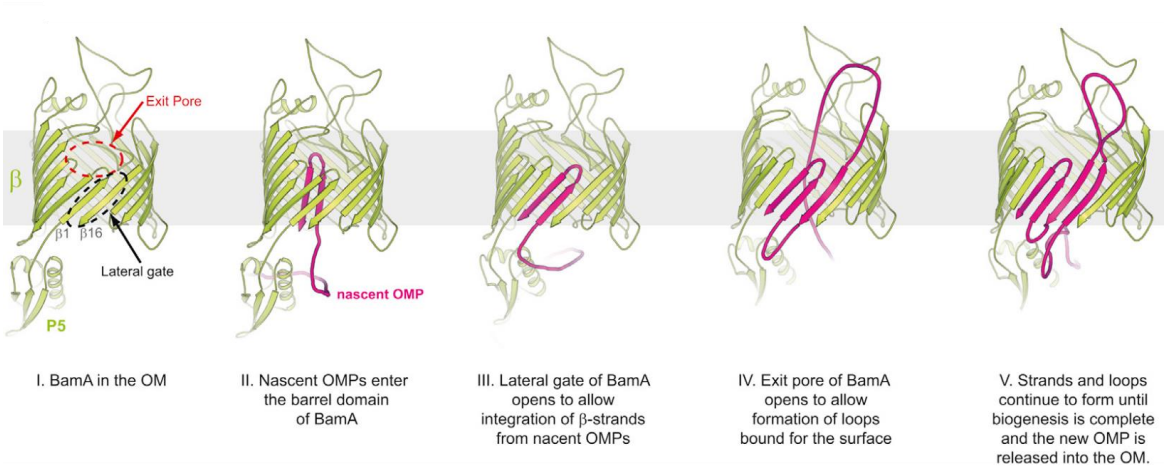


Figure 1.29: Model for BamA-mediated OMP biogenesis proposed by the Buchannan group. Image reproduced from Noinaj *et al*. (2014)[263].

1.5 *In vitro* studies of OMP folding

The field of membrane protein folding has lagged far behind that of soluble protein folding partly due to the lack of available membrane protein structures, although the number of new membrane protein structures has increased rapidly in recent years[89]. In addition, membrane proteins are much less experimentally tractable due to their high hydrophobicity[14,291]. Membrane protein folding is also complicated by the need

to consider the effects of the folding environment. The folding, stability and dynamics of membrane proteins are all influenced by general properties of the bilayer, such as curvature, hydrophobic thickness, and fluidity[292]. The folding environment can also influence the final native state of OMPs; the mitochondrial voltage-dependent anion-selective channel (hVDAC1) forms structures with different amounts of α-helical and β-sheet secondary structure depending upon whether the protein is folded in detergent or lipid[293].

There is some evidence that similar principles are involved in the folding of α-helical and β-sheet membrane proteins[291]. Both classes can fold spontaneously into bilayers to structures which are at free energy minima, can undergo reversible two-state folding under defined conditions, and can be assembled from polypeptide fragments[291,294,295]. However, α-helical and β-barrel proteins may not fold by the same mechanism(s) given the large differences in their structures. The folding of β-barrel membrane proteins may be more cooperative than the folding of α-helical proteins, given that the insertion of single β-strands into a membrane is thermodynamically more unfavourable than the sequential insertion of hydrophobic α-helices[296].

In vitro methods involving purified OMPs, model membranes and chemical chaotropes have been used to study the thermodynamics and kinetics of OMP folding and unfolding for over 20 years[297,298]. Conditions for *in vitro* folding into a large variety of lipid bilayers and detergent micelles have been established for many OMPs, using a variety of buffers, pH values, temperatures and denaturants[291,298]. *In vitro* studies have also begun to explore the influence of chaperones[152,153,299] and BAM components[233,239,283,300-302] on *in vitro* OMP folding. This brief review of the major insights into the kinetics and thermodynamics of OMP folding from *in vitro* studies will focus on results obtained for OmpA and PagP, as these are two of the best characterised OMPs, and experimental results have been obtained for both these OMPs in the current project.

Most OMP folding studies have been performed using the OmpA porin from *E. coli*[298], the most abundant OMP in the OM (~200,000 copies/cell)[83]. OmpA is a 325-residue protein consisting of an N-terminal 171-residue 8-stranded β-barrel domain (Figure 1.30a)[303], and a 154-residue globular C-terminal periplasmic domain (Figure 1.30b)[304] which interacts with the cell wall[305]. OmpA is a multifunctional protein and

has proposed roles as a membrane pore, an immune evasin, an adhesion, and as a receptor for certain bacteriophages, in addition to its structural role[306].

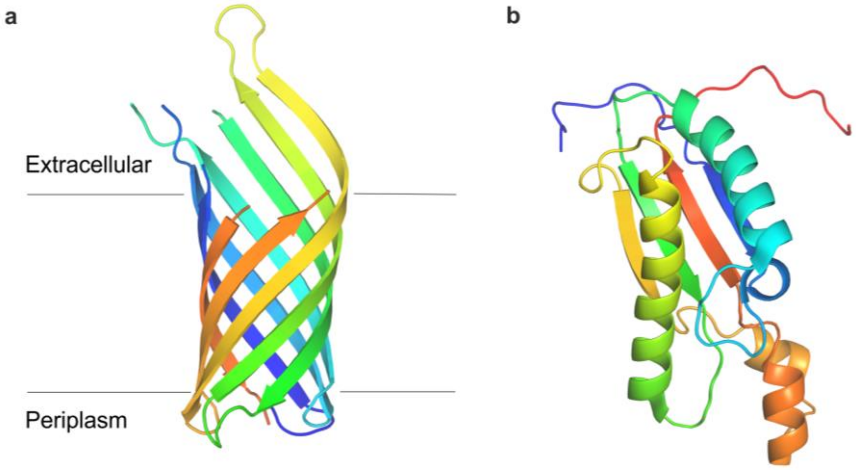


Figure 1.30: Structures of the OmpA N-terminal β-barrel domain and the C-terminal soluble domain. (a) Crystal structure of the N-terminal residues of OmpA (1-171) which form an 8-stranded β-barrel in the OM (PDB: 1QJP[303]). Approximate location of the membrane is indicated by black lines. (b) NMR structure of the C-terminal OmpA soluble domain, residues 172-325 (PDB: 2MQE[39]). Images created with PyMOL.

In landmark studies the Jähnig group first demonstrated that urea-denatured OmpA could be refolded spontaneously *in vitro* into micelles of the detergent β-octylglucoside[307] and lipid vesicles[297]. They took advantage of the fact that folded and unfolded OMPs migrate at different apparent molecular weights if not boiled prior to SDS-PAGE analysis ('heat-modifiability')[308-310], a technique which has proved invaluable in the field. Successful OmpA folding *in vitro*[307,311] allowed the exploration of the factors affecting OmpA folding by systematic variation of the folding conditions and environment[305]. OmpA can fold to its native state in a large variety of detergents, provided the detergent concentration is above its critical micelle concentration (CMC)[312], as well as into lipid containing Small Unilamellar Vesicles (SUVs) (~50 nm) and Large Unilamellar Vesicles (LUVs) (~100 nm)[53,313]. Detailed kinetic studies on the folding of OmpA over a wide range of temperatures (2 ºC – 40 ºC) revealed that OmpA can fold via a sequential pathway containing a number of folding intermediates[314]. Urea-denatured OmpA was proposed to refold into dioleoyl phosphatidylcholine (DOPC) bilayers with at least three distinct kinetic intermediates: (1) a rapidly formed (<1 s) I_W state, formed via hydrophobic collapse as the folding reaction is initiated, as detected by Trp fluorescence, (2) a membrane-

adsorbed state, I_{M1}, and (3) a second-membrane-adsorbed state, I_{M2}, in which β-hairpins may be partially inserted into the membrane[314,315]. The proposed folding reaction pathway (Figure 1.31) was supported by later kinetic experiments using a technique termed time-resolved distance determination by fluorescence quenching (TDFQ)[316]. The locations of tryptophan residues in single tryptophan OmpA mutants within the bilayer could be measured during folding by the quenching of their fluorescence by brominated lipids. By adjusting the position of bromine atoms in the bilayer, Kleinschmidt and colleagues were able to show that all four β-hairpins of the OmpA barrel cross the bilayer simultaneously, providing evidence for a concerted folding and insertion mechanism[317]. Subsequently, the kinetics of OmpA secondary and tertiary structure formation on folding into *di*$C_{12:0}$PC (DLPC) LUVs were investigated by CD and gel electrophoresis, respectively[53]. The rate constants for these processes were the same, again suggesting a concerted folding mechanism[53]. Consistent with this model, recent hydrogen-deuterium exchange (HDX) experiments, which monitored OmpX folding into detergent micelles, found that the rate of hydrogen bond formation between β-strands was the same between all β-strands, and synchronised to tertiary structure formation[318].

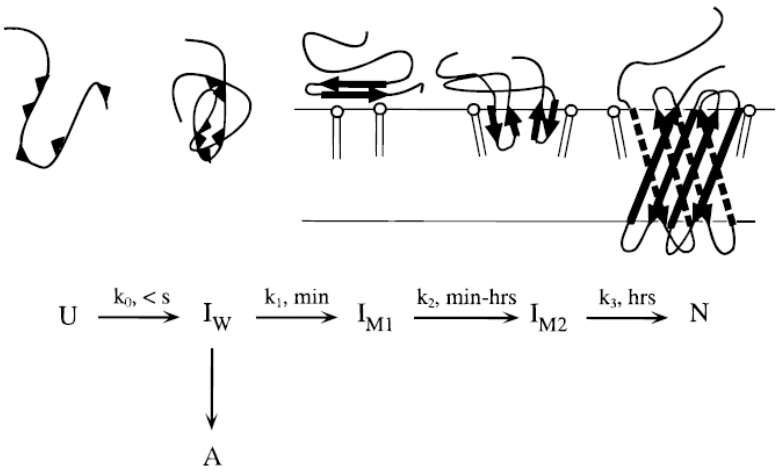


Figure 1.31: Proposed OmpA folding reaction pathway. Following rapid hydrophobic collapse to I_W a fraction of OmpA molecules may form aggregrates (A) while the remainder proceed to a membrane bound intermediate (I_{M1}). A second partially intermediate (I_{M2}) is followed by folding to the native state (N). Image reproduced from Kleinschmidt & Tamm (1996)[314].

OmpA, has also been suggested to be able to fold via parallel pathways in *di*$_{C18:1}$PC (DOPC) vesicles[314,319], similar to results obtained for the 14-stranded OMP FomA[319,320]. Parallel pathways were observed for the folding of OmpA into micelles

of the detergent octyl maltoside[321]. In both cases, the origin of the 'fast' phase has been attributed to direct binding to lipids or surfactant on dilution from denaturant, with the 'slow' phase due to hydrophobic collapse to an unfolded aqueous state prior to adsorbtion[322]. Consistent with this notion, for the parallel pathways observed for PagP folding, the amplitude of the 'slow' phase (i.e. the proportion of molecules following this pathway) increases as the lipid:protein ratio (LPR) is decreased[323].

The rate of association of neighbouring OmpA β-strands during folding was measured by fluorescence quenching in mutants containing a single tryptophan residue and a single nitroxyl-labelled cysteine residue[324]. The results indicated that, in the membrane-adsorbed state, OmpA β-strands are not in close proximity, and that the neighbouring β-strands on the extracellular side appear to associate before those on the periplasmic side[324].

Truncation of the four surface-exposed extracellular loops of OmpA showed that none appear to be important for OmpA folding *in vivo*[325]. OmpA was also found to be tolerant of mutations within individual β-strands provided that: (1) residues are not mutated to proline; (2) most of the lipid facing residues remain hydrophobic, and no charged residue is placed in the central region of the β-strand; and (3) side-chains facing the lumen of the barrel are not too enlarged, presumably due to packaging constraints[326]. Recently, computational methods were used to completely redesign the hydrophobic surface of the OmpA β-barrel[327]. The results generated a mutant with ~60% of the lipid-facing residues of OmpA replaced which was still able to insert and fold into membranes[327].

The *in vivo* folding environment of OMPs is the asymmetric OM bilayer containing phospholipids in the inner leaflet, and LPS in the outer leaflet[21] (Section 1.2). LPS does play a role in assisting the assembly of at least some OMPs *in vivo*[61], however recapitulating asymmetric bilayers *in vitro* is technically challenging[322]. LPS was shown not to be required for OmpA folding *in vitro*[307]. However, the efficiency of refolding of PhoE *in vitro* is improved by LPS[62,328], with this effect likely due to the presence of negative charges in the inner core region (Figure 1.2b)[329]. Steeghs *et al.* were able to isolate a mutant strain of *N. meningitidis* which was unable to synthesise Lipid A, demonstrating that LPS is not essential for OMP folding *in vivo* in at least one organism[330].

Spontaneous refolding into membrane mimetics has now been observed for many different OMPs[291]. In one study the folding behaviour of nine OMPs (OmpX, OmpW, OmpA, PagP, OmpT, OmpLA, FadL, BamA and OmpF) with differing functions and ranging in size from 8-16 β-strands was investigated[313]. Folding conditions were systematically varied and the following interesting observations were made:

(1) Higher pH values promote folding. In DLPC LUVs higher folding efficiency was observed for all OMPs as the folding environment becomes more basic. The OMPs in the study have a pI of between 5-6, so this effect may be due to the increase in their net negative charge as the pH is raised.
(2) Thinner bilayers increase the kinetics of folding. OMPs were folded into three different LUVs containing either *di*$C_{10:0}$PC (DDPC), *di*$C_{11:0}$PC (DUPC), or *di*$C_{12:0}$PC (DLPC) lipids. All OMPs folded fastest in DDPC bilayers, consistent with previous results obtained for OmpA[53].
(3) Increased elastic curvature promotes folding. A comparison of OMP folding into SUVs and LUVs showed that folding efficiency was generally enhanced in SUVs which have a higher elastic curvature, and exhibit more defects than LUVs[305].
(4) No trends emerged when the effect of temperature on folding efficiency was examined. In folding experiments ranging between 30 °C and 50 °C in DDPC LUVs, the folding efficiency of the OMPs either increased, decreased or remained the same as temperature was increased. This varied behaviour was not correlated to β-barrel size.

Much work has been carried out trying to establish conditions for the reversible two-state folding of OMPs in order to calculate their thermodynamic stabilities. Protein stability, or the Gibbs free energy of folding ($\Delta G°_{UN}$), is obtainable when the population of folded and unfolded forms of a protein population can be measured at equilibrium interchanging reversibly within the time frame of an experiment[291]. A protocol resulting in reversible denaturation was first established for the folding of urea-denatured OmpA into SUVs[331]. Equilibrium unfolding experiments were carried out in a range of different bilayers containing different phospholipid head groups and different types of acyl chains at 37.5 °C. A value for $\Delta G°_{UN}$ of -3.4 kcal/mol was first obtained in a reference bilayer of $C_{16:0}C_{18:1}$PC containing a 7.5% mole fraction of negatively charged $C_{16:0}C_{18:1}$PG. Substitution of the $C_{16:0}C_{18:1}$PC lipids for guest lipids allowed the investigation of how the physical properties of the membrane,

such as curvature stress and the hydrophobic mismatch between OMP hydrophobic length and bilayer thickness, modulate OMP stability. Increasing the intrinsic curvature stress by inclusion of $C_{16:0}C_{18:1}$PE lipids increased OmpA stability; substitution with 30% $C_{16:0}C_{18:1}$PE increased $\Delta G°_{UN}$ by ~60%. Substitution of the reference bilayer lipid with lipids that decreased the membrane hydrophobic thickness reduced OmpA stability; addition of 30% *di*$C_{10:0}$PC (DDPC) reduced $\Delta G°_{UN}$ by ~50%. A PagP construct containing a C-terminal His-tag (HT-PagP) was later shown to fold reversibly in DLPC bilayers between urea concentrations of 7 M and 10 M (Section 1.5.1)[332]. HT-PagP stability was calculated as -14.4 kcal/mol, somewhat higher than that reported for OmpA in SUVs (<7 kcal/mol)[331].

Conditions for reversible folding of OmpLA into DLPC liposomes were also obtained by a thorough screen of parameters affecting folding (lipid composition, temperature, salt concentration, incubation time, pH and liposome morphology)[333]. Under reversible folding conditions a $\Delta G°_{UN}$ value of -32.5 kcal/mol was obtained for the free energy of unfolding of OmpLA. The authors found that three factors produced reversible folding: acidic conditions (pH 3.8), the gradual dilution of OmpLA into folding conditions, and the presence of small amounts of the detergent 3-N,N-dimethylmyristylammonio)propanesulfonate (SB3-14) to act as a solubilising 'holdase'. SB3-14 was selected as it does not disrupt the integrity of DLPC LUVs, and has a low CMC (~0.1-0.4 mM, 25 °C)[334] and low aggregation number (83 - the number of molecules per micelle), enabling a low concentration to be used[333]. This reversible OmpLA folding system at pH 3.8 was then used to develop a hydrophobicity scale for the water-to-bilayer transfer free energies of the 20 natural amino acids into a lipid bilayer (the Moon-Fleming scale)[335]. The Fleming group subsequently used the conditions optimised for OmpLA to achieve path-independent folding for two further OMPs, OmpW and PagP[336]. However, the stability of not all OMPs can be obtained with this method as reversible folding was not achieved for OmpX, OmpT and FadL. Despite this, the study allowed the comparison of OMP stabilities obtained under the same experimental conditions for the first time. Values for $\Delta G°_{UN}$ of -18.3 ± 0.5 kcal/mol and -24.4 ± 0.4 kcal/mol were obtained for OmpW and PagP, respectively. The $\Delta G°_{UN}$ value for PagP is approximately 10 kcal/mol more favourable than that obtained previously by Huysmans *et al.*[332]. Fleming and coworkers account for this difference by differences in experimental conditions. In the Huysmans study urea was used; PagP may retain some residual structure in the unfolded state, and was believed to remain adsorbed to the membrane in the

unfolded state in this denaturant[332]. In the Fleming study, the more potent denaturant GuHCl was used, leading to complete PagP dissociation from the membrane in the unfolded state. The authors calculate a value of -10.74 kcal/mol for the free energy of transfer of the PagP sequence from water to the DLPC bilayer using the Wimley-White interfacial scale, thus accounting for the apparent discrepancy in PagP stabilities[336].

An interesting correlation was found between the stabilities of OmpW, PagP and OmpLA and the water-to-bilayer transition free energy of the residues which are in contact with the bilayer, calculated using the Moon-Fleming hydrophobicity scale[335]. The authors suggest that the large thermodynamic stabilities observed for OMPs may serve as a sorting mechanism for OMPs in the periplasm[336]. While the low stability value obtained for OmpA in lipid vesicles (<7 kcal/mol) may argue against this, the Otzen group reported a much higher value of -12-15 kcal/mol for the reversible folding of OmpA in octyl maltoside micelles using GuHCl as the denaturant[321], and a similar value (-14.3 kcal/mol) was obtained for urea-denatured OmpA stability in LDAO micelles[337]. Interestingly, in a study on the folding of OmpA into the ampiphatic polymer (APol) A8-35 the thermodynamic stability was measured as only -1.9 kcal/mol, and a 52 day incubation time was required before the folding and unfolding titration curves overlaid[337]. These results suggest that the increased stability observed for membrane proteins in APols compared to detergents and lipids[338] is due to kinetic trapping of the native state by a very large activation barrier to unfolding, rather than an increase in thermodynamic stability[337].

1.5.1 PagP

PagP (PhoP/Q activated gene P) is a 163-residue OM acyltransferase enzyme which catalyses the transfer of a palmitate residue from the *sn*-1 position of a phospholipid to a Lipid A moiety of LPS[339,340] (Figure 1.32). This reinforces the OM, providing protection against cationic peptides[341], and aids immune evasion by antagonising Toll-Like Receptor 4 (TLR4) signalling[342].

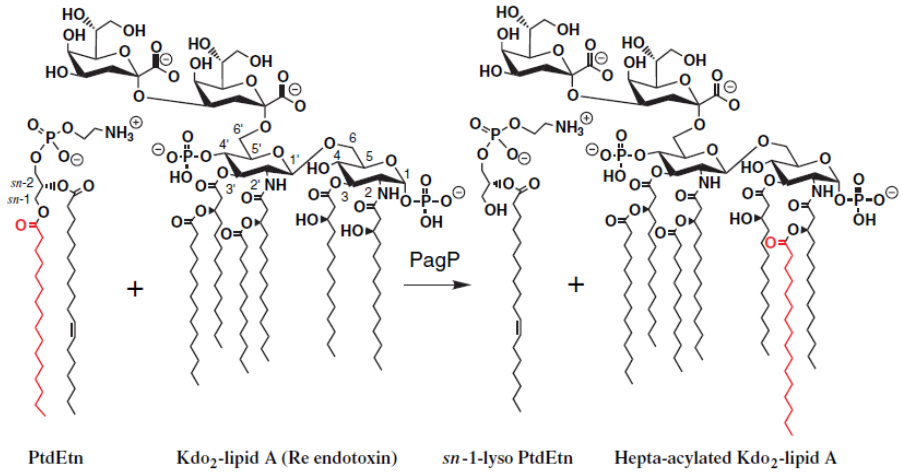


Figure 1.32: Reaction catalysed by PagP. A palmitate residue (red) from the *sn*-1 position of phosphatidylethanolamine (PtdEtn) is transferred to the Lipid A moiety of LPS. Image reproduced from Bishop (2005)[342].

The structure of PagP, solved by NMR[343] and X-ray crystallography[344,345], reveals an 8-stranded β-barrel with a 19-residue N-terminal α-helix which associates with the β-barrel on the periplasmic side of the OM *in vivo* (Figure 1.33). PagP can be produced easily in large quantities in the laboratory, in common with most OMPs[346]. Highlighting this, the yields of intrinsically disordered proteins have been increased by targeting them to inclusion bodies by fusion to PagP, with protein expression yields of >100 mg fusion protein per litre of minimal media obtained[347].

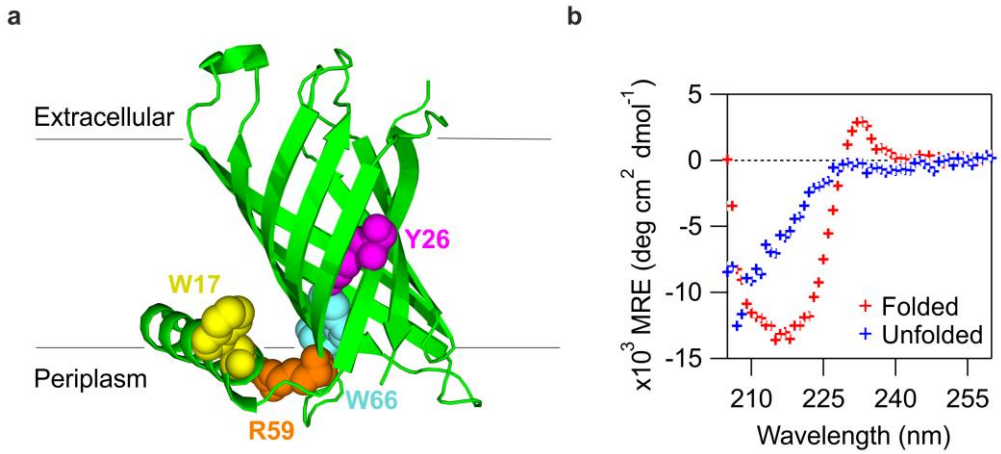


Figure 1.33: Native PagP forms an 8-stranded barrel with a 19-residue N-terminal helix. (a) Crystal structure of PagP (green). Residues which destabilise the barrel on mutation, W17 and R59, are highlighted in yellow and orange, respectively. Residues involved in exiton formation leading to a positive peak in the native far-UV CD spectra at 232 nm, Y26 and W66, are shown as magenta and cyan spheres, respectively. The native structure is tilted approximately 25° to the normal of the membrane. Approximate location of the membrane indicated by black lines. Image created with PyMOL (PDB: 3GP6[345]). (b) Example far-UV CD spectra of folded and unfolded PagP (Section 2.10.4).

A high concentration of urea is required to solubilise an aggregation-prone C-terminal His-tagged construct of PagP[152,348]. High folding efficiency was achieved in cyclofos-7 detergent micelles and DLPC LUVs in the presence of 7 M urea[348]. Huysmans *et al.* investigated the role of the unusual N-terminal α-helix (Figure 1.33) in the folding and stability of PagP using variants in which either the helix was removed (PagPΔ1-19), or mutations were made which interfered with the interactions between the β-barrel and the helix (W17A and R59L)[348]. The formation of secondary and tertiary structure in detergent micelles (cyclofos-7) and DLPC liposomes were monitored by far-UV CD; the spectrum of folded PagP has a negative maximum at 218 nm typical of β-sheet secondary structure, and an additional positive peak at 232 nm which has been attributed to a 'Cotton effect' between residues W66 and Y26 (Figure 1.33)[349]. PagP folding was also monitored by tryptophan fluorescence, Fourier transform infrared spectroscopy (FTIR), changes in electrophoretic mobility on cold SDS-PAGE gels, and gain of enzyme activity[348]. It was found that secondary and tertiary structure of PagP forms simultaneously, as previously reported for OmpA[53]. The results also showed that none of the mutants prevented folding of the β-barrel domain into detergent or liposomes. All mutants, however, were thermodynamically destabilised relative to

wild-type. Interestingly, the decrease in stability of the W17A mutant was dependent on the folding environment and was much more destabilising in the liposome experiments compared with those in cyclofos-7 detergent micelles. Kinetic experiments in DLPC liposomes established that the destabilisation of all mutants was due to an increase in the dissociation rate from the lipid-inserted state. This led to the conclusion that at least one of the functions of the PagP helix is as a 'post-assembly clamp' which docks with the folded PagP β-barrel, increasing its stability[348].

Subsequent work established that PagP undergoes fully reversible two-state folding and unfolding between urea concentrations of 7 M and 10 M in 100 nm DLPC liposomes at a molar LPR of 3200:1[332]. Folding was monitored by tryptophan fluorescence and was found to be independent of protein and lipid concentration (for PagP concentrations of 0.1-0.4 μM and LPRs between 800:1 and 4000:1). Equilibrium denaturation experiments enabled the calculation of the thermodynamic stability of PagP in DLPC LUVs, yielding a free energy of unfolding value, $\Delta G°_{UN}$, of -14.4 ± 0.4 kcal/mol, and a calculated value for M_{UN} (a measure of the surface area buried on folding to the native state) of 1.64. ± 0.05 kcal/mol/M. The urea dependence of PagP folding and unfolding kinetics was measured by initiating folding and unfolding reactions from PagP in 10 M and 7 M urea, respectively. Unfolding reactions fitted well to a single exponential function at all urea concentrations, while the folding kinetics below 7.8 M urea were complicated by the presence of a large burst phase. A chevron plot (which plots the linear dependence of the logarithm of the observed rate constant on denaturant concentration) of the PagP kinetic data was constructed (Figure 1.34). Analysis of the kinetic data provided confirmation of reversible two-state folding as: (1) the denaturation midpoint (C_m), where the folding and unfolding arms meet, was in good agreement with the equilibrium stability data; (2) the sum of the kinetic *m*-values (m_f and m_u) obtained from the slopes of the folding and unfolding arms of the chevron plot were in agreement with the M_{UN} value obtained from equilibrium experiments; and (3) the value for the free energy of unfolding, $\Delta G°_{UN}$, obtained from equilibrium denaturation agreed well with that obtained from the kinetic data (using $\Delta G°_{UN} = -RT(k_f/k_u)$), where R is the gas constant, T is the temperature in Kelvin, and k_f and k_u are the rate constants of folding and unfolding, respectively.

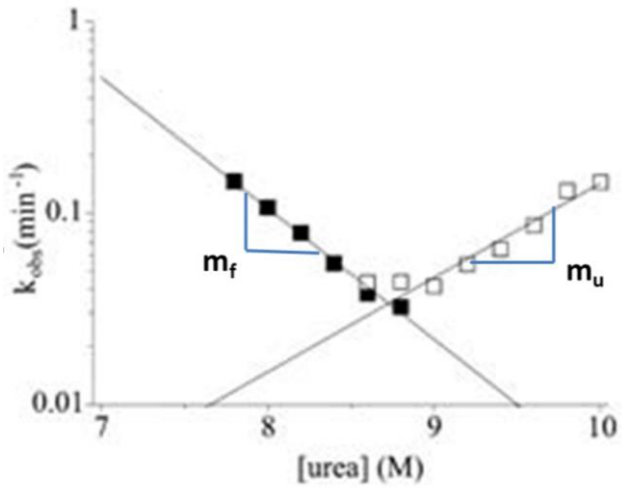


Figure 1.34: Chevron plot of PagP kinetic data. The logarithm of the observed rate constants for folding and unfolding depend linearly on the concentration of denaturant. The kinetic *m*-values m_f and m_u are obtained from the slopes of the folding and unfolding limbs are indicated. Image adapted from Huysmans *et al*. (2010)[332].

Verification of the reversibility of PagP (un)folding under these conditions allowed the construction of the first model of the transition state structure of an OMP using Φ-value analysis. 19 single amino acid mutants across the entire PagP structure were constructed and the proteins purified. The change in stability of the transition state was calculated from kinetic experiments in the unfolding direction (as folding did not fit a single exponential below 7.8 M urea) and the change in stability of the native state was measured by equilibrium urea titration experiments[332]. A *Φ*-value was calculated for each mutant by taking the ratio of these two values, giving a number between 0 and 1. *Φ*-values of 0 suggest that the mutated residue does not make native-like contacts in the transition state, while *Φ*-values of 1 suggest that the residue is as structured in the transition state as it is in the native state[350,351]. The results indicated a highly polarised transition state in which the β-barrel structure is largely formed (Figure 1.35C). Five residues which had a *Φ*-value of >0.5 were present in the C-terminal half of the structure (strands E-H), while five variants with a *Φ*-value of <0.3 were located near the N-terminal helix, consistent with the docking of the helix as the final folding step[348] (Figure 1.35A and Figure 1.35B). Two of the mutants gave rise to negative *Φ*-values, which the authors interpret as an indication that non-native contacts are formed in the transition state in the region of these mutations. Taken together, the results suggest a tilted insertion mechanism in which the C-terminal half of PagP enters the bilayer and partially folds first while the N-

terminal helix remains unstructured[332]. Interestingly, in coarse-grained simulations of OmpA folding into lipid bilayers the β-barrel domain entered the membrane at an angle of approximately 45° to the bilayer normal, hinting that tilted insertion may be a more general OMP folding mechanism[352].

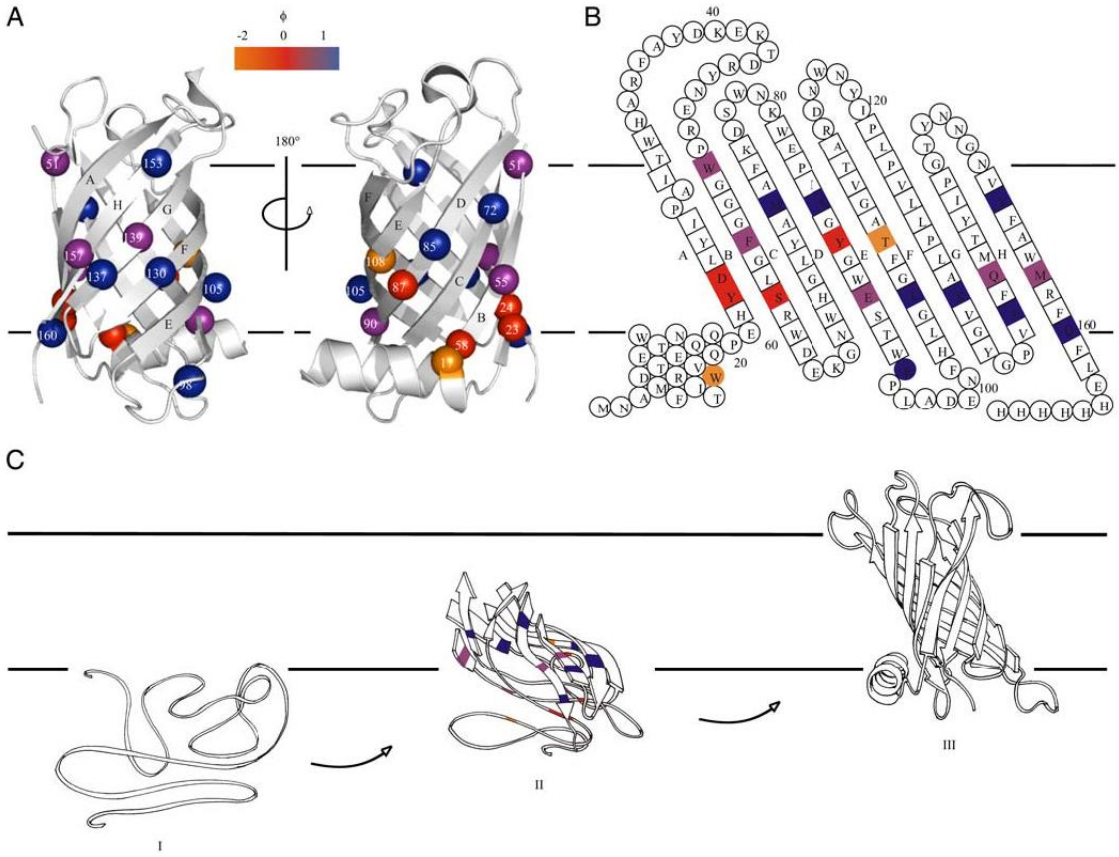


Figure 1.35: *Φ*-value analysis of PagP folding in DLPC LUVs. (a) Location and *Φ*-values of PagP mutated residues in cartoon representation. (b) Topology map and *Φ*-values of PagP mutated residues indicating the mutated residues cover all 8 β-strands. (c) Model of the polarised transition state and proposed tilted-insertion mechanism. *Φ*-values are coloured as indicated in the scale bar. Image reproduced from Huysmans *et al*. (2010)[332].

The *β*-Tanford value (β_T), which is defined as m_f/M_{UN}, gives a measure of how far the transition state is along the reaction coordinate. The β_T for PagP was calculated as 0.6, suggesting that the transition state is approximately midway along the folding reaction coordinate. This contrasts sharply with the much lower value of ~0.1 reported for the α-helical protein bacteriorhodopsin (bR) in SDS, which suggests the bR transition state is close to the unfolded state[353,354]. This difference could be due to differences in residual secondary structure in the denatured states of α-helical

and β-barrel membrane proteins, or may reflect differences in their folding mechanisms.

A subsequent study investigating the folding kinetics of PagP into DLPC liposomes in the presence of 7 M urea from an unfolded state (10 M urea) revealed parallel folding pathways[323]. On dilution from denaturant, a burst phase is observed by tryptophan fluorescence as PagP adsorbs to the lipid bilayer, followed by a time-course as folding and insertion occurs. At high LPRs this time-course fits to a single exponential, while at low LPRs a second exponential with a slower rate constant is present. The effect on PagP folding of altering the properties of the lipid bilayer by the addition of guest lipids was also investigated. Addition of di-lauryolphosphatadylethanolamine (DLPE), which increases the lateral pressure and acyl chain packing of the bilayer, led to an increase in the population of PagP molecules following the slower folding pathway, while addition of di-lauroylphosphatadylserine (DLPS), which increases the overall net negative charge of the membrane, reduced the folding rate of the fast folding pathway. These results not only demonstrate that PagP can fold by different pathways, but also highlight the importance of lipid-protein interactions in modulating these pathways[323].

1.5.2 Studies of BAM complex function *in vitro*

A major breakthrough in the field came in 2010 with the *in vitro* reconstitution of the complete BAM complex by the Kahne group using purified components[233]. It was known that BamB binds BamA separately from BamD, and BamC and BamE depend on BamD for their interaction with BamA[241,247]. Therefore, the BamAB and BamCDE subcomplexes were co-expressed and purified separately, before holocomplex reconstitution in proteoliposomes containing *E. coli* polar lipid extract. The protease OmpT was chosen as the substrate as its rate of folding into proteoliposomes could be monitored by the cleavage of a fluorogenic peptide[355]. The folding rate of OmpT was orders of magnitude higher in the presence of the BAM complex compared with lipid alone. Folding rate enhancement was achieved without the input of an energy source, and the folding rate observed (seconds to minutes) was on a similar time-scale to that observed by *in vivo* pulse-chase experiments for PhoE[356] and LamB[123]. Verification that the increase in fluorescence corresponded to an increase in the level of folded OmpT was carried out by cold SDS-PAGE. After 30 minutes, reactions containing ^{35}S-labelled OmpT were centrifuged and samples from the pellets analysed by cold SDS-PAGE. The results

showed that while the folding yield of OmpT was low (7%) in the presence of the complete BAM complex, in its absence the yield was negligible. Loss of BamB from the complex greatly reduced the folding rate and yield as did loss of BamCDE. The presence of SurA increased the OmpT folding rate in a concentration-dependent manner[233]. However, how much of this effect is due to solubilisation of the substrate, rather than specific targeting to BAM by SurA is unclear. In a later study, it was shown that SurA could be functionally replaced by urea in experiments in which FLAG-tagged BamA was folded by the reconstituted BAM complex[283]. Subsequent work established that each reconstituted BAM complex could catalyse the assembly of multiple substrates[301]. One key question in the evolution of the BAM complex is the 'chicken and egg' problem of how was the first BamA molecule folded, if it is needed to catalyse OMP folding, yet is itself an OMP? This question was resolved by the Kahne group, who reported that BamA, but not OmpA, can be assembled by BAM lipoproteins *in vitro*[283].

Recently it was shown that all five BAM complex members can be co-expressed and co-purified, either from multiple plasmids[357], or a single plasmid[239]. The reconstituted BamABCDE complex expressed from a single plasmid appeared to be more active than that reconstituted from BamAB and BamCDE subunits, and was able to assemble the autotransporter EspP[239]. Iadanza *et al*. (2016) modified the BAM reconstitution method, forming proteoliposomes by dialysis, rather than the dilution method used previously[232]. Using a modified version of the Kahne group OmpT assay this protocol led to substantially greater OmpT folding activity, likely due to greater efficiency of BAM reconstitution. This assay was then used to provide the first *in vitro* demonstration of the importance of BamA lateral gating in BAM function[232].

The presence of prefolded BamA alone, in the absence of BamB-E, was reported to increase the kinetics of OmpA folding into DLPC LUVs at 40 °C[299]. This was confirmed by experiments in which the presence of BamA in DDPC LUVs, containing a mole percentage of 20% DDPE (*di*$C_{10:0}$PE), dramatically increased (>10-fold) the kinetics of OmpX, tOmpA, OmpA and OmpLA folding[300]. This suggests that one function of BamA is to decrease the kinetic energy barrier imposed by native lipid headgroups[300]. It was later found that the mechanism by which BamA alone enhances folding rates is catalytic, rather than stoichiometric[302]. BamA is an unusual catalyst in that it has two substrates, protein and lipid, and in

these studies it is not clear if BamA-assisted OMP folding rate increases are due to its interactions with nascent OMPs, destabilisation/thinning of the bilayer, or both.

1.6 Current questions

This introduction has highlighted the progress made in the OMP folding field, both *in vivo* and *in* vitro, and the remaining challenges ahead. Many key questions remain unresolved including:

1. What are the molecular mechanism(s) by which periplasmic chaperones bind and transport OMPs, and how are their substrates delivered to the BAM complex?
2. What are the functions of the individual components of the BAM complex?
3. How do BAM components interact with each other, chaperones and OMP substrates?
4. What is the molecular mechanism by which the BAM complex catalyses β-barrel folding, insertion and release into the outer membrane?
5. How is OMP assembly coordinated with the biogenesis of other OM components (lipids, LPS, lipoproteins)?

It is also not known whether the BAM complex aids in the oligomerisation of OMP multimers, or how OMP assembly is coordinated with the biogenesis of other OM components (lipids, LPS, lipoproteins), or the molecular details of how the BAM complex is involved in extracellular secretion[9,111]. How similar the folding behaviour and pathways of OMPs observed *in vitro* are to the situation in a living cell is also unclear.

The experiments presented in the following chapters take a classical biochemical approach of purifying the components involved in a biological process and trying to recreate aspects of an *in vivo* process *in vitro* to aid its understanding. The results represent a fascinating journey into the world of OMP folding, which while providing some answers, in keeping with scientific tradition, provoke many more unanswered questions.

2 Materials and Methods

2.1 Materials and Reagents

2.1.1 General chemicals

Purite 18 MΩ water was used in all protocols. 30% (*w/v*) acrylamide:0.8% (*w/v*) bis-acrylamide and 10% (*w/v*) sodium dodecyl sulphate (SDS) was purchased from Severn Biotech Ltd., UK. Agar was purchased from Melford Laboratories, UK. Ready mixed LB-Broth (Miller) was purchased from Merck, Germany. Agarose, imidazole and ethylenediaminetetraacetic acid (EDTA) were purchased from Acros Organics, Belgium. Sodium chloride (NaCl), tris(hydroxymethyl)aminomethane (Tris), glycerol, glucose, sucrose, glacial acetic acid and hydrochloric acid (32% (*w/v*) HCl) were purchased from Fisher Scientific, UK. Carbenicillin disodium salt, dithiothreitol (DTT) and isopropyl β-D-1-thiogalactopyranoside (IPTG) were purchased from Formedium, UK. Triton X-100 (protein grade) was purchased from Merck Millipore, USA. Urea (>99% purity) was purchased from MP Biomedicals, UK or Sigma-Aldrich, USA. Ethidium bromide, magnesium sulphate ($MgSO_4$), magnesium chloride ($MgCl_2$), chloramphenicol, bromophenol blue, guanidine hydrochloride (GuHCl), ammonium persulphate (APS), ammonium acetate, ammonium hydroxide, tetramethylethylenediamine (TEMED), dimethyl sulfoxide (DMSO), tris(2-carboxyethyl)phosphine (TCEP), and ethanol were purchased from Sigma-Aldrich, USA.

2.1.2 Lipids

1,2-didecanoyl-*sn*-glycero-3-phosphocholine (*di*$C_{10:0}$PC) (DDPC), 1,2-diundecanoyl-*sn*-glycero-3-phosphocholine (*di*$C_{11:0}$PC) (DUPC), 1,2-dilauroyl-*sn*-glycero-3-phosphocholine (*di*$C_{12:0}$PC) (DLPC), 1,2-ditridecanoyl-*sn*-glycero-3-phosphocholine (*di*$C_{13:0}$PC) (DTPC), and 1,2-dimyristoyl-*sn*-glycero-3-phosphocholine (*di*$C_{14:0}$PC) (DMPC) lipids were obtained from Avanti Polar lipids (Alabaster, Alabama, USA). HLPC grade chloroform and methanol for lipid storage were purchased from Sigma-Aldrich, USA.

2.1.3 Molecular Biology Materials

E. coli strains BL21(DE3), BL21(DE3)pLysS, XL1-Blue and DH5α were purchased from Stratagene, UK. Competent *E. coli* DH5α cells were also purchased from New England Biolabs (NEB). OrangeG loading dye was purchased from Invitrogen. Vent DNA polymerase, Antarctic Phosphatase and all restriction enzymes were purchased from NEB. T4 DNA ligase (with 10x ligase buffer), 1 kbp and 100 bp DNA ladders, deoxyribonucleotide triphosphates (dNTPs) and 5-bromo-4-chloro-indoyl-β-D-galactopyranoside (X-Gal) were purchased from Promega UK. QIAprep Spin Miniprep and QIAquick gel extraction kits were purchased from Qiagen, UK. Site-directed mutagenesis was performed with the Q5 Site-directed mutagenesis kit (NEB), using primers designed using the online NEBasechanger utility[358]. The plasmids for tOmpA, PagP, OmpA and BamA encoding the mature OMP sequences were kindly provided by Dr Karen Fleming (John Hopkins University, USA). In these plasmids the OMP genes were ligated into pET11a between the NdeI (5') and BamHI (3') restriction sites[313]. Plasmid pET21b (Novagen) containing the Skp gene was a gift from Prof James Bardwell (University of Michigan, USA). Prof Sebastian Hiller (University of Basel, Switzerland), Prof Daniel Kahne (Harvard University, USA), and Dr Susan Buchanan (NIH, USA) kindly provided the HT-Skp-pET28b[171], HT-SurA-pet28b[233] and BamAB-pETDUET-1 plasmids, respectively.

2.1.4 Protein Chemistry Materials

The analytical Superdex S-75 HiLoad™ 26/60 column, Sephacryl S-200 HiLoad™ 16/60 column, Superdex Peptide 10/300 column, 5 ml HisTrap columns, 5 ml HiTrap Q columns and 5 ml HiTrap SP columns were purchased from GE Healthcare, UK. All buffers used during protein purification were filtered before use by vacuum filtration through 0.45 μM filters purchased from Millipore, UK. Small volumes were filtered through 0.2 μM or 0.45 μM Minisart syringe filters purchased from Sartorius, UK. Vivaspin 20 concentrators (MWCO 5 kDa or 10 kDa) were purchased from Sartorius, UK. SnakeSkin dialysis tubing (3.5 kDa MWCO), bicinchoninic acid (BCA) assay kits, and Alexa Fluor 488 C5 maleimide were purchased from Thermo Fisher Scientic, UK.

2.2 Molecular Biology

2.2.1 *E. coli* Bacterial Strains

DH5α:

fhuA2 lac(del)U169 phoA glnV44 Φ80' lacZ(del)M15 gyrA96 recA1 relA1 endA1 thi-1 hsdR17

XL1-Blue:

recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac [F' *proAB lacI*^{q} *Z*ΔM15 *Tn10* (*Tet*^{r})]

BL21(DE3):

F– *omp*T hsdS$_B$ (r_B–, m_B–) gal [dcm] [*lon*] (DE3)

BL21(DE3) pLysS:

F– *omp*T hsdS$_B$ (r_B–, m_B–) gal [dcm] [*lon*] (DE3)

2.2.2 Growth media

E. coli cells were grown in lysogengy broth (LB) (Miller) medium (Merck, Germany) at a concentration of 25 g/L. The medium was autoclaved for 20 min at 120 °C and supplemented with carbenicillin (100 mg/ml in sterile H_2O) to a final concentration of 100 μg/ml when cooled to room temperature. For solid agar plates 25 g/L of LB and 15 g/L of agar were autoclaved for 20 minutes at 120 °C. On cooling to approximately 50 °C, carbenicillin (100 μg/ml) was added and 20-25 ml poured into sterile Petri dishes. For growth of BL21(DE3)pLysS cells, chloramphenicol (34 μg/ml, from a 34 mg/ml stock dissolved in ethanol) was added to the medium for the maintenance of the pLysS plasmid in addition to carbenicillin (100 μg/ml). All antibiotics were filter sterilised prior to use. For cells transformed with pET28b vectors, 100 μg/ml carbenicillin was replaced with 30 μg/ml kanamycin in the above protocols.

For transformation of Supercompetent XL1-Blue cells NZY+ medium was used. NZY+ broth was prepared as follows:

- 10 g NZ amine
- 5 g yeast extract

- 5 g NaCl

A solution contained the above was made up to a final volume of 1 litre with dH_2O and adjusted to pH 7.5 using NaOH. The solution was autoclaved and the following filter sterilised supplements were added:

- 12.5 ml 1 M $MgCl_2$
- 12.5 ml 1 M $MgSO_4$
- 20 ml 20 % (w/v) glucose

2.2.3 Agarose gel electrophoresis

Agarose gels were prepared by heating 150 ml of 1.5% (*w/v*) agarose in 1x TAE buffer in a microwave until all agarose was dissolved. TAE buffer was prepared as a 50x stock, containing 121 g Tris, 28.55 ml glacial acetic acid, 50 ml 0.5 M EDTA in 500 ml (pH 8.0). The solution was left to cool to ~50 ºC before addition of 10 µl per 100 ml of a 10 mg/ml solution of ethidium bromide. Gels were then poured into a 12 x 15 cm gel tray containing a lane comb and allowed to set. Agarose gels were electrophoresed in 1x TAE buffer. DNA samples were mixed 9:1 with a 10x OrangeG loading buffer (50 mM Tris, 60% (*v/v*) glycerol, OrangeG dye powder). Samples containing 1 kb and 100 bp DNA ladders were included to enable size comparison. Gels were run at 100 volts for approximately 1 hour or until sufficient size resolution had been achieved. Gels were imaged using a UV transilluminator (Syngene).

2.2.4 Preparation of competent cells

Cells of the desired strain were streaked onto an LB agar plate containing no antibiotics using a sterile loop and incubated overnight at 37 °C. A single colony was selected and used to inoculate 5 ml of autoclaved LB and incubated overnight at 37 °C with shaking at 200 rpm. This culture was used to inoculate 100 ml of sterile LB which was grown (37 °C, 200 rpm) until the optical density (OD) at 600 nm was 0.4-0.45. Cells were harvested by centrifugation (4000 rpm (4000 *g*), 10 min, 4 °C) in a pre-chilled Beckman JS-5.3 rotor. The pellet was resuspended gently in 10 ml of filter-sterilised, pre-chilled 100 mM $CaCl_2$. After incubation on ice for 10 minutes the cells were centrifuged as described previously and the supernatant removed. The pellet was resuspended gently in 2 ml of pre-chilled 100 mM $CaCl_2$ 30% (*v/v*)

glycerol. Cells were aliquoted (50 μl) into pre-cooled 1.5 ml Eppendorf tubes on dry ice and stored at -80 °C.

2.2.5 Transformation of *E. coli* strains

For transformation from plasmid stocks, 50 μl of laboratory-prepared competent cells were thawed on ice for 10 min then transferred to a pre-chilled round bottomed 14 ml transformation tube. 1-2 μl vector DNA (100-200 ng/μl) was added and the solution gently mixed. After incubation on ice for 30 min, the tubes were heat shocked at 42 °C for 45 sec and returned to ice for 5 min. All 50 μl was plated directly onto LB agar containing 100 μg/ml carbenicillin (or 30 μg/ml kanamycin for cells transformed with pET28b vectors), and incubated at 37 °C overnight. For transformation of ligation reactions, XL-1 Blue Supercompetent cells (Stratagene, UK) or competent DH5α cells (NEB) were used as per manufacturer's instructions. 5 μl of ligation reaction mixture was added per 100 μl or 50 μl of XL-1 Blue Supercompetent, or competent DH5α cells, respectively.

2.2.6 Polymerase Chain Reaction (PCR)

Vent DNA polymerase was used as per manufacturer's instructions. A typical 100 μl PCR contained the following:

- 100 ng dsDNA template
- 0.5 μM of each primer
- 750 μM of each dNTP
- 1 x Thermopolymerase buffer (NEB)
- 2 mM $MgSO_4$
- 1 U Vent DNA polymerase (NEB)
- 1 μL DMSO

The following PCR primers used for the cloning of the barrel domain of BamA (tBamA):

Forward primer: 5'-CACACACAC CATATG agc ttc aac ttt ggt att ggt tac gg – 3'

Reverse primer: 5' – GTGTGTGT GGATCC tta tta cca ggt ttt acc gat gtt aaa ctg g - 3'

NdeI and BamHI restriction sites are highlighted in green and blue text, respectively. Stop codons are highlighted in red text. Nucleotides at the 3' which bind to the template DNA are indicated with uncapitalised text. The PCR cycling conditions used for amplification of tBamA are given in Table 2.1.

Step	Time	Temperature
1	5 minutes	95 ºC
2	30 seconds	95 ºC (melting)
3	30 seconds	54 ºC (annealing)
4	90 seconds	72 ºC (extension)
5	Repeat steps 2-4, 39 times	
6	5 minutes	72 ºC

Table 2.1 Typical cycling times for Vent DNA polymerase PCRs.

2.2.7 Agarose gel DNA extraction

To extract DNA bands after agarose gel electrophoresis the QIAquick gel extraction kit (QIAGEN, UK) was used as per manufacturer's instructions.

2.2.8 Plasmid preparation

For small-scale DNA preparation, 10 ml of LB/carbenicillin was inoculated with a single colony and incubated overnight at 37 °C with shaking at 200 rpm. After obtaining a cell pellet by centrifugation, the plasmid DNA was purified using the QIAprep Spin Miniprep Kit (QIAGEN,UK) or the Wizard Plus SV Miniprep Kit (Promega, UK), both as per manufacturer's instructions. To obtain larger amounts of DNA, 100 ml of LB/carbenicillin was inoculated with a single colony, grown overnight, and after obtaining a cell pellet by centrifugation, the plasmid DNA was purified using the HiSpeed Plasmid Midi Kit (QIAGEN,UK) as per manufacturer's instructions. The 1 ml DNA eluate was concentrated to a volume of 100 μl using a QIAquick spin column (QIAGEN,UK). DNA concentrations were obtained using a Nanodrop 2000 (Thermo Fisher Scientific, UK) as per manufacturers' instructions.

2.2.9 Restriction Digests

The restriction endonucleases NdeI and BamHI were purchased from New England Biolabs (NEB) and used as per manufacturer's instructions. Analytical digests of plasmed minipreps derived from colonies on transformation plates were incubated at 37 °C for 30 min (Table 2.2). Restriction digests of larger volumes of DNA (e.g. pET11a vector DNA (5 μg), or gel extracted PCR products) for use in downstream ligation reactions were incubated at 37 °C for 4-5 h (Table 2.3).

Plasmid DNA (μl)	NEB Buffer 4 (μl)	10x BSA (μl)	Sterile water (μl)	NdeI (μl)	BamHI (μl)
5	2	2	10	0.5	0.5

Table 2.2: **Volumes of reagents for analytical digests of miniprepped DNA from colonies on transformation plates.**

Plasmid DNA (μl)	NEB Buffer 4 (μl)	10x BSA (μl)	Sterile water (μl)	NdeI (μl)	BamHI (μl)
20	3	3	1	1.5	1.5

Table 2.3: **Typical volumes of reagents for digests of vector DNA, or PCR products for use in downstream ligation reactions.**

2.2.10 Plasmid vector descriptions

Details of all plasmids used in this project are given in Table 2.4. DNA sequences of all vectors are available on request from the Radford Lab, University of Leeds.

Plasmid name and vector	Description of insert sequence
BamAB-pETDUET-1	Full-length BamA and BamB genes including signal sequences. Used as DNA template for tBamA and BamA POTRAs cloning
P1-5-pET11a	BamA POTRA domains 1-5 of mature BamA (residues 21-425)*
HT-P1-5-pET11a	BamA POTRA domains 1-5 of mature BamA (residues 21-425) with hexahistidine tag at the N-terminus*
HT-MBP-TEV-POTRAs-pMAL	Maltose-bindng protein (MBP) with an N-terminal 6x His-tag fused to BamA POTRA domains 1-5 of mature BamA (residues 21-425), separated by a TEV cleavage site for removal of HT-MBP*
PagP-pET11a	Mature PagP (residues 26-186)[313]
HT-OmpT-pET11a	Mature OmpT (residues 21-317) preceded by an N-terminal 6x His-tag and TEV protease cleavage site (MH_6ENLYFQG-OmpT)[359]
OmpF-pET11a	Mature OmpF (residues 23-362)*
FadL-PEX-K4	Mature FadL (residues 26-446) gene synthesised by Eurofins Genomics, and delivered in PEX-K4 plasmid[360]
FadL-pET11a	Mature FadL (residues 26-446)*
BamA-pET11a	Mature full-length BamA (residues 21-810)[313]
tBamA-pET11a	C-terminal beta-barrel domain of mature BamA (residues 425-810)*
OmpA-pET11a	Mature full-length OmpA (residues 22-346)[313]
tOmpA-pET11a	N-terminal 171 residues mature OmpA (residues 22-193). I.e. OmpA with the C-terminal soluble domain deleted[313].
Cys-tOmpA-pET11a	N-terminal 171 residues mature OmpA (residues 22-193) with an additional N-terminal cysteine residue*
Skp-pET21b	Full length Skp gene (residues 1-161) including signal sequence for periplasmic expression
HT-Skp-pET28b	Mature Skp (residues 21-161) preceded by an N-terminal 6x His-tag and thrombin protease cleavage site (MGSSH_6SSGLVPRGSH-Skp)[171]
HT-SurA-pET28b	Mature SurA (residues 21-428) preceded by an N-terminal 6x His-tag and thrombin protease cleavage site (MGSSH_6SSGLVPRGSH-SurA)[233]

Table 2.4: Description of all vectors used in this thesis. All gene sequences are from *E. coli*. *Vector was produced in the current project using NdeI and BamHI restriction enzymes.

2.2.11 DNA sequencing

All DNA sequencing was carried out by Beckman Coulter Genomics. The T7 promoter forward primer (TAA-TAC-GAC-TCA-CTA-TAG-GG) and the T7 terminator reverse primer (CTA-GTT-ATT-GCT-CAG-CGG-TG) were used for sequencing of all pET vectors.

2.2.12 DNA ligations

Ligations of DNA inserts into pET11a were carried out using T4 DNA ligase in 1x ligation buffer with overnight incubation at 4 °C. A typical ligation reaction contained the following:

- Digested pET11a 1 µl
- Digested insert DNA 7 µl
- 10x T4 DNA ligase buffer 1 µl
- T4 DNA ligase (3 U/µl) 1 µl

Prior to ligation, digested pET vectors were dephosphorylated at the 5' ends using Antarctic phosphatase as per manufacturer's guidelines, to prevent vector-only re-ligation. Digested insert DNA and digested vector DNA were purified by agarose gel electrophoresis followed by gel extraction prior to the ligation reaction, to remove restriction endonuclease and Antarctic phosphatase reagents. 5 µl of ligation reaction mixture was used for transformation into XL1-Blue Supercompetent cells (Section 2.2.5).

2.2.13 Cloning of OmpF, tBamA and OmpT

Codon optimised synthetic genes (Eurofins, Germany) of the mature sequences of OmpF (residues 23-362) and FadL (residues 26-446) were cloned into pET11a (Novagen, UK) between the NdeI (5′) and BamHI (3′) restriction sites. To create the tBamA construct, residues 425-810 of BamA were amplified by PCR, using plasmid BamAB-pETDUET-1 as the template DNA, and the resultant product then ligated into pET11a after restriction digestion, as described above. The OmpT mature sequence (residues 21-317) was amplified by PCR from *E. coli* XL1-blue cells to include an N-terminal 6x His-tag and a Tobacco Etch Virus (TEV) protease cleavage site (MH_6ENLYFQG-OmpT), and subsequently cloned into the pET11a vector as

described above. The cloning of OmpT and OmpF were performed by Dr L. M. McMorran (University of Leeds) and Dr A. N. Calabrese (University of Leeds), respectively.

2.3 Outer membrane protein expression and purification

2.3.1 Expression and purification of untagged outer membrane proteins

The relevant plasmid was transformed into BL21(DE3) cells and 5 ml LB medium containing 100 μg/ml carbenicillin was inoculated with a single colony from the transformation plate and incubated at 37 °C with shaking (200 rpm) overnight. This starter culture was added to 500 ml LB medium containing 100 μg/ml carbenicillin and grown at 37 °C with shaking (200 rpm). Expression was induced with 1 mM IPTG when the culture reached an OD_{600} of 0.5-0.6, and then harvested after 4h by centrifugation (5000 *g*, 15 min, 4 °C). Cells were resuspended in 20 ml 50 mM Tris-HCl, pH 8.0, 5 mM EDTA, 1 mM phenylmethylsulfonyl fluoride (PMSF), 2 mM benzamidine, and lysed by sonication (6 x 1 min bursts with 1 min cooling on ice between each sonication). The insoluble fraction was collected by centrifugation (25000 *g*, 30 min, 4 °C), resuspended in 20 ml 50 mM Tris-HCl, pH 8.0, 2% (*v/v*) Triton-X-100 and incubated for 1 h at room temperature, with gentle agitation. The insoluble fraction was again pelleted (25000 *g*, 30 min, 4 °C) and the inclusion bodies were washed twice by resuspending in 50 mM Tris-HCl, pH 8.0, incubating for 1 h at room temperature with gentle agitation, followed by centrifugation (25000 *g*, 30 min, 4 °C). The inclusion bodies were solubilised in 25 mM Tris-HCl, 6 M GuHCl, pH 8.0 and centrifuged (20000 *g*, 20 min, 4 ºC)**.** The supernatant was filtered (0.2 μM syringe filter, Sartorius, UK) and purified further by gel filtration in 6 M GuHCl, 25 mM Tris-HCl, pH 8.0 using a Superdex 75 HiLoad™ 26/60 column (or Sephacryl S-200 HiLoad™ 16/60 column for preparations of BamA) connected to the ÄKTA Prime chromatography system. The column was washed with two column volumes of dH_2O and equilibrated with two column volumes of 6 M GuHCl, 25 mM Tris-HCl, pH 8.0. For each gel filtration run, ~4 ml of ~5 mg/ml of OMP was loaded into a 5 ml loop. An example programme used is shown in Table 2.5.

Breakpoint (ml)	Flow rate (ml/min)	Fraction size (ml)	Injection valve position	Autozero
0	2	0	Load	No
10	2	0	Inject	Yes
20	2	0	Load	No
105	2	3	Load	No
355	2	0	Load	No

Table 2.5: AKTA Programme parameters for gel filtration of OMPs. Gel filtration was carried out using a Superdex 75 HiLoad 26/60 column in 6 M GuHCl, 25 mM Tris-HCl, pH 8.0.

Samples from peak fractions were buffer exchanged into 8 M urea using Zeba spin desalting columns (Thermo Scientific, UK), or subjected to TCA precipitation prior to analysis by SDS-PAGE to assess purity. Peak fractions were pooled, concentrated to ~500 µM using Vivaspin 20 (5 kDa MWCO) concentrators (Sartorius, UK), snap-frozen in liquid nitrogen, and stored at -80 °C.

2.3.2 Expression and purification of His-tagged OmpT

Expression and inclusion body purification for His-tagged OmpT was carried out as for untagged OMPs. OmpT inclusion bodies were then solubilised in 50 mM sodium phosphate, 6 M GuHCl, 250 mM NaCl, pH 8.0, and then centrifuged (25000 *g*, 20 min, 4 ºC). The supernatant was filtered (0.2 µM syringe filter, Sartorius, UK) and loaded on to a 5 ml HisTrap column (GE Healthcare, UK) equilibrated with 50 mM sodium phosphate, 6 M GuHCl, 250 mM NaCl, pH 8.0. The column was washed with 10 column volumes of 50 mM sodium phosphate, 6 M GuHCl, 250 mM NaCl, 5 mM imidazole, pH 8.0. OmpT was eluted with 50 mM sodium phosphate, 6 M GuHCl, 250 mM NaCl, 250 mM imidazole, pH 8.0, and peak fractions were pooled and dialysed overnight against H_2O at 4 ºC. The resultant precipitated protein was collected by centrifugation (25000 *g*, 20 min, 4°C), solubilised in 25 mM Tris-HCl, 6 M GuHCl, pH 8.0 (~500 µM), snap-frozen in liquid nitrogen, and stored at -80 °C. Expression and purification of His-tagged OmpT was performed by T.G. Watkinson.

2.4 Expression and purification of BamA POTRA domains

A pMAL-c5X (NEB) plasmid was modified previously by addition of an N-terminal 6x His-tag and replacement of the thrombin cleavage site with a TEV cleavage site. This vector was used to make a construct that contained maltose-binding protein (MBP) fused to BamA POTRA domains 1-5 of mature BamA (residues 21-425) with an additional C-terminal cysteine residue, separated by a TEV cleavage site. This vector was produced previously by T. Crosskey (University of Leeds)[361]. The C-terminal cysteine residue was removed by site-directed mutagenesis to produce the HT-MBP-TEV-POTRAs construct (Figure 2.1).

Figure 2.1: Schematic of HT-MBP-TEV-POTRAs construct. HT: His-tag, TEV: TEV cleavage site.

HT-MBP-TEV-POTRAs-pMAL was transformed into BL21(DE3) cells, and a single colony used to inoculate an LB starter culture containing 100 μg/ml carbenicillin which was grown overnight (37 ºC, 200 rpm). 10 ml starter culture was used to inoculate 2x 1 L of LB, supplemented with 100 μg/ml carbenicillin. Cells were grown at 37 °C with shaking (200 rpm) until the culture reached an OD_{600} of ~0.6, at which point expression was induced with 1 mM IPTG. After 4 h expression cells were harvested by centrifugation (5000 *g*, 15 min, 4 °C). Cells were resuspended in 20 mM Tris-HCl, pH 8.0, 300 mM NaCl, 20 mM imidazole, EDTA-free protease inhibitor tablets (Pierce Biotechnology, USA), and lysed using a cell disrupter (Constant Cell Disruption Systems, UK). Following centrifugation to remove cell debris (20 min, 4 °C, 39000 *g*), the lysate was applied to 5 ml HisTrap columns and washed with 20 mM Tris-HCl, pH 8.0, 300 mM NaCl, 20 mM imidazole. HT-MBP-TEV-POTRAs was eluted with a 0-500 mM imidazole gradient with 20 mM Tris-HCl, pH 8.0, 300 mM NaCl, 20 mM imidazole. Fractions containing HT-MBP-TEV-POTRAs were pooled, then dialysed to into 20 mM Tris-HCl, pH 8.0, 300 mM NaCl. To prevent aggregation following TEV cleavage the protein was diluted to a volume of ~100 ml with 20 mM Tris-HCl, pH 8.0, 300 mM NaCl. Following the addition of 14.4 mM β-mercaptoethanol (BME), His-tagged TEV protease (provided by Dr D. Walsh (University of Leeds)) was added at a 1:10 TEV: HT-MBP-TEV-POTRAs molar ratio and incubated overnight at 4 ºC. The cleavage reaction was passed over 5 ml

HisTrap columns and the flow through, containing cleaved BamA POTRAs was collected. BamA POTRAs were concentrated to ~100 μM using Vivaspin 20 (5 kDa MWCO) concentrators (Sartorius, UK), aliquoted, snap-frozen in liquid nitrogen and stored at -80 °C. The expression and purification of BamA POTRAs was performed by Dr A. N. Calabrese (University of Leeds).

2.5 Expression and purification of Skp

The pET21b plasmid, containing the full-length Skp gene, including the N-terminal signal sequence, was transformed into BL21[DE3] cells (Stratagene, UK). After inoculation with a 10 ml starter culture (LB, supplemented with100 μg/ml carbenicillin), cells were grown in 12 x 1 L LB medium containing 100 μg/ml carbenicillin at 27 °C with shaking (200 rpm) until the culture reached an OD_{600} of ~0.6 (after ~6 h). Cultures were induced with 25 μM IPTG, expressed overnight, and harvested by continuous centrifugation at 15000 rpm (Heraeus Contifuge, Rotor 8575, Thermo Fisher Scientific, UK). The cell pellet was gently resuspended in 50 mM Tris-HCl, 5 mM EDTA, 50 mM NaCl, pH 7.5 at 4 °C with gentle agitation for ~48 h. 1 mg/ml polymyxin B sulphate was added to the resuspended cells, and then incubated for 1 h at 4 °C with gentle agitation. Spheroplasts were sedimented by centrifugation (12 000 *g*, 20 min 4 °C) and the supernatant dialysed against 20 mM Tris-HCl, 100 mM NaCl, pH 8.0 (Buffer A) overnight at 4 °C. The periplasmic extract was filtered (0.4 μm syringe filter, Sartorius, UK) and loaded onto a HiTrapQ (5 ml) anion exchange column (GE Healthcare, UK). The flow through from this column was loaded onto a HiTrap SP (5 ml) cation exchange column (GE Healthcare, UK) equilibrated with Buffer A. The column was washed with 5 column volumes of Buffer A and eluted with a gradient (0-100 %) of 20 mM Tris-HCl, 750 mM NaCl, pH 8.0 (Buffer B). Peak fractions were dialysed against 20 mM Tris-HCl, pH 8.0 and concentrated to ~50 μM (trimer) using Vivaspin 20 (5 kDa MWCO) concentrators (Sartorius, UK). Aliquots were snap-frozen in liquid nitrogen and stored at -80 °C.

2.6 Expression and purification of His-tagged Skp

His-tagged Skp was expressed and purified using a protocol adapted from Burmann *et al*.[171]. The pET28b plasmid, containing the Skp gene with an N-terminal 6xHis-tag and thrombin cleavage site, was transformed into BL21[DE3]pLysS cells (Stratagene, UK). Following inoculation with a 10 ml starter culture (LB, supplemented with 30 μg/ml kanamycin), cells were grown in 2x 1L LB medium

containing 30 µg/ml kanamycin at 37 °C with shaking (200 rpm) until the culture reached an OD_{600} of ~0.6. The temperature was then lowered to 20 °C and expression induced with 0.4 mM IPTG. Following overnight expression (~18 h) cells were harvested by centrifugation, resuspended in 25 mM Tris-HCl, pH 7.2, 150 mM NaCl, 20 mM imidazole, containing a cocktail of EDTA-free protease inhibitors (Roche), and lysed using a cell disrupter (Constant Cell Disruption Systems, UK). Following centrifugation to remove cell debris (20 min, 4 °C, 39000 *g*), the lysate was applied to 5 ml HisTrap columns (GE Healthcare) and washed with 25 mM Tris-HCl, pH 7.2, 150 mM NaCl, 20 mM imidazole. His-tagged Skp was denatured on-column with 25 mM Tris-HCl, 6 M GuHCl, pH 7.2, and eluted with a 0-500 mM imidazole gradient over 50 ml in 25 mM Tris-HCl, 6 M GuHCl, pH 7.2. Fractions containing Skp were pooled and the protein refolded by dialysis against 25 mM Tris-HCl, pH 7.2, 150 mM NaCl. Refolded His-tagged Skp was concentrated to ~50 µM (trimer) using Vivaspin 20 (5 kDa MWCO) concentrators (Sartorius, UK), aliquoted, snap-frozen in liquid nitrogen and stored at -80 °C.

2.7 Expression and purification of His-tagged SurA

His-tagged SurA was expressed and purified with the same protocol as for His-tagged Skp (Section 2.6), and stored at a concentration of ~200 µM.

2.8 Protein expression yields

Expression yields for outer membrane proteins were >50 mg/L of bacterial culture. Yields for purified His-tagged Skp, His-tagged SurA and BamA POTRA domains were also >50 mg/L. Yields for Skp expressed and purified from the periplasm were ~5 mg/L.

2.9 General protein methods

2.9.1 Sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE)

Tris-tricine buffered SDS-PAGE gels (Table 2.6). were used to monitor protein purification and to visualise the ratio of folded to unfolded OMPs in band shift assays. Protein samples were diluted two-fold in 2x loading buffer (50 mM Tris-HCl pH 6.8, 2% (*w/v*) SDS, 0.1% (*w/v*) bromophenol blue, 10% (*v/v*) glycerol). All samples

except those used in cold SDS-PAGE band shift assays were boiled for 5 min prior to loading. Precision Plus Protein™ Dual Xtra Standards (Bio-Rad, UK) were loaded into one lane to aid size determination and identification of protein bands. Gels were electrophoresed with the inner reservoir of the gel tank containing cathode buffer (100 mM Tris, 100 mM tricine, 0.1% (*w/v*) SDS, pH 8.25) and the outer reservoir containing anode buffer (200 mM Tris-HCl, pH 8.9). A current of ~30 mA was applied until the samples entered the resolving gel, at which point the current was adjusted to 60 mA. Semi-native SDS-PAGE gels were made as detailed in Table 2.6, except without the addition of SDS. Semi-native SDS-PAGE 6x loading buffer contained 50 mM Tris-HCl pH 6.8, 0.1% (*w/v*) SDS, 0.1% (*w/v*) bromophenol blue, 30% (*v/v*) glycerol). Semi-native SDS-PAGE gels were run in a cold cabinet at 4 ºC for ~12 h at 14 mA to avoid denaturation of the BamA barrel[129]. All gels were stained using Instant Blue stain (Expedeon, UK) and imaged using the Syngene InGenius gel documentation system (Syngene, UK). Densitometry quantification of folded and unfolded OMP bands in cold SDS-PAGE assays was performed using ImageJ[362].

Reagent	Resolving gel volumes (ml)	Stacking gel volumes (ml)
30% (*w/v*) acrylamide:0.8% (*w/v*) *bis*-acrylamide	7.5	0.83
3 M Tris, 0.3% (*w/v*) SDS (pH 8.45)	5.0	1.55
H_2O	0.44	3.72
Glycerol	2.0	-
10% (*w/v*) ammonium persulphate	0.1	0.2
N,N,N',N'-tetramethyl-ethylenediamine (TEMED)	0.01	0.01

Table 2.6: Volumes of reagents for Tris-tricine buffered SDS-PAGE gels. The volumes indicated are sufficient for casting two 8 cm by 10 cm mini-gels using a 1.5 mm spacer.

2.9.2 Trichloroacetic acid (TCA) precipitation of proteins

TCA precipitation of purified outer membrane proteins was used to remove GuHCl and concentrate the sample prior to SDS-PAGE analysis. The procedure was carried out in a cold room at 4 °C. Equal volumes of 12.5 % (*w/v*) TCA and sample solution were vortexed and incubated on ice for 20 min. Solutions were then centrifuged (13000 rpm, 10 min, 4 °C) in a microcentrifuge and the supernatant removed. 1 ml of ice cold ethanol was added to wash the precipitated protein and the centrifugation step repeated. The supernatant was again discarded and the remaining ethanol removed by placing tubes in a 100 °C heat block for approximately 5 min. The pellets were then resuspended in SDS loading buffer and boiled for 10 min prior to SDS-PAGE analysis.

2.9.3 Determination of protein concentration

The protein concentrations of all OMPs, BamA POTRA domains, and SurA were determined spectrophotometrically using their absorbance at 280 nm. Theoretical molar extinction coefficients at 280 nm, calculated using the ExPASy Protparam server[363], are shown in Table 2.7.

Skp has a very low molar extinction coefficient at 280 nm (1490 M^{-1} cm^{-1}) as its sequence contains no tryptophan residues, and only one tyrosine residue. Therefore, for increased accuracy[364], Skp concentrations were determined using a bicinchoninic acid (BCA) assay (Thermo Fisher Scientific, UK), according to the manufacturer's instructions.

Protein	Molar extinction coefficients at 280 nm (M^{-1} cm^{-1})
PagP	82,390
tOmpA	46,870
Cys-tOmpA	46,870
OmpA	52,955
OmpT	79,760
OmpF	54,210
FadL	94,770
tBamA	101,315
BamA	140,165
BamA POTRAs	38,850
SurA	29,450

Table 2.7: Molar extinction coefficients at 280 nm of purified proteins.

2.10 Characterisation of PagP in DLPC liposomes (Chapter 3)

2.10.1 Liposome preparation

DLPC was dissolved in a 90:10 (*v/v*) chloroform:methanol solution in glass test tubes. Solvent was then removed by drying under a gentle stream of N_2, followed by further drying in a desiccator under high vacuum for >3 h. The resulting thin lipid film was resuspended at a concentration of 40 mM in 50 mM sodium phosphate, pH 8.0 and left to stand at room temperature for 30 min. 100 nm LUVs were prepared by extruding the lipid suspension 11 times through 0.1 μm polycarbonate membranes

(Nucleopore, Whatman, Clifton, NJ) using a mini-extruder (Avanti, Alabaster, AL, USA), and used within 72 h.

2.10.2 SDS-PAGE band shift assays

100 µl samples were made containing 16 mM (total lipid) DLPC liposomes, 5 µM PagP, between 1 M and 8 M urea in 1 M increments in 50 mM sodium phosphate pH 8.0. The final LPR was 3200:1. Samples were allowed to equilibrate overnight at 25 ºC. 2x SDS loading buffer was added immediately prior to gel loading, with the exception of a boiled sample (in 1 M urea) which was boiled in 2x SDS loading buffer for 10 min prior to loading.

2.10.3 Fluorescence emission spectra

Fluorescence emission spectra were measured using a Photon Technology International Fluorimeter (Ford, West Sussex, UK). Samples contained 0.4 µM PagP, 12.8 mM DLPC (LPR 3200:1), 7 M or 10 M urea for folded and unfolded samples, respectively, in 50 mM sodium phosphate, pH 8.0. Samples were incubated at 25 °C overnight prior to measurement. Both excitation and emission slit widths were set to 3 nm. Each spectrum was recorded from 295 nm to 400 nm in 1 nm increments, using an excitation wavelength of 280 nm.

2.10.4 CD spectra

Far-UV CD spectra of PagP were acquired on a Chirascan plus circular dichroism spectrometer (Applied PhotoPhysics) with a bandwidth of 1 nm, a scan speed of 20 nm min^{-1}, a step size of 1 nm and a pathlength of 0.1 mm. The average of eight scans was taken to enhance signal to noise. PagP samples contained 10 µM PagP in 50 mM sodium phosphate, pH 8.0. An LPR of 800:1 was used to reduce light scattering. Samples were refolded for 16 h before measurement. Corresponding blank spectra, containing all reagents except for PagP, were subtracted for each sample. The Mean Residue Ellipticity (MRE) at each wavelength was obtained by first calculating the Mean Residue Weight (MRW):

$$MRW = \frac{M}{N-1}$$

where M is the molecular mass of the protein in Daltons, and N is the number of amino acids it contains. The MRE is then given by:

$$[\theta]_{MRE} = \frac{MRW \times \theta_\lambda}{10 \times d \times c}$$

where $[\theta]_{MRE}$ is the MRE, θ_λ is the measured ellipticity at a particular wavelength, d is the pathlength in cm, and c is the concentration in g/ml.

2.11 Preparation of liposomes

DUPC, DLPC, DTPC and DMPC lipids were obtained as a powder, dissolved in a 80:20 (*v:v*) chloroform:methanol mixture at 25 mg/ml, and stored at -20 ºC until use. Appropriate volumes were transferred to glass test tubes and an even lipid film was created by drying with a gentle stream of nitrogen while being shaken moderately in a 42 ºC water bath. Lipid films were further dried in a vacuum desiccator for >3 h, followed by resuspension in 50 mM glycine-NaOH, pH 9.5 to a concentration of 40 mM. Resuspended lipids were vortexed briefly and allowed to stand for 30 min. After vortexing again, lipids were subjected to 5 freeze-thaw cycles using liquid nitrogen. Large Unilamellar Vesicles (LUVs) (100 nm) were prepared by extruding the lipid suspension a minimum of 11 times through a 0.1 μm polycarbonate membrane (Nuclepore, New Jersey, USA) using a mini extruder (Avanti Polar Lipids, Alabama, USA). In the case of DMPC, the mini extruder was pre-warmed and the extrusion performed at 37 ºC to be well above the transition temperate for DMPC (24 ºC)[365]. Liposomes were stored at 4 ºC. DUPC liposomes were used within 48 h of preparation. DLPC, DTPC and DMPC liposomes were used within 72 h.

2.12 CD spectroscopy of folding factors

Far-UV CD spectra were acquired on a Chirascan plus circular dichroism spectrometer (Applied PhotoPhysics) with a bandwidth of 2.5 nm, a scan speed of 0.5 nm s^{-1}, a step size of 1 nm and a pathlength of 1.0 mm. CD of soluble proteins (Skp, SurA and BamA POTRA domains) were acquired in 0.24 M urea, 50 mM glycine-NaOH, pH 9.5, at 25 ºC, using a protein concentrations of ~0.2 mg/ml. Samples for CD of BamA in LUVs composed of DLPC, DTPC or DMPC, contained 1.5 μM BamA, 1.2 mM lipids (LPR 800:1), 0.24 M urea, 50 mM glycine-NaOH, pH

9.5, and were measured at 25 ºC. BamA-containing samples were folded overnight at 25 ºC. An LPR of 800:1 was used to reduce light scattering.

To enhance signal to noise the average of three or eight scans was taken for soluble folding factors or BamA, respectively. Corresponding blank spectra, containing all reagents except the protein components, were taken and subtracted. Mean residue ellipticities (MREs) were calculated as in Section 2.10.4.

2.13 BamA band shift assays in DLPC, DTPC and DMPC

BamA folding efficiency in DLPC, DTPC and DMPC liposomes was assessed by semi-native SDS-PAGE band shift assays[129] (Section 2.9.1). Samples contained 0.8 μM BamA 1.28 mM lipids (LPR 1600:1), 0.24 M urea, 50 mM glycine-NaOH, pH 9.5, and were folded overnight at 25 ºC prior to analysis.

2.14 Kinetic folding assays

Kinetic measurements were carried out using a Quantum Master Fluorimeter (Photon Technology International, UK) controlled by FelixGX software v4.3. For each experiment, four separate samples were run in a four cell changer maintained at 25 ºC by a peltier-controlled temperature unit. Tryptophan fluorescence of samples was excited at a wavelength of 295 nm, and fluorescence emission was monitored at 335 nm. 295 nm rather than 280 nm was chosen as the excitation wavelength as this minimises the fluorescence intensity contribution from Skp, which contains a single tyrosine residue, but no tryptophan residues. The excitation slit widths were set to 0.3-0.6 nm and the emission slit widths were set to 5 nm. The high emission:excitation slit width ratio was important to minimise photobleaching on the experimental timescale. Folding was initiated by dilution of denaturant and manual mixing in quartz cuvettes (10 x 4 mm, Hellma Analytics, UK). Cuvettes were shaken ~three times following OMP dilution before insertion into the fluorimeter.

2.14.1 Kinetic assays in DUPC liposomes (Chapters 4 and 5)

OMPs were buffer exchanged from 25 mM Tris-HCl, 6 M GuHCl, pH 8.0 into 50 mM glycine-NaOH, 8 M urea, pH 9.5 using Zeba spin desalting columns (Thermo Scientific, UK) and diluted to 80 μM. OMP folding reactions in the absence of additional folding factors were initiated by diluting OMPs manually from this 80 μM

unfolded protein stock in 8 M urea to a final concentration of 0.4 μM protein and 0.24 M urea in the presence of 1.28 mM DUPC liposomes (a lipid:protein molar ratio (LPR) of 3200:1), in 50 mM glycine-NaOH, pH 9.5. The final volume for each sample was 500 μl. For experiments in which tOmpA was folded into DUPC liposomes in which BamA, tBamA, or OmpA were prefolded, 0.8 μM of BamA, tBamA, or OmpA were folded by dilution into liposomes from a 100 μM stock for >1.5 h prior to addition of tOmpA to initiate the folding reaction.

For experiments involving preincubation with either Skp, SurA or BamA POTRAs, OMPs were pre-incubated with these factors for approximately 1 min prior to addition of liposomes. OMPs were diluted and mixed from an 80 μM stock in 8 M urea to a final concentration of 2.4 μM, in the presence of either a two-fold molar excess or SurA or POTRA domains, or a 1:1, 2:1, or 4:1 molar ratio of Skp:OMP, in 0.24 M urea, 50 mM glycine-NaOH, pH 9.5 (no lipids). This stock was then further diluted 6x in presence of 1.28 mM DUPC in 0.24 M urea, 50 mM glycine-NaOH, pH 9.5 to begin the assay. Experiments in which these mixtures were added to DUPC liposomes which contained prefolded BamA, tBamA, or OmpA were performed as for tOmpA alone in the presence of these prefolded OMPs (above). In experiments in which Skp-tOmpA mixtures were added to a solution containing free SurA or BamA POTRAs, as well as DUPC liposomes, SurA and BamA POTRAs were pre-added to liposomes at a concentration of 0.8 μM (a two-fold molar excess over tOmpA). The final volume for each sample was 540 μl. Note that at the concentrations of Skp utilised here, Skp has been shown to be in a dynamic equilibrium between folded monomer subunits and trimers[163]. All Skp concentrations referred to here are trimer equivalents.

For each experiment with a particular liposome batch, four samples were measured concurrently. A minimum of three replicates were globally fitted using IgorPro 6.0 (Wavemetrics, Oregon, USA) to extract rate constant(s), forcing the fits to share the same rate constant(s). Transients were fitted either to a single exponential function:

$$y = A_1 \cdot e^{-k_1 t} + c \qquad (2.1)$$

or to a double exponential function:

$$y = A_1 \cdot e^{-k_1 t} + A_2 \cdot e^{-k_2 t} + c \quad (2.2)$$

Where k_1 and k_2 are rate constants, A_1 and A_2 are their associated amplitudes, and c is a constant. Transients were fitted to a double exponential function if a satisfactory fit was not obtained to a single exponential function as judged by inspection of residuals. Experiments were performed for each condition using three separate liposome batches, and reported errors are the standard deviation of rate constants between liposome batches. The standard deviation, σ, for each rate constant was defined as:

$$\sigma = \sqrt{\frac{\Sigma_i (x_i - \bar{x})^2}{n - 1}} \quad (2.3)$$

where x_i is a rate constant for a particular condition, $\bar{x}$ is the mean rate constant for that condition, and n is the number of replicates.

2.14.2 Kinetic assays for tOmpA in DLPC, DTPC and DMPC in the presence or absence of BamA (Chapter 5)

Stocks of BamA and tOmpA at 100 μM and 80 μM, respectively, were prepared in 8 M urea, 50 mM glycine-NaOH, pH 9.5 (Section 2.14.1). Folding experiments were initiated by dilution of tOmpA to a concentration of 0.4 μM protein and 0.24 M urea in the presence of 1.28 mM DLPC, DTPC and DMPC liposomes (LPR 3200:1), in 50 mM glycine-NaOH, pH 9.5, and monitored by fluorescence spectroscopy as described in Section 2.14. The final volume for each sample was 500 μl. In BamA-containing experiments, 0.8 μM BamA (two-fold molar excess) was prefolded overnight at 30 ºC. As the tOmpA alone experiments exhibited a lag in kinetics a sigmoidal fit to the data was used to guide the eye:

$$y = base + \frac{max}{1 + e^{\left(\frac{\mathrm{xhalf} - x}{rate}\right)}} \quad (2.4)$$

This equation is implemented in the IgorPro 6.0 set of standard fitting functions; $base$ corresponds to the value of y when x is small, $base + max$ corresponds to the value of y when x is large, $xhalf$ corresponds to the value of x when y is at $(base + max)/2$, and $rate$ corresponds to the gradient of the rise in the sigmoid,

with a smaller value for rate leading to a faster rise. Data for tOmpA folding into DLPC or DTPC in the presence of BamA were fitted to a single exponential function, equation (2.1), while data for tOmpA folding into DMPC in the presence of BamA was fitted to a double exponential function, equation (2.2). To compare the tOmpA folding data in the presence and absence of BamA quantatively, the t_{50} value, the time taken to reach 50 % of the total fluorescence change on folding was used. Software was written in the Python programming language[366] to extract t_{50} values. For each transient the minimum fluorescence value was located, and the maximum value was defined by fitting a horizontal baseline to the final section of the data. Three separate liposome batches were used for each lipid type, and for each condition (lipid ± BamA) four transients were used for the t_{50} calcuation i.e. a total of 12 transients per condition. Errors were calculated as the standard error of the mean (*S.E.M*) defined as:

$$S.E.M = \frac{\sigma}{\sqrt{n}} \tag{2.5}$$

where σ is the standard deviation, equation (2.3), and *n*, is the number of samples. To calculate the error in the fold change in t_{50} value with or without the presence of BamA, the standard error of the mean was propagated using the following:

$$\delta R = |R| \cdot \sqrt{\left(\frac{\delta \mathrm{X}}{X}\right)^2 + \left(\frac{\delta \mathrm{Y}}{Y}\right)^2} \tag{2.6}$$

where δR is the error in the fold change, $|R|$ is the fold change value, X and Y are the mean t_{50} values with or without BamA, respectively, and δX and δY are the *S.E.M.* values with or without BamA, respectively.

2.15 Fluorescence emission spectra (Chapter 4)

Fluorescence emission spectra were acquired on the same instrument as the kinetic assays in Section 2.14. Each spectrum was recorded from 305 nm to 400 nm in 1 nm increments, using an excitation wavelength of 295 nm. All spectra were acquired at 25 ºC and all samples contained 50 mM glycine-NaOH, pH 9.5, in a sample volume of 500 μl. OMPs from an 80 μM stock in 8 M urea were diluted to a final concentration of 0.4 μM in the presence of a 2-fold molar excess of Skp in 0.24 M

urea, 8 M urea, or buffer alone in 0.24 M urea. Folded samples were prepared by dilution of an 80 μM OMP stock to 0.4 μM in the presence of 1.28 mM DUPC liposomes (molar LPR 3200:1) in 0.24 M urea and incubated at 25 ºC for ~1.5 h prior to acquisition of the fluorescence emission spectra.

2.16 Temperature denaturation of BamA and OmpA

BamA or OmpA was folded by dilution from a 100 μM stock in 8 M urea, 50 mM glycine-NaOH, pH 9.5, to a final concentration of 0.8 μM OMP, in the presence of 0.24 M urea, 1.28 mM DUPC LUVs in 50 mM glycine-NaOH, pH 9.5, at 25 ºC. The final volume for each sample was 500 μl. A temperature ramp experiment was performed in in which a fluorescence emission spectrum was recorded from 320 nm to 370 nm in 1 nm increments, using an excitation wavelength of 280 nm, every 1 ºC between 20 ºC and 90 ºC. The average wavelength (λ_{av}) of the spectrum was calculated for each measured temperature using:

$$\lambda_{av} = \frac{\sum_{i=320}^{i=370} \lambda_i I_i}{\sum_{i=320}^{i=370} I_i} \qquad (2.7)$$

where λ_i and I_i are the wavelength, and the fluorescence intensity at λ_i, respectively. To extract the temperature at which 50% of molecules are unfolded (apparent T_m) the data were fitted to the following modified Gibbs-Helmholtz equation[367,368]:

$$\lambda_{av} = \frac{(aT + b)\left(\frac{1}{Q_R}\right) e^{\left(\frac{\Delta H}{R}\right)\left(\frac{1}{T_m} - \frac{1}{T}\right)} + cT + d}{1 + \left(\frac{1}{Q_R}\right) e^{\left(\frac{\Delta H}{R}\right)\left(\frac{1}{T_m} - \frac{1}{T}\right)}} \qquad (2.8)$$

where *a* and *b* are the slope and intercept of the pre-transition baseline, respectively, *c* and *d* are the slope and intercept of the post-transition baseline, respectively, ΔH is the change in enthalpy, *R* is the gas constant (8.314 $J.K^{-1}.mol^{-1}$), T_m is the transition temperature in Kelvin, *T* is the temperature in Kelvin, and Q_R is the quantum yield ratio between the folded and unfolded states. The Q_R term is required to account for the observation that the average wavelength (λ_{av}) does not vary linearly with the proportion of folded and unfolded molecules[152,331], and is defined as:

$$Q_R = \frac{\sum_{i=320}^{i=370} I_i\,(F)}{\sum_{i=320}^{i=370} I_i\,(U)} \qquad (2.9)$$

Where $\sum I_i(F)$ and $\sum I_i(U)$ are the sum of the fluorescence intensities measured between 320 and 370 nm in the folded and unfolded state, respectively. Temperature ramp experiments were performed on a Chirascan plus circular dichroism spectrometer (Applied PhotoPhysics) in fluorescence mode. Data fitting was performed using IgorPro 6.0 (Wavemetrics, Oregon, USA).

2.17 Mass Spectrometry

Skp:OMP complexes were prepared by rapid dilution of the denatured OMP (400 μM in 8 M urea, 50 mM glycine-NaOH, pH 9.5) to a final concentration of 5 μM into a solution of Skp (5 μM in 50 mM glycine-NaOH, pH 9.5). The samples were then buffer exchanged into 200 mM ammonium acetate, pH 10 using Zeba spin desalting columns (Thermo Scientific, UK) immediately prior to MS analysis. nanoESI-IMS-MS spectra were acquired using a Synapt HDMS hybrid quadrupole-travelling wave-time of fight mass spectrometer (Waters Corporation, UK) using platinum/gold-plated borosilicate capillaries prepared in-house. Typical instrument parameters were: capillary voltage 1.2-1.6 kV, cone voltage 40 V, trap collision voltage 6 V, transfer collision voltage 10 V, trap DC bias 20 V, backing pressure 4.5 mBar, IMS gas pressure 0.5 mBar, travelling wave height 7 V, travelling wave velocity 250 ms^{-1}. Data were processed using MassLynx v4.1, Driftscope 2.5 (Waters Corporation, UK) and Massign[369]. CCSs were estimated by a calibration approach[370-372] using arrival time data for ions with known CCSs (β-lactoglobulin A, avidin, concanavilin A and yeast alcohol dehydrogenase, all Sigma Aldrich, UK). Estimated modal CCSs are shown as mean ± standard deviation of three independent experiments. Theoretical CCSs for globular proteins with a given effective gas phase density were calculated according to published methods using the following equation[373]:

$$\Omega = \pi(\sqrt[3]{3m/4\pi\rho} + R_{gas})^2 \qquad (2.10)$$

where Ω is the CCS, m is the mass of the protein or protein complex, ρ is the effective gas-phase density, and R_{gas} is the radius of the buffer gas molecule (set to 1.55 Å as nitrogen (N_2) was the buffer gas used). All mass spectrometry

experiments and analyses, including production the mass spectrometry figures in Chapter 4, were performed by Dr A. N. Calabrese.

2.18 Modelling of Skp-OMP complexes

All modelling was performed with the PyMOL Molecular Graphics System (v1.7rc1)[374] using scripts written in Python 2.7.5. To generate models of the Skp:OMP complexes, the missing residues in chains B and C of the Skp crystal structure (PDB: 1U2M[150]) were modelled from chain A. For the Skp:tOmpA/PagP model the OMP was modelled as a sphere of radius 20 Å with its origin positioned at the geometric centre between the α-carbon atoms of residue 50 of each Skp chain. For the Skp:OMP model with expanded Skp subunits, each chain was positioned around a sphere of radius 25 Å representing the larger OMP. The flexible tips of each subunit (residues 51-101) were modelled hinged slightly inwards to wrap around the substrate. The side-by-side parallel and antiparallel 2:1 Skp:OMP models were created by duplication of the Skp:tOmpA/PagP model and appropriate rotation and translation. The interlocking trimer 2:1 Skp:OMP model was generated by duplication of the Skp model with expanded subunits and appropriate rotation and translation of the duplicated Skp. Theoretical CCS values were generated using the calibrated trajectory method implemented in the software IMPACT (Ion Mobility Projection Approximation Calculation Tool)[375].

2.19 Molecular Dynamics Simulations

Molecular dynamics (MD) simulations were prepared using the AmberTools 14 suite of programs, and performed using AMBER and the ff14SB forcefield[376,377]. To simulate apo-Skp in water, a Skp model was first generated from the Skp crystal structure (PDB: 1U2M[150]) with the residues absent in chains B and C modelled from chain A. Following addition of hydrogen atoms using xleap, Skp was placed in a TIP3P water box with a 10.0 Å cutoff and the system neutralised with a total of 15 Cl^- ions. The system was equilibrated by performing an initial energy minimisation, followed by 80 ps of restrained MD during which the system was heated to 300 K with gradual release of restraints. This was followed by an unrestrained MD simulation of 100 ns.

Simulations of the collapse of the extended chains of tOmpA and tBamA were carried out using a Generalised Born/Solvent Accessible surface area (GB/SA)

implicit solvent model[378,379]. Use of an implicit solvent model speeds up exploration of conformational space by at least an order of magnitude due to the neglect of frictional forces from collisions with water molecules[380], leading to rapid adoption of a collapsed configuration from the initially linear structure**.** The polypeptide starting structures were generated in xleap and, after initial energy minimisation, were simulated for 3 ns. The starting models for the simulation of Skp-tOmpA and Skp_2-tBamA *in vacuo* were created in PyMOL by positioning OMPs, after simulated collapse, within the cavity of Skp trimer structures in an 'open' conformation taken from the explicit solvent apo-Skp simulation. All simulations except apo-Skp in explicit water were performed in triplicate. Each *in vacuo* simulation of Skp-tOmpA and 2:1 Skp:tBamA was performed using a starting OMP structure from a different simulation. *In vacuo* simulations of apo-Skp were performed using three different starting structures selected from the simulation of apo-Skp in explicit water. For all *in vacuo* simulations the system was equilibrated by performing an initial energy minimisation, followed by eight steps of restrained MD during which the system was heated to 300 K with gradual release of restraints. This was followed by an unrestrained MD simulation of 100 ns.

To simulate the 1:1 Skp:tOmpA and 2:1 Skp:tBamA complexes in solution, starting models were generated as detailed above for the *in vacuo* simulations. The Skp:OMP complexes were placed in a TIP3P water box with a 10.0 Å cutoff and the system neutralised with a total of 10 Cl^- ions (1:1 Skp:tOmpA) or 12 Cl^- ions (2:1 Skp:tBamA). The systems were equilibrated by performing an initial energy minimisation, followed by 80 ps of restrained MD during which the system was heated to 300 K with gradual release of restraints. This was followed by an unrestrained MD simulation of 100 ns. Simulations were repeated in triplicate for each complex.

Theoretical collision cross-sections (CCSs) of final structures at the end of simulations were calculated, for comparison with IMS-MS data, using the trajectory method and the software IMPACT[375]. CCS values for all structures following *in vacuo* simulations were obtained after 100 ns of unrestrained simulation. CCS values for tOmpA and tBamA in implicit solvent were obtained after 3 ns of unrestrained simulation. The integration time-step was 2 fs and atomic positions were saved every 500 steps (1 ps). The software VMD[381] was used to compute backbone root mean square deviation (RMSD) and to render videos of the

simulations. Analysis of radius of gyration changes over trajectories were carried out with ptraj[382].

MD simulations made use of time on the ARC2 supercomputer facility at the University of Leeds, and time on ARCHER granted via the UK High-End Computing Consortium for Biomolecular Simulation, HECBioSim.

2.20 Microscale thermophoresis (MST) binding experiments

2.20.1 Labelling of tOmpA with Alexa Fluor 488

An N-terminal cysteine residue was introduced in the tOmpA sequence in pET11a by site-directed mutagenesis (using the Q5 site-directed mutagenesis kit as per manufacturers' instructions), to create Cys-tOmpA-pET11a (prepared by Tom Crosskey, University of Leeds)[361]. Cys-tOmpA was expressed and purified as detailed in Section 2.3.1 (Tom Crosskey, University of Leeds) and covalently labelled with Alexa Fluor 488 dye via maleimide chemistry. Alexa Fluor 488 C5 maleimide (Thermo Fisher Scientific, UK) dissolved in DMSO (10 mg/ml) was added to a sample containing 50 μM Cys-tOmpA, 6 M GuHCl, 0.5 mM TCEP 25 mM Tris-HCl, pH 7.2, to a final concentration of 0.5 mM. The total sample volume was 500 μl. The labelling reaction was left overnight at 4 ºC then was loaded onto Superdex Peptide 10/300 column equilibrated with 6 M GuHCl, 25 mM Tris-HCl, pH 7.2 to remove the excess free dye. Samples were collected every 1 ml and peak protein fractions tested for dye labelling using a Nanodrop 2000 (Thermo Fisher Scientific, UK). The labelling efficiency was estimated as ~50%. Cys-tOmpA was created, expressed and purified by Tom Crosskey.

2.20.2 MST protocol

From a 200 μM SurA stock solution in 50 mM glycine-NaOH, pH 9.5, a series of two-fold serial dilutions were performed to obtain sixteen 15 μl samples. Labelled Cys-tOmpA was buffer exchanged into 8 M urea, 50 mM glycine-NaOH, pH 9.5, to a concentration of 1.6 μM. This stock was diluted 16-fold to a concentration of 100 nM with 50 mM glycine-NaOH, pH 9.5, then immediately added to the sixteen SurA-containing samples in 15 μl aliquots (30 μl total sample volume). Therefore, the final concentrations in the samples were 50 nM Cys-tOmpA, 100 μM – 3 nM SurA, 0.25

M urea, 50 mM glycine-NaOH, pH 9.5. Samples were rapidly added to capillaries by capillary action then read using a Monolith NT.115 MST machine (NanoTemper, Germany).

2.20.3 MST data fitting to a quadratic binding equation

The observed signal (S_{obs}) in the MST experiment arises from contributions from both the unbound and bound fraction of labelled tOmpA molecules (F_B and F_u). Therefore:

$$S_{obs} = F_B S_B \times F_U S_U \quad (2.11)$$

Where S_B and S_U are the signals from the bound and free states, respectively. As the sum of the bound and unbound fractions must equal 1 we can write:

$$F_U = 1 - F_B \quad (2.12)$$

Substituting (2.12) into (2.11) gives:

$$S_{obs} = F_B S_B \times (1 - F_B) S_U \quad (2.13)$$

Expanding and rearranging:

$$S_{obs} = F_B S_B + S_U - F_B S_U$$

$$S_{obs} = S_U + (S_B - S_U) F_B \quad (2.14)$$

The equilibrium dissociation constant K_d is defined as:

$$K_d = \frac{[A_{free}][B_{free}]}{[AB]} \quad (2.15)$$

where [A_{free}] is the concentration of one binding partner (fluorescently labelled tOmpA) and [B_{free}] is the concentration of the other binding partner (SurA).

$$[A_0] = \left[A_{free}\right] + [AB]$$

$$[B_0] = \left[B_{free}\right] + [AB]$$

where $[A_0]$ and $[B_0]$ are the total concentrations of *A* and *B*, respectively. Therefore:

$$\left[A_{free}\right] = [A_0] - [AB] \qquad (2.16)$$

$$\left[B_{free}\right] = [B_0] - [AB] \qquad (2.17)$$

Substituting (2.16) and (2.17) into (2.15) gives:

$$K_d = \frac{([A_0] - [AB]) \times ([B_0] - [AB])}{[AB]} \qquad (2.18)$$

Expanding and rearranging:

$$K_d[AB] = [A_0][B_0] - [A_0][AB] - [AB][B_0] + [AB]^2$$

$$0 = [A_0][B_0] - [A_0][AB] - [AB][B_0] - K_d[AB] + [AB]^2$$

$$0 = [AB]^2 - [AB]([A_0] - [B_0] - K_d) + [A_0][B_0] \qquad (2.19)$$

Equation (2.19) is of the form $ax^2 + bx + c = 0$ therefore using the quadratic formula:

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a} \qquad (2.20)$$

$$[AB] = \frac{([A_0] + [B_0] + K_d) \pm \sqrt{([A_0] + [B_0] + K_d)^2 - 4[A_0][B_0]}}{2} \qquad (2.21)$$

The fraction bound is defined as:

$$F_B = \frac{[AB]}{A_0} \qquad (2.22)$$

Substituting (2.14) and (2.22) into (2.21) gives:

$$S_{obs} = S_U + (S_B - S_U) \times \frac{([A_0] + [B_0] + K_d) \pm \sqrt{([A_0] + [B_0] + K_d)^2 - 4[A_0][B_0]}}{2[A_0]}$$

(2.23)

Equation (2.23) was fitted to MST data for SurA (molecule *B*) binding to fluorescently labelled tOmpA (molecule *A*) to obtain a K_d for the interaction. Data fitting was carried out using IgorPro 6.0 (Wavemetrics, Oregon, USA).

3 Development of a fluorescence assay for the study of OMP folding kinetics in low urea concentrations

3.1 Introduction

To aid understanding of the pathways and physicochemical principles of OMP folding over the last 25 years *in vitro* studies have measured OMP folding into a wide variety of amphiphiles and surfactants[383]. Kinetic assays of OMP folding have, in the vast majority of cases, measured folding using cold SDS-PAGE gels[298]. This technique exploits the fact that folded and unfolded OMPs migrate at different apparent molecular weights if left unboiled in SDS-containing sample buffer prior to gel loading. SDS resistance of the native state ('heat modifiability'), which has been known since the 1970s[308-310], allows the direct quantification the fraction folded within a sample, due to the increased stability and compaction of the native state[320]. The technique can also allow the resolution of native dimers and trimers if these are present[311,328]. To measure folding kinetics, following initiation of a folding reaction a sample is taken at different time points to which SDS is added which rapidly binds to the OMP, and is thought to prevent any subsequent folding. The fraction folded at each time point is then quantified by gel densitometry[320]. Far-UV CD has also been used to measure OMP folding kinetics by quantifying the appearance of β-sheet secondary structure[53,348,365] (or additionally, in the case of PagP, native tertiary structure[348]). No studies have examined the kinetics of OMP folding in the presence of additional folding factors by CD, likely due to the complication that the secondary structure of folding factors also contributes to the CD signal.

The work in this chapter describes efforts to obtain conditions suitable for monitoring the *in vitro* folding kinetics of a range of different OMPs using intrinsic fluorescence, as a first step towards comparing how the influence of cellular factors on folding kinetics varies with OMP substrate. An important factor influencing OMP folding efficiency is the final OMP concentration on dilution from denaturant, as higher OMP concentrations increase the rate of aggregation relative to the folding rate. The assay developed here, based on intrinsic tryptophan fluorescence, allows real-time monitoring of folding at much lower OMP concentrations (sub-μM) than normally used in CD or cold SDS-PAGE based studies (low μM concentrations). Fluorescence-based kinetic assays are complementary to cold SDS-PAGE assays, the latter, although a discontinuous assay, and limited as to the number of points per

transient, provides a read out of the absolute fraction of folded SDS-resistant molecules at any point. In some cold SDS-PAGE based studies, notably those involving autotransporters or OMPs folding in the presence of the whole BAM complex[136,239] (Section 1.5.2), western blotting has been used to detect folded populations, allowing low concentrations to be used (e.g. 0.1 μM for the autotransporter EspP[239] and 0.5 μM for FLAG-tagged BamA, respectively[136,283]).

A key problem in the experimental tractability of OMPs is their aggregation propensity. Denaturants, most commonly urea, help to maintain OMPs in a soluble folding-competent form. Depending on the OMP under study, and the nature of the folding environment, a high final denaturant concentration may be required to achieve high folding efficiency. For example, the *Neisserial* OMP Opa_{60} required 4 M urea to fold at >95% efficiency in DDPC ($di_{C10:0}$PC) vesicles at pH 12[384]. Further, a highly aggregation-prone His-tagged PagP construct has been shown to fold into DLPC ($di_{C12:0}$PC) LUVs with the greatest efficiency in 7 M urea[348]. In subsequent work on an untagged PagP construct, a high folding efficiency with exponential kinetics was observed into DLPC LUVs in the presence of 2 M urea, allowing the investigation of the effects of the chaperones Skp and SurA on PagP folding[152]. While it was shown that SurA was stable in conditions ≤ 2 M urea, this was not the case for Skp, although Skp-PagP complexes formed in a low urea concentration (0.24 M) were resistant to subsequent addition of 2 M urea[152]. The addition of high concentrations of denaturant does not appear to perturb synthetic liposomes structure; Rayleigh-Gans-Debye (RDG) scattering experiments showed that DLPC LUVs remained intact in titrations of up to 6 M GuHCl[333]. However, it is less clear how the inclusion of molar concentrations of urea may affect the activity of proteins involved in assisting OMP folding, therefore a low urea concentration (0.24 M) was used in the assay developed here.

Structurally, OMPs generally have short turns on their periplasmic side between the β-strands of their barrels (with TolC a notable exception[80]), and longer loops on their extracellular side[81,89,385]. OMPs have a wide variety of sizes, with known barrel sizes ranging from 8 to 26 β-strands[89,386]. They may also have non-membrane embedded domains on either side of the membrane[89]. All this structural diversity is necessary for the multitude of functions OMPs perform. It influences OMP stability and dynamics, and also the folding kinetics, pathways, and mechanisms en route to the native state. Detailed explorations of the *in vitro* folding mechanisms of OMPs

initially focused almost exclusively on OmpA[383,387], limiting opportunities for comparing and contrasting the folding behaviour of different OMPs. In recent years the *in vitro* folding behaviour of other OMPs have been studied, including the 8-stranded OMPs PagP[152,313,323,332,336,348] and OmpX[300,313,318,388], the 10-stranded OmpT[313], the 12-stranded OmpLA[333,335,389], the 14-stranded OMPs FomA[319,320] and OmpG[390], the 16-stranded OmpF[311,313] and BamA[136,283], and the 19-stranded mitochondrial OMP, hVDAC1[293]. However, conditions in OMP folding studies often need to be optimised for each protein studied limiting opportunities for direct comparison[322], and relatively few studies have looked at the *in vitro* influence of folding factors on OMP folding[136,152,153,233,239,283,299,300]. In particular, no studies have compared the effects of chaperones on the *in vitro* folding behaviour of a range of OMPs.

A study of the aggregation propensity of eight different OMPs in their unfolded aqueous (U_{AQ}) state using sedimentation velocity analytical ultracentrifugation (AUC) revealed variation in the amount of self-association between different OMPs[164]. Interestingly, some of these variations do not have any obvious explanations. While it could be expected that OMPs which have an additional soluble domain (OmpA and BamA) would exhibit less self-association, it is unclear why OmpT, which has a large extracellular region, has a much greater tendency to self-associate than FadL, which has a particularly low folding efficiency *in vitro*[313]. PagP required the greatest addition of urea (5 M) to remain monomeric at the low micromolar concentrations used[164]. The addition of salt (KCl) increased the levels of OMP self-association[164], consistent with experiments on Opa_{50} and Opa_{60} from *Neisseria gonorrhea*, which showed that salt lowers folding efficiency by favouring aggregation[384]. For this reason, no salt titrations were performed when optimising the kinetic assays described in this chapter. High pH was shown to reduce self-association between all OMPs[164], consistent with the notion that at a pH level far from the theoretical pI of *E. coli* OMPs (pI ~4-6), solubility is maintained by the increased negative charge on the protein. This also provides an explanation as to why high pH also increases OMP folding efficiency *in vitro*[313,391,392]; decreased aggregation allows a greater percentage of molecules to fold[164]. Consistent with this, in the assay developed here using PagP, it was also found that higher pH increased folding efficiency (Sections 3.1.1 and 3.1.2). *In vivo* the periplasm is at a similar pH to the extracellular environment[49,84] as the OM is porous to small ions such as protons, providing an additional reason why OMPs are heavily dependent on the proper functioning of the periplasmic chaperone network (Section 1.3)[393].

In the work described in this chapter OmpF was cloned, expressed and purified by A. N. Calabrese, and OmpT was cloned and produced by L. M. McMorran and T. G. Watkinson, respectively.

3.1.1 Expression and purification of PagP

To develop a kinetics assay for the study of the folding of multiple OMPs in the presence of folding factors, the system was first optimised for the 8-stranded OMP PagP. PagP, lacking its N-terminal signal sequence, was expressed as insoluble inclusion bodies, which were isolated, solubilised in denaturant, then further purified by gel filtration in its denatured state (6 M GuHCl, 25 mM Tris-HCl, pH 8.0, Section 2.3.1). A typical A_{280} trace from PagP gel filtration is shown in Figure 3.1a, with SDS-PAGE of peak fractions indicating highly pure PagP (Figure 3.1b).

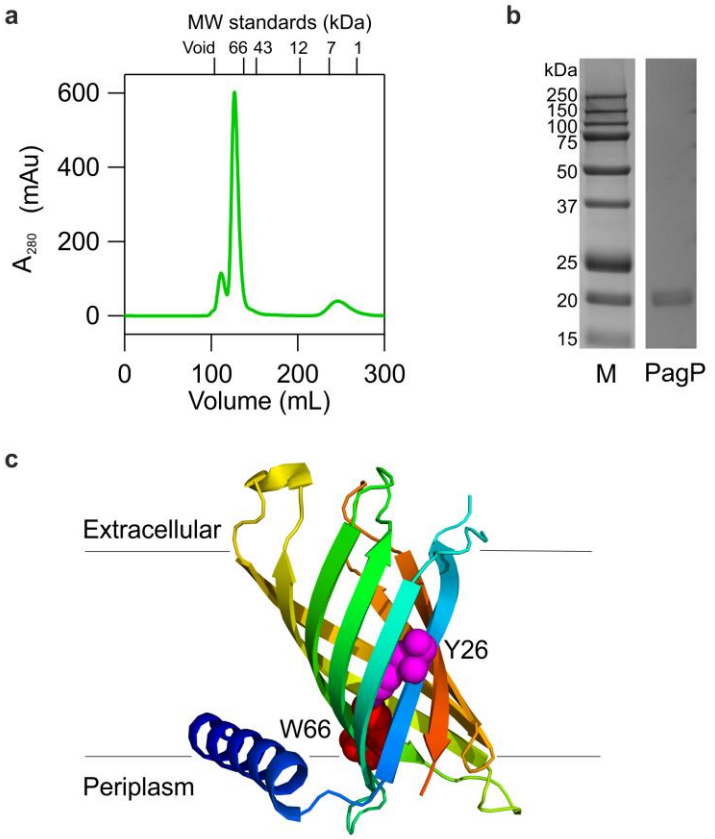


Figure 3.1: Purification of PagP. (a) Purification of PagP by gel filtration (Superdex 75 HiLoad™ 26/60 column) following isolation from inclusion bodies. (b) SDS-PAGE gel lane of pooled peak fractions from gel filtration of solubilised PagP inclusion bodies (elution volumes ~110-140 mL in (a)). (c) Crystal structure of PagP (PDB: 3GP6[345]). Residues involved in exiton formation leading to a positive peak in the native far-UV CD spectra at 232 nm, Y26 and W66[349], are shown as magenta and red spheres, respectively. The native structure is tilted approximately 25° to the normal of the membrane. Black lines indicate approximate location of the membrane.

To verify that PagP, purified in an unfolded stated, could fold to its native state three techniques were used to examinine the relative populations of folded and unfolded OMPs in a sample: fluorescence emission spectroscopy, CD spectroscopy and cold SDS-PAGE band shift assays. These experiments were performed with previously published conditions[332,348] in 100 nm DLPC liposomes, 50 mM sodium phosphate, pH 8.0, with an LPR of 3200:1 at 25 ºC. The fluorescence emission spectrum for folded PagP is blue-shifted with a λ_{max} at a lower wavelength and higher intensity compared to that of unfolded PagP, with a maximum difference at ~335 nm (Figure 3.2a). The CD spectrum of unfolded PagP is consistent with a random coil set of conformations, while the CD spectrum of folded PagP has a negative maximum at

218 nm, typical of β-sheet secondary structure[394] (Figure 3.2b). An additional positive peak at 232 nm, which has been attributed to a 'Cotton effect' between residues W66 and Y26[349] (Figure 3.1c), indicates the formation of native PagP tertiary structure. Overnight PagP folding experiments were performed under the same conditions in the presence of urea concentrations between 1 and 8 M and the fraction of folded protein assessed by a band-shift assay (Figure 3.2c). The results demonstrate that at high urea concentrations (6-8 M) the vast majority (>90%) of PagP is folded, while at lower urea concentrations (1-3 M) most of PagP is unfolded (>80%).

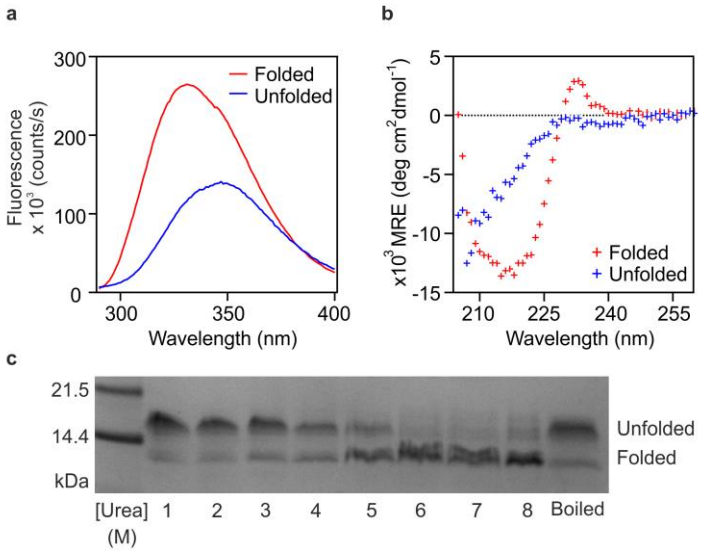


Figure 3.2: Characterisation of PagP folding in DLPC liposomes at pH 8.0. (a) Fluorescence emission spectra of folded and unfolded PagP. (b) CD spectra of folded and unfolded PagP. (c) Folding yields of PagP by SDS-PAGE band shift assay at concentrations of urea between 1 M and 8 M. Folded and unfolded samples in (a) and (b) contained 7 M and 10 M urea, respectively. The concentrations of PagP in (a), (b) and (c) were 0.4 μM, 10 μM and 5 μM, respectively. All samples were folded overnight in 50 mM sodium phosphate, pH 8.0, at 25 ºC.

3.1.2 Effects of lipid chain length on PagP folding yields in PC liposomes

Evidence in the literature suggests that the *in vitro* folding efficiency of OMPs varies greatly with liposome composition, with a general trend of increasing efficiency with decreasing acyl chain length for folding into phosphatidylcholine LUVs[53,313]. Previous work by Dr. Alice I. Bartlett (University of Leeds)[395] demonstrated that PagP folds

with the greatest efficiency at pH 9.5 (4 μM PagP, 1 M urea, 10 mM glycine-NaOH, DLPC liposomes, LPR 3200:1, 37 ºC). Accordingly, this pH was used to investigate the impact of decreasing the acyl chain lengths on PagP folding efficiency. Further, the LPR was kept at the high value of 3200:1 as increasing the lipid:OMP ratio correlates with increased folding efficiency[53], presumably by increasing the likelihood of interaction between OMPs and a membrane surface, compared with self-association and potentially aggregation.

Overnight PagP folding experiments were performed in the presence of differing concentrations of urea (0.24-8 M) with liposomes composed of DDPC (*di*$C_{10:0}$PC), DUPC (*di*$C_{11:0}$PC) or DLPC (*di*$C_{12:0}$PC) (Figure 3.3). The results showed that in high urea concentrations (7-8 M) PagP folding efficiency is high (>90%) regardless of lipid type (Figure 3.4). However, at low urea (0.24-1 M) the majority of PagP is unfolded (>60%) when folded into DLPC liposomes (Figure 3.4c). By contrast, when liposomes are composed of lipids with a reduced acyl chain length (DDPC or DUPC, Figure 3.4a,b) at low urea concentrations (1 M) the majority of PagP remaining in solution is folded (>80%). PC lipids with a chain length of lower than C9 form micellular structures[396], and as the folding efficiency for DDPC liposomes was close to 100%, PC lipids with an acyl chain length of C9 were not investigated.

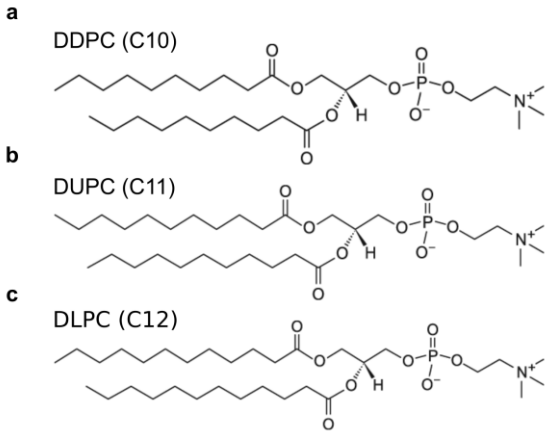


Figure 3.3: Phosphatidylcholine lipids used in folding assay optimisation. Chemical structures of (a) 1,2-*d*i-decanoyl-*sn*-glycero-3-phosphocholine, *di*$C_{10:0}$PC (DDPC), (b) 1,2-*d*i-undecanoyl-*sn*-glycero-3-phosphocholine, *di*$C_{11:0}$PC (DUPC), and (c) 1,2-*d*i-lauroyl-*sn*-glycero-3-phosphocholine, *di*$C_{12:0}$PC (DLPC).

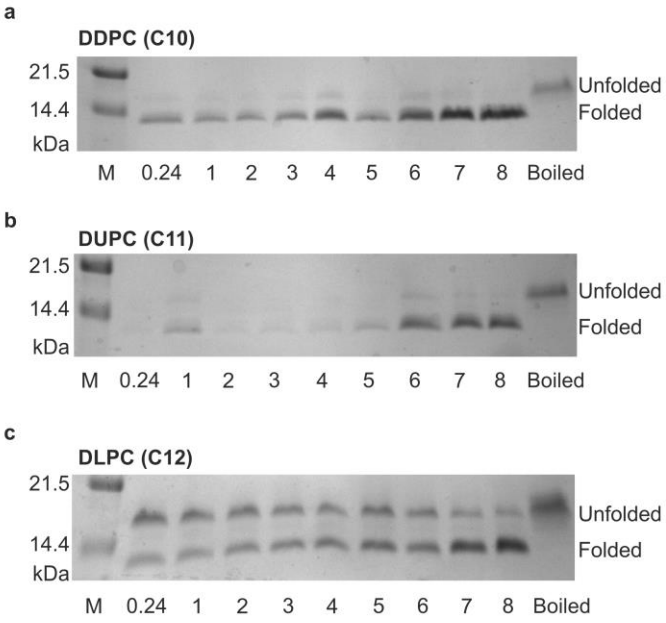


Figure 3.4: PagP folding yields in PC liposomes of different acyl chain lengths as a function of urea concentration. SDS-PAGE band shift assay in concentrations of urea between 0.24 M and 8 M urea in liposomes composed of (a) DDPC, (b) DUPC, and (c) DLPC. Samples containing 4 μM PagP, 50 mM glycine-NaOH, pH 9.5 were folded overnight at 37 ºC.

3.1.3 Effects of lipid chain length on PagP folding kinetics in PC liposomes

Next, fluorescent spectroscopy was used to examine the effects of acyl chain length on PagP folding kinetics in the presence of a low urea concentration (0.24 M) (Figure 3.5). Folding into DLPC liposomes was incomplete over a 2 h time-course (Figure 3.5c), consistent with previous observations[152]. An exponential fit of this data could not be made over this time period. By contrast, PagP folding into DDPC was rapid with the reaction mostly complete within a few minutes, and the data well described by a double exponential function (Figure 3.5a). However, a previous comparative study of the folding kinetics of nine different OMPs indicated that PagP is a relatively slow folding OMP[313]. The aim of the present study was to compare the folding behaviour of PagP with that of much faster folding OMPs such as OmpA and OmpF[313] in the presence of folding factors which may increase their folding rate. Therefore, it was likely that the folding of some OMPs under these conditions using DDPC liposomes would be too fast to be measurable by manual mixing. However, PagP folding into DUPC liposomes was also well described by a double exponential function and occurred on a substantially slower timescale (Figure 3.5b).

Therefore, the results show that very small changes in lipid composition (the additional of a single carbon atom to the acyl chain of phosphatidylcholine lipids) dramatically alters the *in vitro* efficiency and kinetics of PagP folding, consistent with previous results[313,323]. Further, for measuring the folding kinetics of multiple OMPs it appears that the bilayer thickness of LUVs made from PC lipids with a chain length of C11 are ideal under the conditions employed, in that PagP folding was not 'too fast' or 'too slow' to measure the rates reliably. Therefore, this lipid was selected for further studies to compare the folding rate constants of a variety of OMPs under the same conditions.

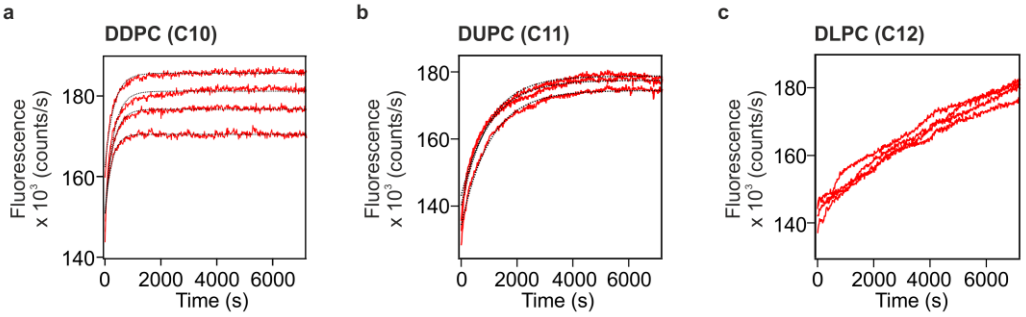


Figure 3.5: PagP folding kinetics in PC liposomes of different acyl chain lengths in 0.24 M urea. (a) DDPC, (b) DUPC, and (c) DLPC. Samples contained 0.4 μM PagP, 1.28 mM lipids, 0.24 M urea, 50 mM glycine-NaOH, pH 9.5, at 37 ºC. Fits to a double exponential equation are shown as black dotted lines in (a) and (b).

3.2 Extension of the OMP folding kinetics assay to additional OMPs

3.2.1 OMP selection

Having optimised the assay conditions for measuring the folding kinetics of PagP, additional OMPs were selected to compare their folding behaviour under identical conditions. OMPs were selected according to the following criteria: (1) a crystal structure is available, (2) each OMP had been previously folded *in vitro*, (3) all are found in the OM of *E. coli*, (4) they are structurally and functionally distinct, and (5) they cover a range of barrel sizes (from 8 to 16 β-strands). Details of the OMPs chosen are given in Table 3.1 and their three-dimensional structures are shown in Figure 3.6.

OMP	No. of β-strands	Molecular weight* (kDa)	Function	Reference
tOmpA	8	18.875	Adhesin/invasin/evasin	Pautsch *et al*. (1998)[397]
PagP	8	19.110	Acyl transferase	Hwang *et al*. (2002)[343]
OmpT	10	35.284‡	Protease	Vandeputte-rutten *et al*. (2001)[398]
FadL	14	46.038	Fatty acid transporter	Van den Berg *et al*. (2004)[399]
OmpF	16	37.216	Porin	Cowan *et al*. (1992)[400]
tBamA	16	43.240	Insertase/secretase	Noinaj *et al*. (2013)[11]

Table 3.1: Model OMPs investigated have a range of functions, sizes and number of β-strands. *Molecular weight of the mature OMP sequences, lacking their cleavable N-terminal signal sequences. ‡Molecular weight for the OmpT construct used here including an N-terminal hexahistidine tag and a TEV protease cleavage site.

The OMPs tOmpA (Figure 3.6a) and PagP (Figure 3.6b) and tBamA (the transmembrane domain of BamA, Figure 3.6f) have already been described in Sections 1.5, 1.5.1, and 1.4.1, respectively.

OmpT is a 10-stranded endoprotease with specificity for paired basic residues[401]. It is used as a defence against colicin-mediated attack[402] and is also implicated in bacterial pathogenicity[403]. OmpT is involved in processing proteins of the autotransporter family, and cleaves the antimicrobial peptide proteamine, which is secreted by urinary tract epithelial cells[403]. The crystal structure of OmpT[398] (Figure 3.6c) revealed a 10-stranded β-barrel with a large extracellular region including long loops which contain the active site residues[404].

FadL is a 14-stranded OMP involved in the transport of fatty acids[405,406]. FadL shows selectivity for long chain fatty acids (LCFAs); it was found to bind myristic acid (C14:0) with an order of magnitude lower K_d than the low μM affinity observed for oleic acid (C18:1) and palmitic acid (C16:0)[407]. LCFAs acquired from the environment via FadL serve as sources of carbon and energy[408]. Additionally, removal of proinflammatory host LCFAs implicates FadL in immune evasion[409]. FadL was selected for study on the basis that, in the absence of additional factors, it

folds with a particularly poor folding yield *in vitro* (~25% in DDPC LUVs) compared to eight other OMPs folded under the same conditions (all >60% in DDPC LUVs)[313]. It was reasoned that FadL may be a useful model for studying the improvement in folding mediated by, for example, periplasmic chaperones. Similar to OmpT, a substantial percentage of the FadL barrel is not membrane-embedded and extends into the extracellular space (Figure 3.6d).

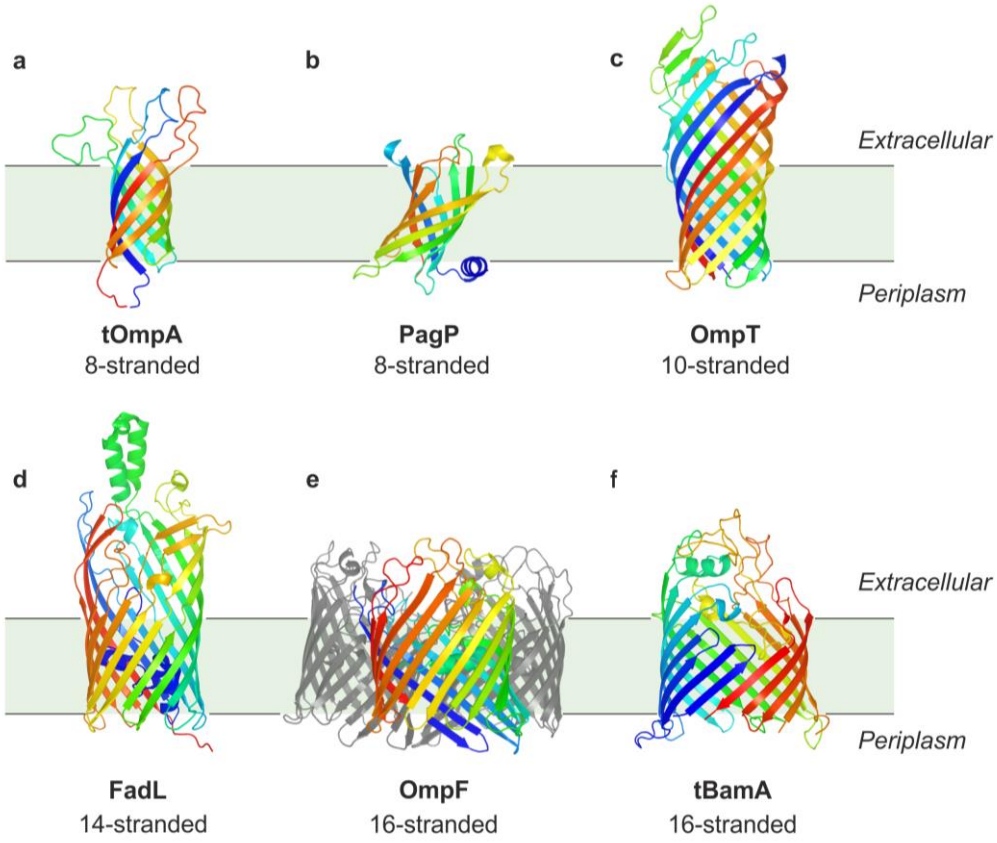


Figure 3.6: Three-dimensional structures of OMPs used in this study. (a) tOmpA (PDB:1G90[410]), (b) PagP (PDB:1THQ[344]), (c) OmpT (PDB:1I78[398]), (d) FadL (PDB:1T16[399]) (e) OmpF (PDB:2ZFG[411]), and (f) tBamA (PDB:4N75[287]). All structures are to scale. OmpF is a functional trimer; for clarity only one subunit is coloured and the other two are shown in grey. The OM is represented in pale green, and approximate positions of OMPs within the membrane are judged by the position of residues in the 'aromatic girdles' at either side of the membrane[412].

OmpF is a trimeric porin and is one of the most abundant *E. coli* OMPs with an estimated ~100,000 copies per cell[49,83]. Each OmpF monomer forms a narrow, slightly cation selective pore[85], that allows the passage of hydrophilic molecules of <600 Da across the OM. It is important in multidrug resistance (MDR) as antibiotics,

such as the β-lactams and fluoroquinines, enter cells via the OmpF pore[413]. OmpF is also used as the entry point for toxins produced by other bacteria; following initial binding of colicin E9 to BtuB, intrinsically disordered colicin regions are threaded through two pores of an OmpF trimer in opposite direction to aid translocation of the cytotoxic colicin domain[414].

Expression plasmids for tOmpA, PagP, OmpT and OmpF were obtained from previous work. However, the project required the cloning of plasmids for the expression of FadL and tBamA, details of which are given below.

3.2.2 Subcloning of FadL

A codon-optimised gene encoding the mature sequence of *E. coli* FadL was synthesised by Eurofins Genomics, Germany. The synthetic gene was delivered in Eurofins Genomics' standard high-copy number PEX-K4 plasmid[360], and therefore required subcloning into a suitable expression vector. To obtain the expression vector backbone, a midiprep of PagP-pET11a was performed and ~5 µg of the resulting DNA was double digested with NdeI and BamHI restriction enzymes. Next, the vector was dephosphorylated and the reaction was run on an agarose gel, together with non-digested and singly digested controls, and the double digested bands gel extracted (Figure 3.7a). The FadL-PEX-K4 plasmid was also midiprepped and ~20 µg of the DNA produced was double digested and run on an agarose gel. The band corresponding to the FadL gene was gel extracted (Figure 3.7b). An overnight ligation reaction was performed between extracted, double digested pET11a backbone and the extracted FadL gene. This was then transformed into competent DH5α cells and the resulting colonies midiprepped and sent for sequencing to confirm the presence of the FadL gene.

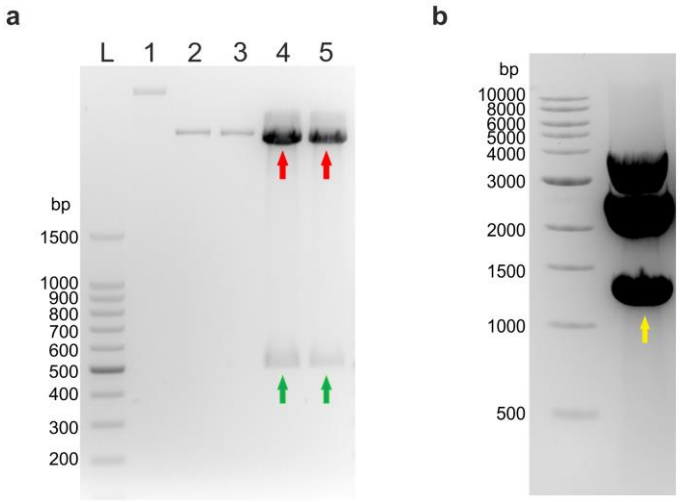


Figure 3.7: Subcloning of FadL. (a) Restriction digests of pET11a-PagP. Lane L: 100 bp ladder. Lane 1: Undigested PagP-pET11a. Lane 2: BamHI digest of pET11a. Lane 3: NdeI digest of pET11a. Lane 4,5: NdeI/BamHI double digest of pET11a. The success of the double digestion reaction is indicated by the different gel mobility of the double digest reaction (lanes 4,5) compared to that of the digests with a single restriction enzyme (lanes 2,3). Additionally, bands corresponding to the PagP insert size (~500 bp) are present in the double digest reaction lanes (indicated by green arrows), which are not present in the lanes containing the vector digested with a single restriction enzyme (lanes 2,3). The bands corresponding to the double digested pET11a vector backbone (indicated by red arrows) were gel extracted for subsequent ligations. (b) NdeI/BamHI double digest of the FadL-PEX-K4 vector. The band corresponding to the FadL gene (~1300 bp, indicated by a yellow arrow) was gel extracted for subsequent ligation. The two bands above this at ~3500 bp and ~2500 bp correspond to undigested vector and digested vector backbone, respectively.

3.2.3 Cloning of tBamA

To clone the gene for tBamA PCR primers were designed for the region encompassing residues 425-810 of BamA, and a PCR reaction performed (Section 2.2.6), using the BamAB-pETDUET-1 vector as a template (Table 2.4). The PCR reaction was run on an agarose gel and the product gel extracted and digested with NdeI and BamHI restriction enzymes (Section 2.2.9). The digested tBamA product was run on an agarose gel, gel extracted, and subsequently ligated into the pET11a vector backbone, digested with NdeI and BamHI, obtained above (Section 2.2.12). Following transformation of the ligation reaction into supercompetent XL1-Blue cells, DNA minipreps were obtained from eight colonies. These eight minipreps were analysed by restriction digestion (Figure 3.8) and those containing an insert of the correct size (1169 bp) were sent for sequencing.

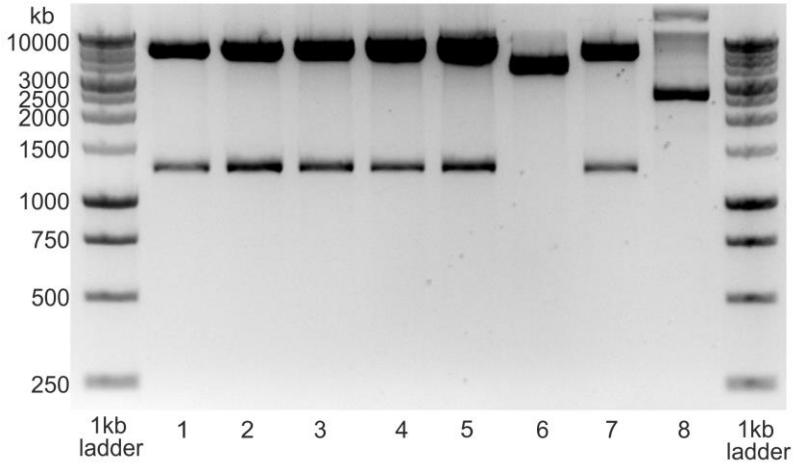


Figure 3.8: Cloning of tBamA into pET11a. Agarose gel analysis of double digests (NdeI/BamHI) of plasmid DNA obtained from colonies transformed with a ligation reaction between a PCR product of the mature tBamA gene sequence and pET11a. Clones in lanes 1-5 and 7 contained an insert of the approximately correct size (1169 bp). Sequencing confirmed the correct tBamA sequence was present in the clone in lane 1.

3.3 Expression and purification of additional OMPs

The pET11a expression vectors containing the sequences for tOmpA, PagP, OmpT, FadL and tBamA were transformed into competent BL21(DE3) cells (Section 2.2.4) and expressed as detailed in Section 2.3. All OMPs were subsequently successfully purified from inclusion bodies to a purity of >95% (Figure 3.9) as outlined in Section 2.3.

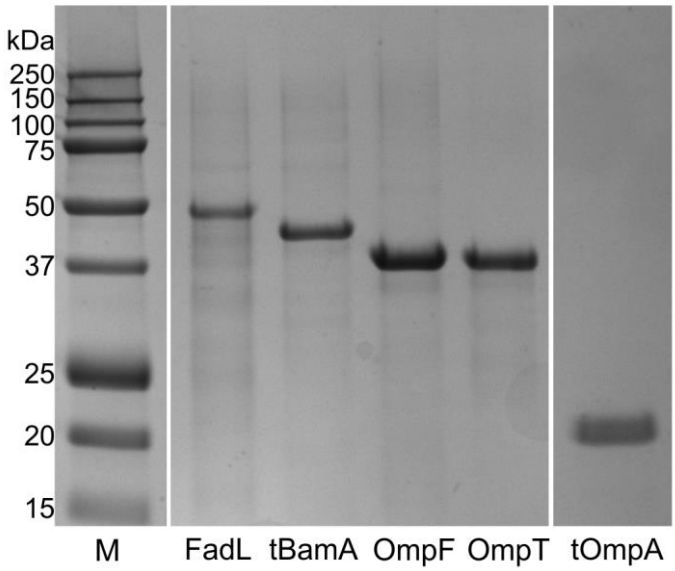


Figure 3.9: Purification of tOmpA, OmpT, OmpF, tBamA and FadL. SDS-PAGE analysis of additional OMPs studied following purification by inclusion body isolation and gel filtration.

3.4 Kinetic assays in the absence of lipids

To verify that the fluorescence transients obtained for the six OMPs studied represent insertion and folding into a membrane environment, and not some other process that may give rise to a fluorescence change, control assays were performed in the absence of liposomes. For the 8-stranded OMPs tOmpA and PagP, and the 16-stranded OMPs OmpF and tBamA no change in Trp fluorescence was observed on dilution from 8 M to 0.24 M urea (Figure 3.10). Further, there are no increases in noise and/or decreases in signal over time, which would provide evidence of aggregation, indicating the formation of a stable soluble form, as previously reported[152,160].

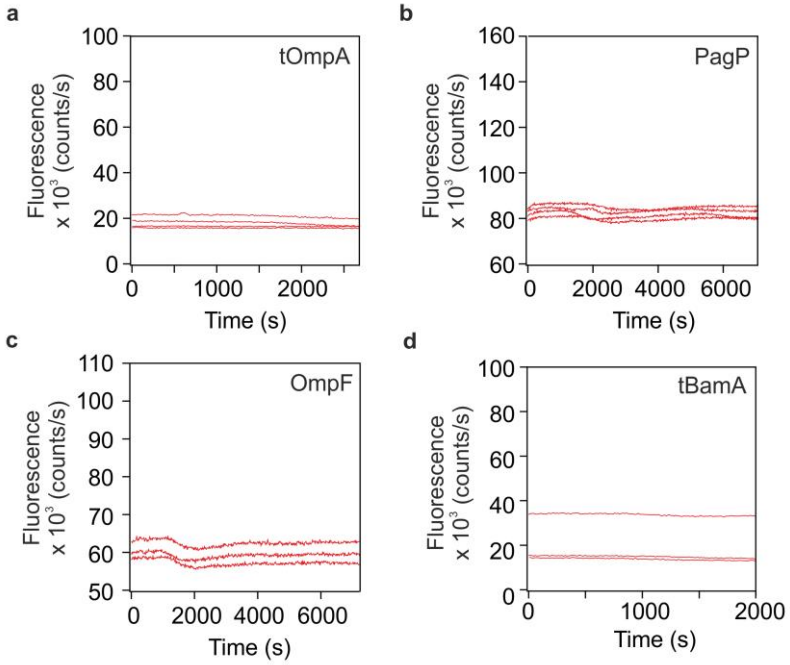


Figure 3.10: OMP folding transients in the absence of lipids. Example kinetic traces for (a) tOmpA, (b) PagP, (c) OmpF, and (d) tBamA, in the absence of lipids, monitored by fluorescence emission spectroscopy. Assays were performed with OMP concentrations of 0.4 µM, in 0.24 M urea, 50 mM glycine-NaOH, pH 9.5, at 25 ºC. At least three replicates are shown for each protein.

However, in the cases of OmpT and FadL folding transients were observed in the absence of lipids (Figure 3.11a,b), with good fits to the data obtained with single or double exponential functions, respectively. In the case of OmpT this presumably reflects folding of the large extracellular domain which contains three tryptophan residues (Figure 3.11c). Similarly, FadL contains two tryptophan residues which are located beyond the membrane on the extracellular side (Figure 3.11d) that may contribute to the observed fluorescence change. Additionally, FadL has three short N-terminal helices which plug the barrel in the native state[398] (Figure 3.6d) which possibly aid the formation of partially folded soluble conformations. Therefore, kinetic studies of OmpT and FadL folding into liposomes were not pursued due to the difficulty of separating and interpreting changes in fluorescence caused by membrane insertion and folding, and those observed in the absence of lipids.

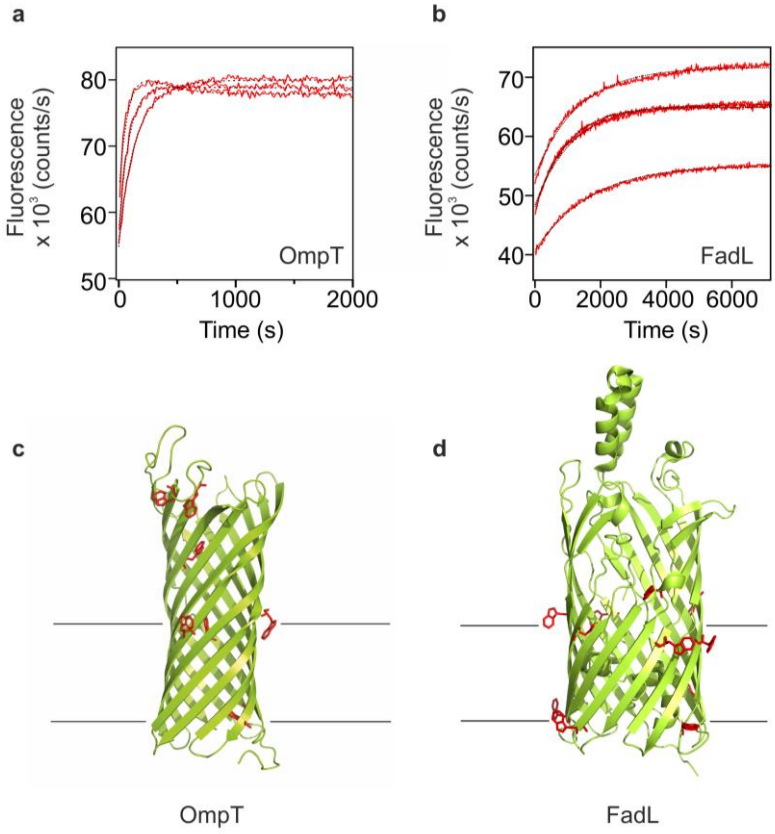


Figure 3.11: Folding transients for OmpT and FadL in the absence of lipids. Kinetic folding traces for (a) OmpT, and (b) FadL, in the absence of lipids, monitored by fluorescence emission spectroscopy. The data for OmpT and FadL are well described by a single and double exponential, respectively, indicated by black dashed lines. Assays were performed with OMP concentrations of 0.4 µM, 50 mM glycine-NaOH, pH 9.5, at 25 ºC, in either in 0.24 M or 2 M urea (in (a) and (b), respectively). At least three replicates are shown for each protein. (c) Crystal structure of OmpT (PDB:1I78[398]). (d) Crystal structure of FadL (PDB:1T16[399]). Tryptophan residues are shown in stick representation and highlighted in red.

3.5 Highly reproducible exponential kinetics for tOmpA, PagP, OmpF and tBamA under identical conditions

Finally, folding experiments into liposomes were performed for the substrates tOmpA, OmpF and tBamA using conditions identified for PagP (Section 3.1.3) (0.4 µM OMP, 50 mM glycine-NaOH, 0.24 M urea, LPR 3200:1, pH 9.5, 25 ºC). Exponential folding transients could be measured for these three OMPs under these conditions (Figure 3.12). The data for tOmpA could be fitted to a single exponential, while the data for PagP, OmpF and tBamA required an additional exponential term. In each experiment, four samples were run concurrently. A minimum of three of these transients were globally fitted, forcing them to share the same rate constant,

but allowing their amplitudes to vary. The amplitude data was not considered further due to the variability introduced by stochastic variation in the amount of aggregation in folding experiments, and between the results acquired on different days. The extracted folding rate constants show very minor variation between liposome batches (Table 3.2).

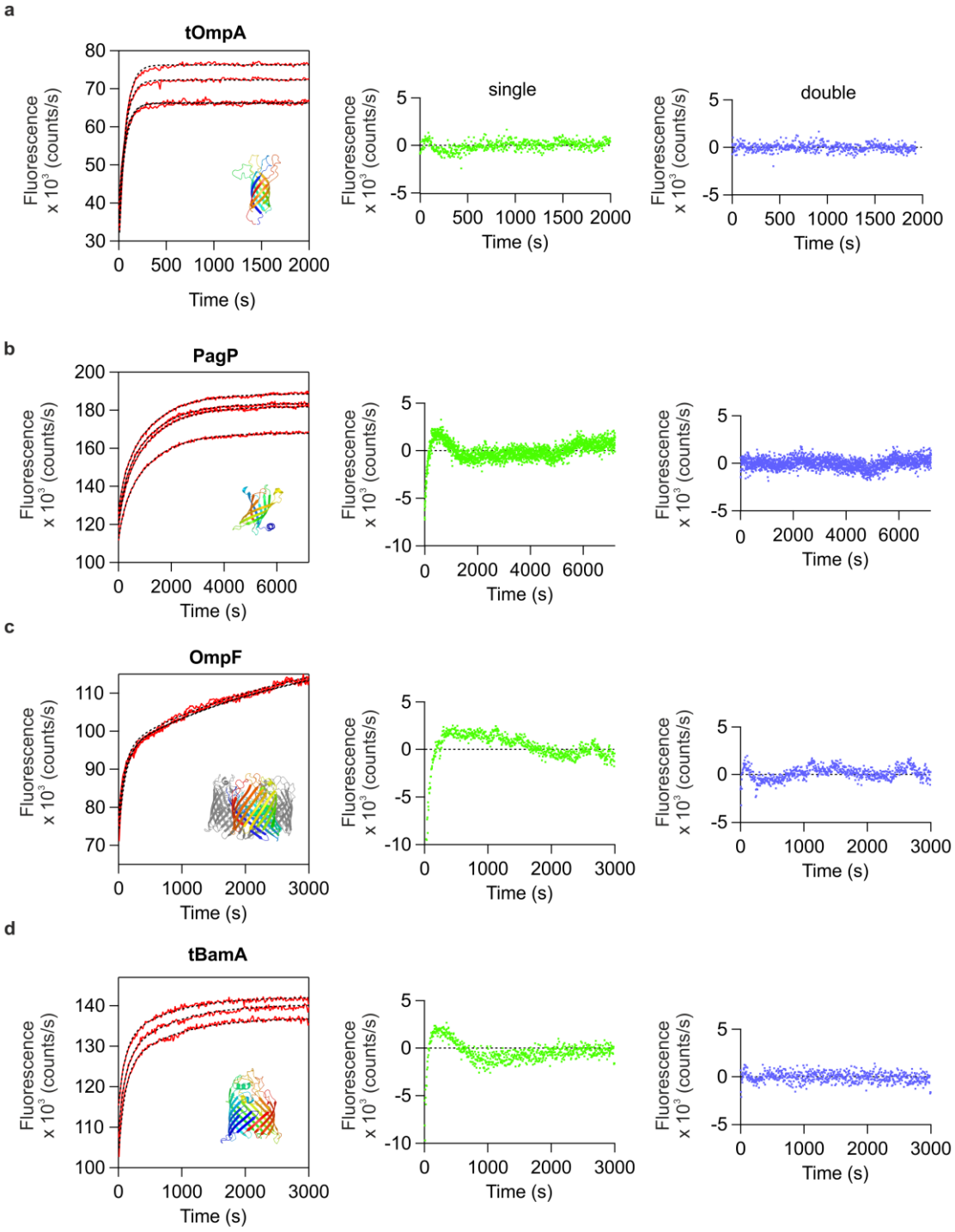


Figure 3.12: The OMPs tOmpA, PagP, OmpF and tBamA fold with highly reproducible exponential kinetics. Folding transients for (a) tOmpA, (b) PagP, (c) OmpF, and (d) tBamA, folding into DUPC liposomes monitored by fluorescence emission spectroscopy (left). Residuals for fits to a single (middle) or double (right) exponential function are shown. Assays were performed with OMP concentrations of 0.4 µM, in 0.24 M urea, 50 mM glycine-NaOH, LPR 3200:1, pH 9.5, at 25 ºC. At least three replicates are shown for each protein. Exponential fits to the data (a single exponential for tOmpA (a) and double exponentials for PagP, OmpF and tBamA (b)-(d)) are indicated by black dashed lines.

OMP	k_1 ($x10^{-3}$ s^{-1})	k_2 ($x10^{-3}$ s^{-1})	A_1 (range)	A_2 (range)
tOmpA	15.7 ± 0.6	N/A	-27300, -51081	N/A
PagP	3.8 ± 0.8	0.8 ± 0.1	-13587, -30984	-28220, -46976
OmpF	10.5 ± 0.6	0.5 ± 0.2	-5728, -43673	-4448, -61239
tBamA	12.8 ± 2.4	1.6 ± 0.3	-11272, -35841	-11307, -37961

Table 3.2: Measured rate constants for OMP folding into DUPC liposomes. Data are shown as the mean ± the standard deviation of the rate constants obtained from three separate folding experiments, each using independently prepared batches of liposomes. For each batch of liposomes at least three folding transients were fitted globally to obtain the rate constants shown. Amplitude data was variable and shown as a range between minimum and maximum values. N/A: The kinetic traces for tOmpA were adequately described by a single exponential.

Several alterations to the method previously used in our laboratory for the storage and preparation of DLPC liposomes[152] were made, which aided the reproducibility of experiments with DUPC liposomes: (1) purchased solid lipids were dissolved in 80:20 ethanol:methanol, aliquoted, and stored in solution, instead of being stored in solid form and weighed out for each liposome batch; (2) an additional step was added to the protocol for making liposomes. Following resuspension of lipid films an additional step of 5x freeze/thaw cycles was performed to aid resuspension prior to extrusion (Section 2.11); and (3) DUPC liposomes were used within 48 h of preparation.

3.6 Discussion

This chapter presents the first examination of the *in vitro* folding kinetics of multiple OMPs using fluorescence spectroscopy. OMP folding behaviour is highly dependent on the properties of the lipid bilayer[53], and here LUVs made from saturated PC lipids were used as the folding environment. Studies examining the principles of OMP folding often employ PC lipids[53,298,313,320], despite the fact that Gram-negative bacteria contain no PC[34]. The positive spontaneous curvature of bilayers formed from PC lipids may be the physical reason for the high OMP folding efficiencies achieved in these lipids[32,415]. Liposome size, as well as composition, influences *in vitro* OMP folding efficiency and kinetics[53,297,313]. LUVs are made by extrusion through filters 100-400 nm in size, while SUVs (~40-50 nm) are generally made by sonification of LUVs[322]. OMP folding efficiency is generally higher in SUVs which

have a greater elastic curvature than LUVs[313,416]. This may be due the increased stress within SUVs leading to bilayer defects, which exposes hydrophobic surfaces thereby aiding OMP folding[320]. Early work found that OmpA could not fold in LUVs composed of DMPC ($di_{C14:0}$PC) under the conditions used, but would readily fold into SUVs of PC lipids with chain lengths of up to C18[297,391]. However, this effect is OMP dependent as liposome size had no influence on the yield of folded OmpF trimers[311]. The liposomes employed here were 100 nm LUVs. LUVs were selected as opposed to SUVs, due to their lower curvature strain. However, given the typical dimensions of an *E. coli* cell (~2-4 μm x ~1 μm[417]) the average curvature of the OM *in vivo* is likely much lower. Indeed, recently it was found that nascent OMPs are inserted into the central region of the OM in *E. coli*, the spatial location of the OM which exhibits the lowest curvature[37,290].

Examination of the effects of lipid chain length on PagP folding demonstrated that thinner bilayers promote greater folding yields and faster kinetics (Figure 3.4 and Figure 3.5), likely due to increased fluidity in the membrane[322]. This is consistent with previous folding experiments in which OMPs are diluted from denaturant[313], and with experiments which used a cell-free transcription-translation system[392].

A low protein concentration (0.4 μM) and a high LPR (3200:1) was used throughout this chapter to decrease the probability of aberrant protein-protein interactions, which may decrease yields. The kinetics of OMP folding are strongly dependent on the concentration of liposomes, following a second order rate law, with increased LPR resulting in faster folding (as demonstrated for OmpA folding into DLPC liposomes)[53].

There are clear differences in the kinetic behaviour of the four OMPs compared here (Figure 3.12). This variation does not correlate with protein size, consistent with previous results[313]; while the 8-stranded tOmpA folds faster (~2-times) than the 16-stranded OMPs, tBamA and OmpF in DUPC LUVs, the 8-stranded PagP folds ~2-times more slowly than the 16-stranded OMPs. Differences in folding behaviour likely results from variations in states along the pathway(s) to the native state. The structural diversity of the four OMPs in the native state may be reflected in their ensemble of denatured states. For the 8-stranded OmpX NMR studies found that the urea unfolded state (in 6.5 M urea) contains regions of structure, including a helical segment (residues 73-82) and a hydrophobic cluster (residues 137-145) both

of which were present in ~25 % of molecules[418], with these results supported by MD simulations[419]. Different levels and types of structure in the denatured states may result in variation between OMPs in collapsed misfolded (U_{AQ}) states on dilution from denaturant. In turn, differences in U_{AQ} states may affect interaction with lipid bilayers, all of which may affect folding rates.

Additionally, it is not obvious why the kinetics of tOmpA folding under the conditions employed here can be adequately fitted to a single exponential model, while the kinetics of PagP, OmpF and tBamA folding require a second exponential term (Figure 3.12 and Table 3.2). Where multiple exponential kinetics are seen in folding transients, this may reflect the formation of on- or off-pathway folding intermediates[314,383], or be due to parallel folding pathways[319-321,323]. It is also possible that the slow folding phase observed for OmpF (Figure 3.12c) is the result of trimerisation[311]. *In vivo* experiments determined that monomers of the trimeric transporter LamB stably fold into the OM prior to trimerisation[123], and folded dimers as well as trimers have been identified in OmpF folding experiments *in vitro*[311,313]. However, a further complication in interpreting fluorescence data for OmpF is that it contains only two tryptophans which exhibit different fluorescence properties in the native state[420]. One of these tryptophans (W61) is located at the trimer interface[421], consistent with the notion that the slow folding phase observed is the result of trimerisation. Future work could investigate the origins of this multiphase kinetic behaviour, and the reasons for differences between OMPs.

A key finding in this chapter is that the use of tryptophan fluorescence is not appropriate for monitoring the folding status of all OMPs. Importantly, the assay developed here does not report directly on the formation of secondary or tertiary structure, but on the change in polarity of OMP Trp residues. The underlying assumption is that changes in observed fluorescence are due to interactions with and/or burial within the bilayer. Accordingly, it was found for OmpA folding into DLPC liposomes that the kinetics of Trp burial were faster than those of secondary and tertiary structure measured by CD and cold SDS-PAGE[53]. Available evidence suggests that on dilution from denaturant OMPs undergo a rapid collapse to a misfolded or partially folded state[160,173,391] occurring in the dead time of the manual mixing experiments performed here. The presence of a folding transient when monitoring the OmpT and FadL in the absence of lipids (Figure 3.11b) suggests a structural rearrangement from the collapsed state which gives rise to signal changes

independent of lipid interaction, thereby limiting the usefulness of intrinsic fluorescence as a probe for bilayer folding and insertion for these OMPs.

In summary, here it is demonstrated that folding kinetics could be reliably measured for set of OMPs with different molecular weights, functions and number of β-strands (Table 3.1) under identical conditions into DUPC liposomes using fluorescence spectroscopy. This allows the influence of folding factors on folding to be directly compared and contrasted for different OMPs, which is the focus of the following chapter.

4 Skp is a multivalent chaperone of outer membrane proteins

4.1 Introduction

The final folded location of β-barrrel OMPs in the OM is distal to their site of synthesis in the cytosol[422], and OMPs are inherently aggregation prone due the patterns of alternating hydrophobic and hydrophilic residues in their sequences[81]. Therefore, chaperones are required at all stages of OMP assembly to prevent OMP self-association[164], and to maintain unfolded OMPs in a soluble, folding-competent state[359]. It has been reported that poly(Val-Lys) peptides form stable β-sheet containing structures[423], and that *de novo* designed peptides with alternating polar and non-polar residues have a high propensity to form amyloid-like fibres[114]. Accordingly, it was recently found that OMPs, which contain similar alternating polar and non-polar residues, can also form amyloid-like fibres, as judged by EM, and binding of the amyloid-specific dye Thioflavin T (ThT) to OmpA aggregates in the absence of chaperones[424]. As fibrillar protein species are strongly associated with cellular toxicity[425], this further highlights the importance of chaperones in controlling OMP assembly pathways.

Following synthesis and translocation across the inner membrane, OMPs must traverse the periplasm (~165-210 Å[17,22,426]) before reaching the OM, where they are folded and inserted by the β-barrel Assembly Machinery (BAM) complex[9,13,68,152] (Section 1.4). In assisting OMPs across the periplasm chaperones prevent aberrant interactions, which could lead to aggregation, and maintain OMPs in a folding-competent state. Additionally, in the case of some OMPs, such as OmpA, attached soluble domains within the same polypeptide chain can fold independently and act to chaperone the β-barrel domain[427]. Recent kinetic modelling studies suggest that OMPs may make many hundreds of interactions with chaperones as they traverse the periplasmic space[124]. The thermodynamics and kinetics of chaperone-OMP complexes, as well as the synthesis and degradation rates of all components, must be finely balanced to ensure the correct flux in the system and the maintainance of OM homeostasis[17,124,336]. Despite the availability of crystal structures for the chaperones implicated in OMP assembly (Skp, SurA, FkpA, and DegP[10]) the mechanism(s) by which they bind and release substrates, in the absence of ATP[21,428], is poorly understood[10].

In the case of the trimeric 'holdase'[152] chaperone Skp, OMPs are sequestered within a hydrophobic cage to prevent their misfolding and aggregation during transport across the periplasm (Section 1.3.1). Skp has broad substrate specificity[155], with reported affinities for its substrates in the low nanomolar range[160,336]. Skp is a functional homotrimer (referred to herein as Skp) with a 'jellyfish'-like architecture[149,150] (Figure 4.1), consisting of three α-helical 'legs' that extend 60 Å away from the 'body' domain; a 9-stranded β-barrel which mediates trimerisation[149,150].

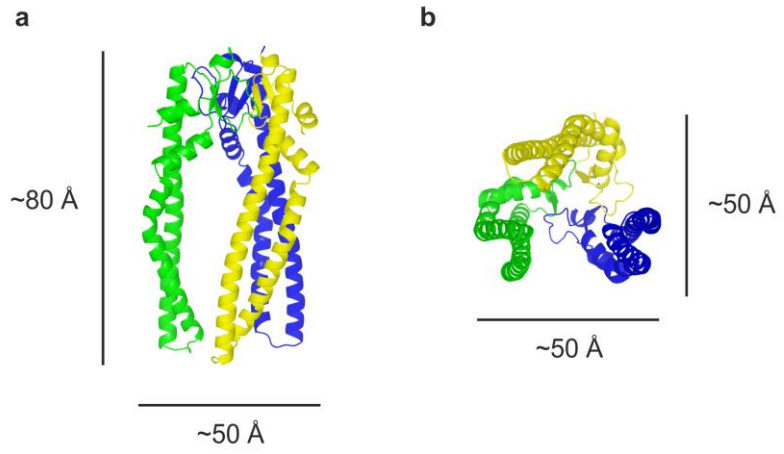


Figure 4.1: Dimensions of the Skp chaperone. Crystal structure of Skp (PDB: 1U2M[150]) from (a) the side, (b) the bottom. Missing residues in chains B and C were modelled from chain A using PyMOL. Residues in chains A, B, and C are shown in green, blue and yellow, respectively.

The three subunits of Skp form a hydrophobic cavity inside which OMP clients are bound[150,155,161,171], with previous studies suggesting a 1:1 stoichiometry for all Skp:OMP complexes[156,160,166]. These 1:1 stoichiometries have been proposed in studies using tryptophan fluorescence (Skp complexes formed with tOmpA (19 kDa)[160], NalP (32 kDa)[160], OmpG (33 kDa)[160], OmpA (35 kDa)[160] and BamA (89 kDa)[160]), SEC-MALS (OmpA (35 kDa)[170]), NMR (OmpX (16 kDa) and tOmpA (19 kDa)[171]) and fluorescence correlation spectroscopy (OmpC (38 kDa)[156]). The Skp hydrophobic cavity has been estimated to be able to accommodate folded proteins of ~25 kDa[150], but many OMPs known to interact with Skp are considerably larger (e.g. the 22-stranded BtuB and the 26-stranded LptD are 66 kDa and 87 kDa, respectively)[12]. Therefore, this raises the question of what are the structural alterations that must occur for Skp to accommodate its larger substrates?

To investigate the mechanism by which Skp sequesters OMPs of different sizes, this chapter presents kinetic studies of OMP folding, complemented with electrospray ionisation-ion mobility spectrometry-mass spectrometry (ESI-IMS-MS) analyses, computer modelling and molecular dynamics simulations. The interactions are examined between Skp and five diverse OMPs introduced in the previous chapter: (1) tOmpA, the 8-stranded transmembrane domain of OmpA[387]; (2) PagP, an 8-stranded acyl transferase enzyme[342]; (3) OmpT, a 10-stranded protease[398]; (4) OmpF, a 16-stranded trimeric porin[411]; and (5) tBamA, the 16-stranded transmembrane domain of the BamA OMP insertase[16] (Figure 3.6, Table 3.1).

The mass spectrometry experiments described were performed and analysed by Dr A. N. Calabrese (University of Leeds). The *in vacuo* molecular dynamics simulations of Skp were initiated by P. W. A. Devine (University of Leeds). All simulations were performed and analysed by B. Schiffrin, with the exception of one replicate of the Skp-tOmpA simulations in explicit solvent (Section 4.6) which was carried out by Dr S. A. Harris (University of Leeds). The work presented in this chapter has been published in: Schiffrin, B., Calabrese, A. N. *et al.*, Skp is a multivalent chaperone of outer membrane proteins, *Nature Structural & Molecular Biology*, **23**, 786-793, 2016[173].

4.2 Higher Skp:OMP ratios are required to prevent folding of larger OMPs

The work presented in Chapter 3 determined optimised conditions for measuring exponential folding kinetics for PagP and successfully extended the assay to the OMPs tOmpA, OmpF and tBamA (Section 3.5). Here, this approach was exploited to compare the kinetic behaviour of these OMPs in the presence of Skp.

To verify that the OMPs selected for study are indeed folded at the end of the observed transients (Figure 3.12), fluorescence emission spectra of each OMP were measured in 8 M urea and compared with spectra of membrane inserted (folded) OMP obtained at the end point of the folding reaction (Figure 4.2, red and grey spectra). Spectra of tOmpA, PagP, OmpF and tBamA folded into liposomes show a characteristic blue-shift in λ_{max} and increase in fluorescence intensity compared with spectra of the unfolded proteins in 8 M urea, indicating that a substantial fraction of all four OMPs fold successfully into the 100 nm DUPC liposomes. Next, to determine whether the OMPs are able to interact with Skp under the experimental

conditions selected, fluorescence emission spectra of each OMP were measured in 0.24 M urea in the absence of liposomes with or without a two-fold molar excess of Skp (Figure 4.2, cyan and purple spectra). In the presence of Skp, the spectra of tOmpA, OmpF and tBamA show decreases in λ_{max} and/or increases in fluorescence intensity compared with spectra of these OMPs in buffer alone, demonstrating that these unfolded OMPs interact with Skp. For PagP, no change in fluorescence is observed in the presence of Skp compared with buffer alone, however previous work, using pull-down experiments, has shown that these proteins do interact under these conditions[152]. The unusual fluorescence emission spectrum for PagP in buffer alone may be due to the presence of the 19-residue N-terminal helix which contains two Trp residues[348].

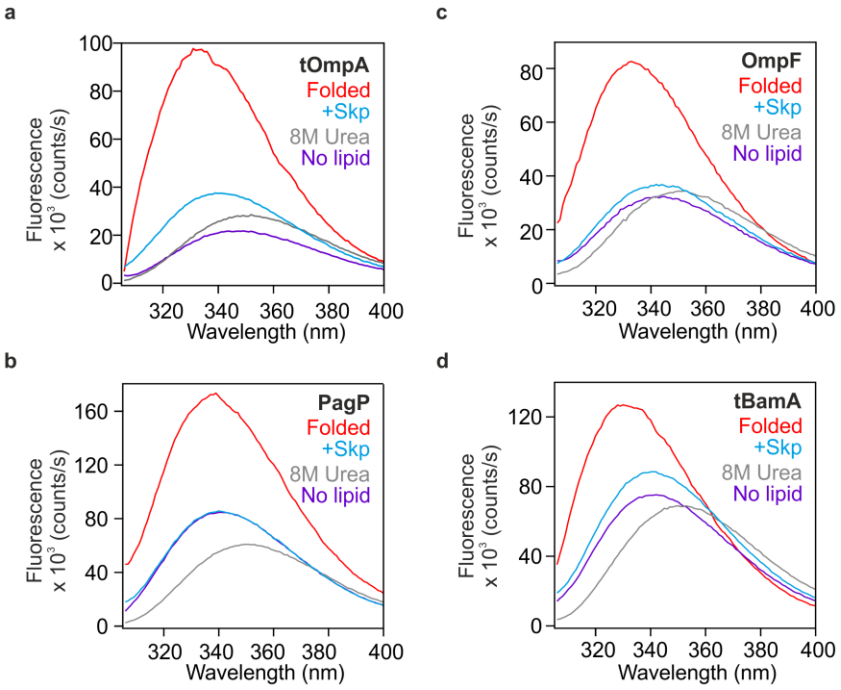


Figure 4.2: tOmpA, PagP, OmpF and tBamA interact with Skp and fold into DUPC liposomes under selected conditions. Fluorescence emission spectra of OMPs folded in DUPC liposomes (red lines), unfolded in 8 M urea (grey lines), in buffer alone in the absence of lipids (purple lines), and in the presence of a 2-fold molar excess of Skp (cyan lines) for (a) tOmpA, (b) PagP, (c) OmpF, and (d) tBamA. Final OMP concentrations were 0.4 µM, with a molar lipid:protein ratio of 3200:1 (in folded samples), in 0.24 M urea, 50 mM glycine-NaOH, pH 9.5, at 25 ºC.

The effect of Skp on OMP folding kinetics was next investigated. In the absence of Skp, all four OMPs fold with single or double exponential kinetics (Figure 4.3, left), allowing extraction of the folding rate constants (k_1 and k_2) (Figure 4.3, right, Table

4.1). The folding kinetics of OMPs that had been pre-incubated with Skp at molar ratios ranging from 1:1 to 4:1 (chaperone:client) (where Skp concentrations are in trimer equivalents) were then measured by adding the pre-formed complexes to DUPC liposomes and monitoring OMP folding by fluorescence spectroscopy (Figure 4.3, left, Figure 4.4). Concentrations of Skp are given as trimer equivalents as AUC experiments have shown that at the concentrations utilised here, Skp is in a dynamic equilibrium between folded monomer subunits and trimers[163]. The results show that at a substrate concentration of 0.4 µM, a 2-fold molar excess of Skp is sufficient to prevent tOmpA and PagP folding into DUPC liposomes (Figure 4.3a,b, Figure 4.4a,b). It has been shown previously that while incubation of Skp with PagP at a 2:1 molar ratio prevents OMP folding over a 2 h timescale, overnight incubation with liposomes allows OMP release and folding to the equivalent yields as in the absence of Skp[152]. For both of the 8-stranded OMPs tOmpA and PagP, a 1:1 Skp:OMP ratio retards folding, while a 2:1 ratio prevents folding over the timescale of the experiment. By contrast, pre-incubation of OmpF and tBamA with a 2-fold excess of Skp decreases the folding rate, but does not prevent folding (Figure 4.3c,d, Figure 4.4c,d). However, pre-incubation with a 4-fold excess of Skp inhibits folding of these larger proteins over the timescale of the experiment (Figure 4.3c,d, Figure 4.4c,d. right). As the K_d for Skp binding is similar for all OMPs measured (0.3 - 83 nM)[160], the results suggest, therefore, that complete sequestration of larger OMP barrels requires the binding of more than one copy of Skp.

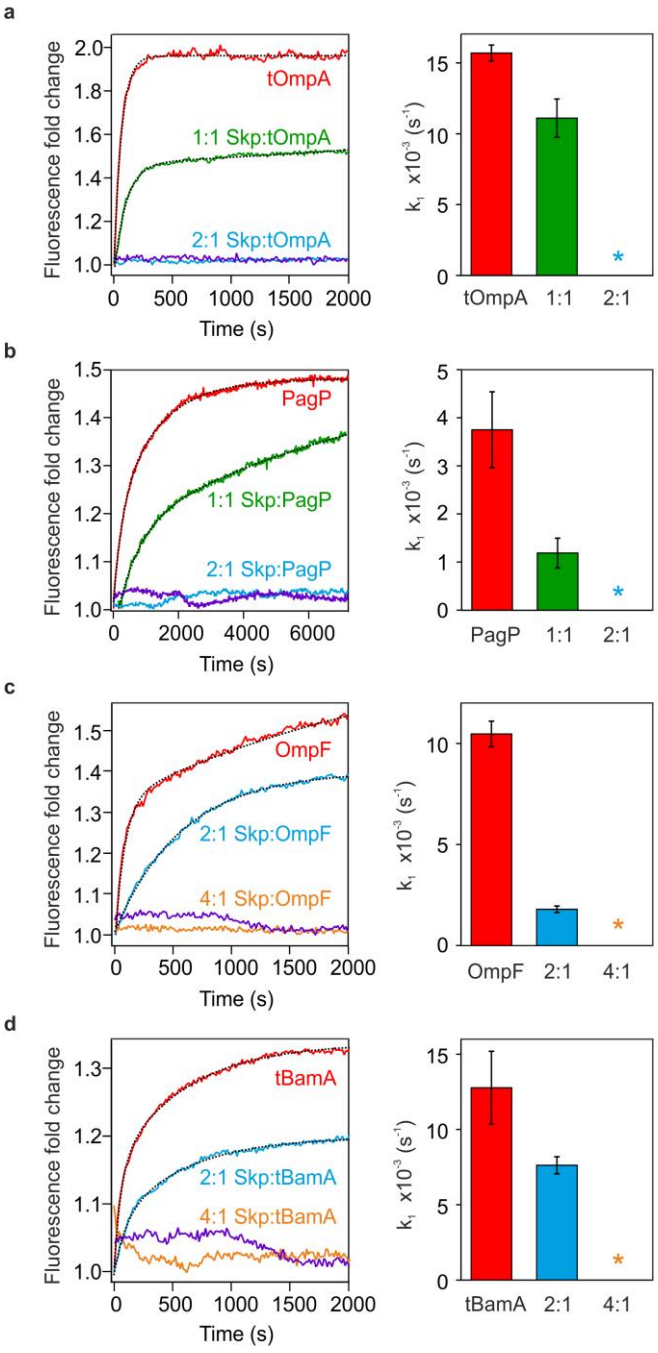


Figure 4.3: Different Skp:OMP ratios are required to sequester OMPs and inhibit folding. Folding data for (a) tOmpA, (b) PagP, (c) OmpF, and (d) tBamA. (left) Kinetic traces for the folding of OMP alone into DUPC liposomes (red lines), and in the presence of Skp at a 1:1 molar ratio (tOmpA and PagP, green lines, parts a and b), a 2:1 molar ratio (cyan lines, parts a–d), a 4:1 molar ratio (OmpF and tBamA, orange lines, parts c and d), or in the absence of lipids (purple lines). Single or double exponential fits to the data are shown as dotted lines. (right) Rate constants for OMP folding in the absence or presence of Skp at the OMP:Skp ratio indicated. Where the data were best described by a double exponential fit, the rate constants for the second, slower, phase are not shown, but are given in Table 4.1. A star indicates a Skp:OMP ratio where folding did not occur on the timescale of these experiments. Data are shown as mean ± s.d. of three independent experiments using three separate liposome batches.

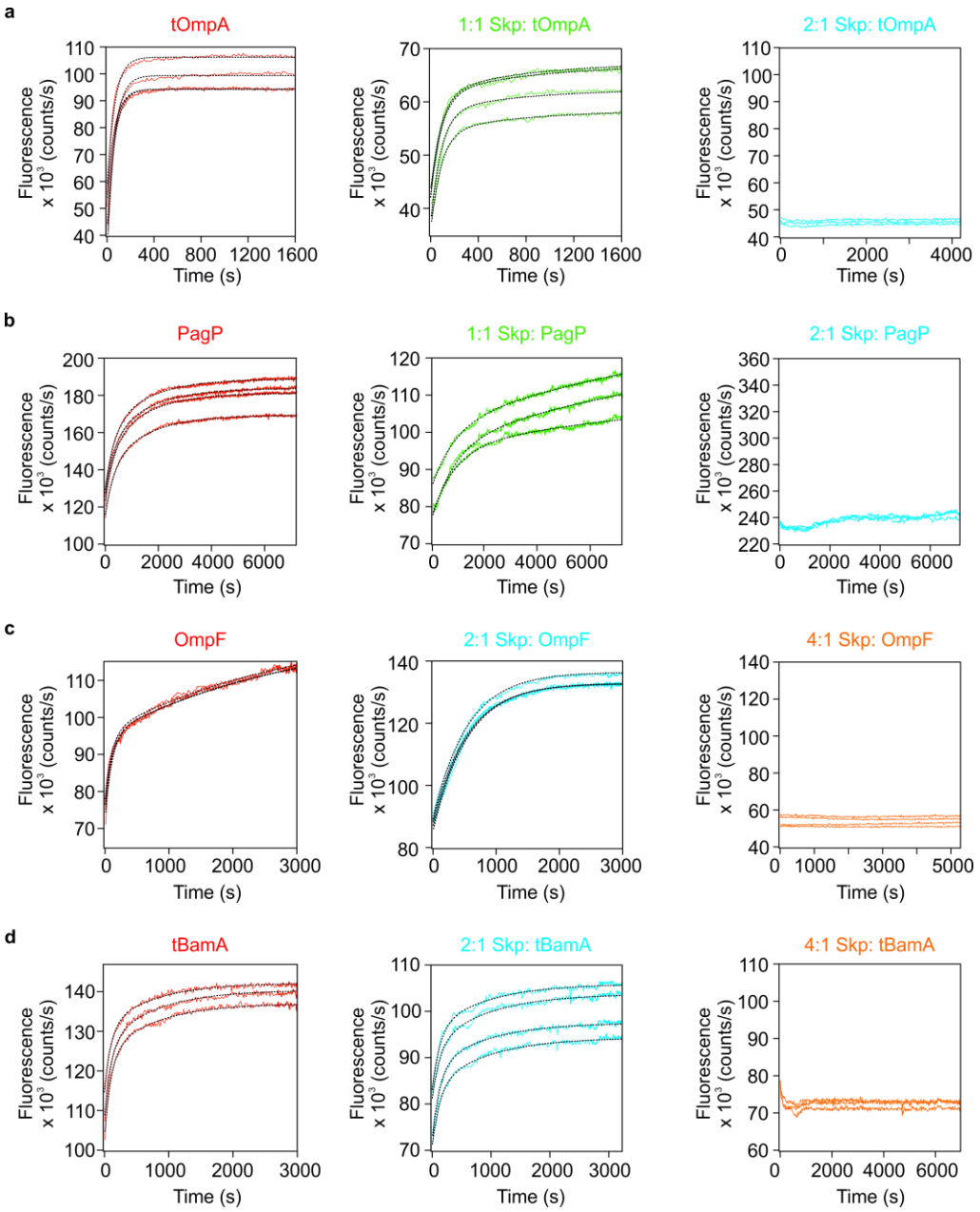


Figure 4.4: 16-stranded OMPs require pre-incubation with a greater molar excess of Skp than 8-stranded OMPs to inhibit folding into synthetic liposomes. Example kinetic folding data for (a) tOmpA, left panel and tOmpA in the presence of a 1:1, centre panel, or 2:1, right panel, molar ratio of Skp:tOmpA; (b) PagP, left panel and PagP in the presence of a 1:1, centre panel, or 2:1, right panel, molar ratio of Skp:tOmpA; (c) OmpF, left panel and OmpF in the presence of a 2:1, centre panel, or 4:1, right panel, molar excess of Skp; (d) tBamA, left panel and tBamA in the presence of a 2:1, centre panel, or 4:1, right panel, molar excess of Skp. Pre-incubated Skp-OMP complexes were added to DUPC liposomes and OMP folding was monitored by fluorescence spectroscopy. Final OMP concentrations were 0.4 μM, with a molar lipid:protein ratio of 3200:1, in 0.24 M urea, 50 mM glycine-NaOH, pH 9.5. A minimum of three transients are shown in each panel. Exponential fits to the data are indicated by black dashed lines.

OMP	OMP:Skp ratio	k_1 ($x10^{-3}$ s^{-1})	k_2 ($x10^{-3}$ s^{-1})
tOmpA	OMP alone	15.7 ± 0.6	N/A
	1:1	11.1 ± 1.4	1.3 ± 0.1
	1:2	No folding	No folding
PagP	OMP alone	3.8 ± 0.8	0.8 ± 0.1
	1:1	1.2 ± 0.3	0.1 ± 0.02
	1:2	No folding	No folding
OmpF	OMP alone	10.5 ± 0.6	0.5 ± 0.2
	1:2	1.8 ± 0.2	N/A
	1:4	No folding	No folding
tBamA	OMP alone	12.8 ± 2.4	1.6 ± 0.3
	1:2	7.6 ± 0.6	1.4 ± 0.4
	1:4	No folding	No folding

Table 4.1: Measured rate constants for OMP folding into liposomes after pre-incubation with varying molar ratios of Skp. Data are shown as the mean ± the standard deviation of the rate constants obtained from three separate folding experiments, each using independently prepared batches of liposomes. For each batch of liposomes at least three folding transients were fitted globally to obtain the rate constants shown. N/A: The kinetic traces for the condition were adequately described by a single exponential.

4.3 ESI-MS reveals differing Skp:OMP stoichiometries according to OMP size

The experiments described above suggest that the binding of more than one Skp is necessary to prevent the folding of larger OMPs. To gain insights into the architectures of these complexes, ESI-MS coupled with IMS was used to analyse different Skp:OMP assemblies within multicomponent mixtures. Skp complexes with tOmpA, PagP, OmpT, OmpF and tBamA were prepared in the absence of detergent or lipid (Section 2.17) and, following buffer exchange into a MS-compatible buffer, ESI-MS spectra were acquired using carefully tuned instrumental conditions (Section 2.17), allowing non-covalent interactions to be retained *in vacuo* (non-covalent or "native" MS)[370,429,430] (Figure 4.5, Table 4.2). The results show that the complexes of all five OMPs with Skp are sufficiently stable to survive the ESI process and be transferred into the gas phase for analysis. Binding to Skp maintains the solubility of these OMPs prior to mass spectral analysis, which is typically

achieved by solubilising with detergents or other media[431,432], or by using harsh denaturing conditions[433,434].

Several reports have suggested that Skp binds unfolded OMPs ranging from 16-89 kDa (8-16 β-strands in the native state) with a 1:1 stoichiometry[156,160,166,171]. By contrast, the ESI mass spectra shown in Figure 4.5 reveal that the stoichiometry of these assemblies is dependent on the size of the OMP client, consistent with the kinetic traces shown in Figure 4.3. Thus, tOmpA and PagP (8-stranded OMPs) bind only one Skp, whilst the larger OMPs, OmpT, OmpF and tBamA (10- and 16-stranded OMPs) bind up to two copies of Skp, at least as suggested by ESI-MS. Peaks corresponding to monomeric Skp subunits were also observed in the spectra at *m/z* ~2000 (Figure 4.5), indicating either that some dissociation of the assembly is occurring in-source, and/or reflecting a population of monomeric Skp subunits in solution[163,167]. Interestingly, a 2:1 assembly is the predominant complex observed in the spectrum for the largest 16-stranded OMP studied, tBamA (Figure 4.5f), despite being formed by mixing Skp with tBamA at a 1:1 molar ratio. To investigate whether the absence of the POTRA domains in tBamA alters complex formation, the interaction of Skp with full-length BamA was also investigated (Figure 4.6). This sample also displays a 2:1 Skp:OMP stoichiometry in addition to a 1:1 complex, which has been reported previously[160]. Interestingly, the 1:1 stoichiometry previously reported for Skp in complex with full-length BamA was obtained using fluorescence studies[160]. This suggests that the binding of additional copies of Skp does not increase OMP fluorescence, which is saturated at a 1:1 Skp:OMP ratio.

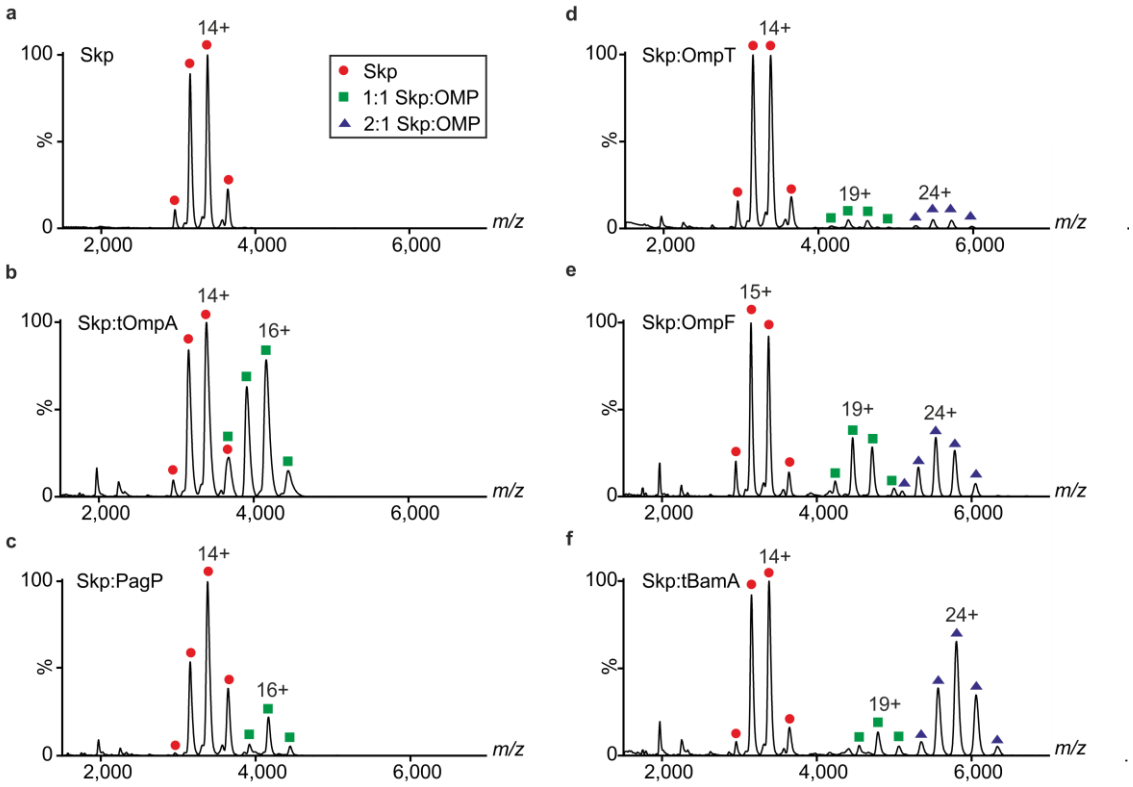


Figure 4.5: Skp:OMP complexes have different stoichiometries. ESI mass spectra of (a) Skp (5 μM) and Skp pre-incubated in the presence of 5 μM (b) tOmpA, (c) PagP, (d) OmpT, (e) OmpF and (f) tBamA. OMPs were diluted from a denatured state (in 50 mM glycine-NaOH pH 9.5, 8 M urea) into Skp-containing solutions (final OMP and urea concentrations of 5 μM and 0.2 M, respectively), and incubated at room temperature for 5 min before buffer exchange into 200 mM ammonium acetate, pH 10. Spectra are annotated with red circles (●) (Skp), green squares (■) (1:1 Skp:OMP) and blue triangles (▲) (2:1 Skp:OMP). The most abundant charge state is labelled for each distribution. Observed masses for the complexes are summarised in Table 4.2. Complementary IMS data are displayed in Table 4.3.

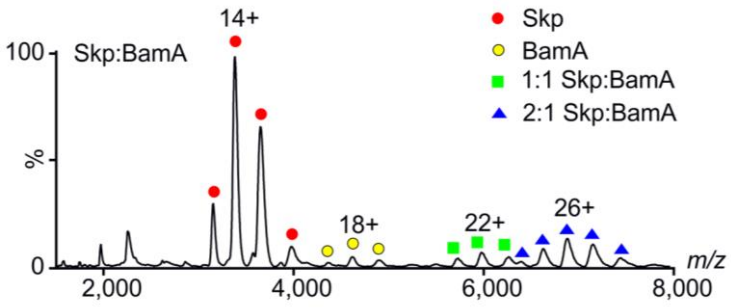


Figure 4.6: BamA forms complexes with Skp with a 2:1 stoichiometry. ESI mass spectra of Skp (5 μM) pre-incubated in the presence of BamA (5 μM). Experiments were performed, and peaks labelled, as detailed in Figure 4.5.

Complex	Expected Mass (Da)	Observed Mass (Da)
Skp	47,075	47,125 ± 10
1:1 Skp:tOmpA	65,818	65,898 ± 12
1:1 Skp:PagP	66,185	66,208 ± 15
1:1 Skp:OmpT	82,355	83,304 ± 20
2:1 Skp:OmpT	129,430	130,542 ± 14
1:1 Skp:OmpF	84,220	84,210 ± 18
2:1 Skp:OmpF	131,295	131,337 ± 11
1:1 Skp:tBamA	90,315	90,183 ± 5
2:1 Skp:tBamA	137,390	137,303 ± 8

Table 4.2: Observed and expected masses for the complexes studied. For mass measurements, mass spectrometry conditions were optimised empirically for each sample to decrease peak widths, by increasing the cone voltage and trap collision energy, (typically 100 V and 40 V, respectively), relative to the gentler conditions used to acquire IMS data.

4.4 Insights into the structure and dynamics of Skp:OMP complexes using ESI-IMS-MS

The folding kinetics and non-covalent mass spectrometry data described above provide evidence that Skp binds its OMP clients of larger size by increasing the Skp:OMP stoichiometry. Next, ESI-IMS-MS was used to study the structures and dynamics of the complexes identified. IMS measures the mobility of ions through an inert gas-filled chamber under the influence of a weak electric field, with the drift time (the time taken to traverse the ion mobility cell) of an ion dependent principally on its size and charge[370,371]. Conceptually, a useful analogy is with gel electrophoresis, in which differences in size and charge of different species are used for their separation and measurement, based on their mobility through a medium[435]. Similar to the manner in which a marker lane of protein standards is used for calibration to obtain the masses of observed bands in a gel electropherisis experiment, here, using travelling wave IMS-MS[370,371] calibration of drift time data using a set of globular standards allowed rotationally averaged collision cross-sections (CCSs) for each ion in the ESI mass spectra to be obtained. The CCS parameter is a measure of the interaction area between an ion and the buffer gas, which is related to the conformation of an ion, as well as its structure[436,437]. From a

Protein Data Bank (PDB) file containing the complete structural information for a protein or protein complex a theoretical CCS value can be calculated, here using the software IMPACT[375]. For globular assemblies, the error between measured CCS values and theoretical CCS values calculated from atomic coordinates has been estimated to be below 3%[438]. Hence, CCS data from IMS were acquired for Skp and all Skp-OMP complexes identified (Figure 4.7, Table 4.3), and compared with the theoretical CCS value of the Skp crystal structure, or with CCS values for models of the Skp:OMP complexes for which there are no high resolution structural data (Section 4.5).

Figure 4.7 displays the CCS distributions of the observed ions originating from Skp and 1:1 Skp:OMP complexes, normalised to spectral intensity (a representative dataset is shown from three independent experiments). The modal CCSs are also plotted as a function of charge state in Figure 4.7g (Table 4.3). Interestingly, the CCSs of the Skp ions (black squares in Figure 4.7g) are substantially smaller than expected based on the published crystal structure (the modal CCS at the lowest observed charge state, which is least affected by Coulombic repulsion[439], was 37.9 nm^2, approximately 25 % lower than the expected value of 45.7 nm^2 derived from the crystal structure (red dotted line (ii) in Figure 4.7g)). To investigate this, molecular dynamics simulations of apo-Skp *in vacuo* were performed (Figure 4.8), revealing that the assembly collapses in the gas phase, presumably due to the large, empty central cavity present in the structure. The CCS of the structure at the end of the simulation, 37.3 ± 1.9 nm^2 (black dotted line (i) in Figure 4.7g) is in excellent agreement with the CCS measured by IMS (37.9 nm^2, Table 4.4)

Binding of Skp to the 8-stranded OMPs tOmpA and PagP (red and green circles in Figure 4.7g) results in ions with increased CCS compared with Skp alone. The CCSs of the ions observed (45.6 nm^2 and 45.8 nm^2 for Skp-tOmpA and Skp-PagP, respectively, at the lowest observed charge state) are similar to the CCS predicted from the crystal structure of Skp alone (45.7 nm^2), supporting the notion that these assemblies are specific complexes in which the OMP is located within the central Skp cavity[161,171], preventing collapse of the chaperone in the gas phase. Consistent with this, these Skp:OMP complexes sample a narrower conformational ensemble when compared with Skp alone, measured by the width at half height of the mobility peaks observed[440,441], (Figure 4.7a-c). The data suggest that the formation of a complex with tOmpA and PagP packs the hydrophobic cavity of Skp resulting in a narrower conformational ensemble. This is consistent with NMR experiments, which

have demonstrated that Skp binding to tOmpA and OmpX results in stable NMR-visible complexes with reduced Skp flexibility when compared with Skp alone[171].

The 1:1 Skp:OMP assemblies of the larger OMPs studied (OmpT, OmpF and tBamA) (orange, blue and purple circles in Figure 4.7g) have increased CCSs (51.4-54.2 nm^2) compared with the Skp:tOmpA/PagP assemblies (~46 nm^2). These data, together with mismatch between the volume of Skp's cavity and the volume likely to be occupied by larger OMP clients in the 'fluid globule' state[171], suggest that the central cavity expands in size to encapsulate these species, consistent with recent SANS data[172]. However, given that increased Skp:OMP ratios are required to prevent folding of these larger OMPs (Figure 4.3), and that 2:1 Skp:OMP complexes are detected by IMS-MS, these data also indicate that Skp expansion is insufficient to fully sequester these larger OMPs in a 1:1 complex.

The increase in CCS as a function of molecular weight for Skp and all complexes, including 2:1 Skp:OMP assemblies, is plotted in Figure 4.7h. This data can be fitted to a globular model (Section 2.17), irrespective of client size, with the complexes exhibiting an effective gas phase density[375] of 0.33 $Da.Å^{-3}$. This value is similar to those reported for other protein complexes[371], and is consistent with recent calculations of the CCSs of globular proteins in the PDB[375].

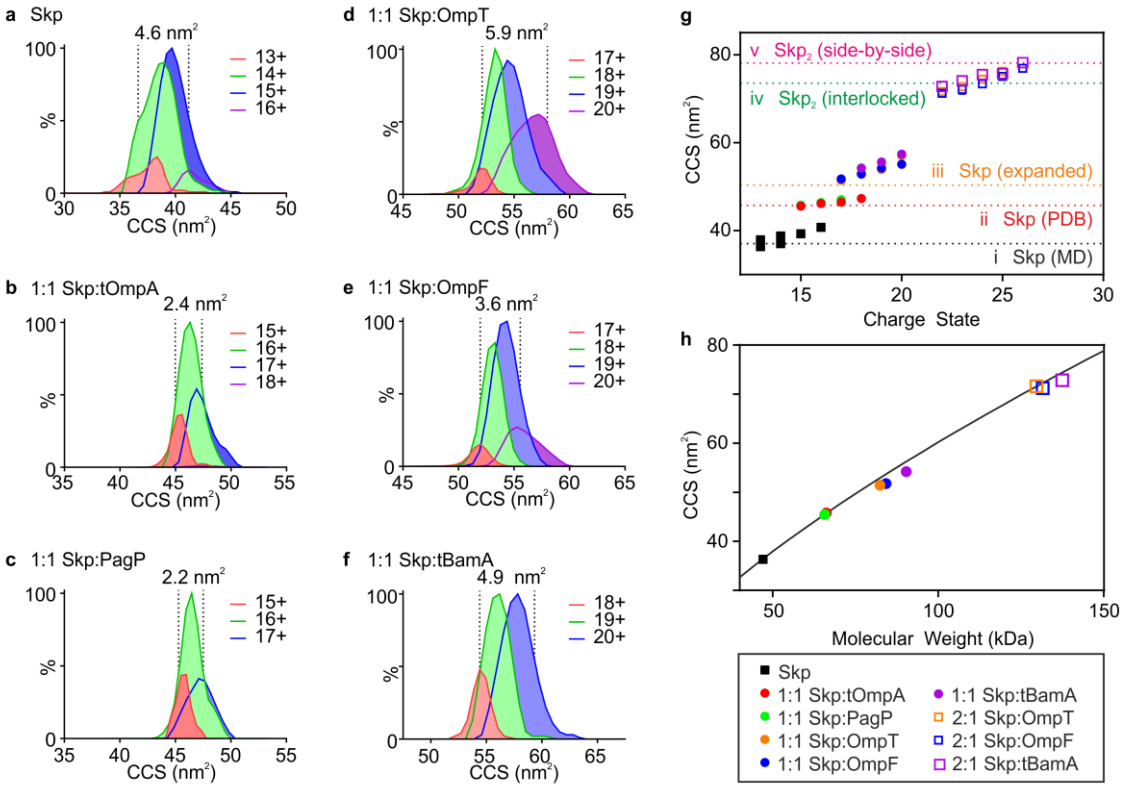


Figure 4.7: Collision cross-section distributions of Skp and Skp:OMP complexes. CCS distributions (peak heights normalized to MS peak intensity) of (a) Skp and 1:1 Skp:OMP complexes with (b) tOmpA, (c) PagP, (d) OmpT, (e) OmpF and (f) tBamA. The width at half height (normalized for spectral intensity) for each distribution is indicated (dashed lines)[440,441], as well as the charge states which are represented by each CCS distribution. (g) Plot of observed CCSs of the assemblies as a function of charge state (mean of three replicates shown, note that standard deviation values are smaller than the symbol size used): Skp, black squares; 1:1 Skp:tOmpA, red circles; 1:1 Skp:PagP, green circles; 1:1 Skp:OmpT, orange circles; 1:1 Skp:OmpF, blue circles; 1:1 Skp:tBamA, purple circles; 2:1 Skp:OmpT, orange open squares; 2:1 Skp:OmpF, blue open squares; 2:1 Skp:tBamA, purple open squares. Dashed lines indicate CCSs estimated from (i) the collapsed structure of Skp from the MD simulation (black) (Figure 4.8), (ii) the crystal structure of Skp (red), (iii) a model 1:1 Skp:OMP complex with an expanded central cavity (orange), (iv) a model 2:1 Skp:OMP complex comprising two interlocking Skp molecules (green) and (v) a model $(Skp)_2$:OMP complex arranged in a side-by-side fashion (pink) (Figure 4.9). (h) Plot of the lowest measured CCS of each assembly as a function of molecular weight (symbols as in (g)). The line represents the expected CCSs for globular proteins with an amino acid density of 0.33 Da.Å^{-3}.

Skp		1:1 Skp:tOmpA		1:1 Skp: PagP		1:1 Skp:OmpT		1:1 Skp:OmpF	
z	**CCS**	***z***	**CCS**	***z***	**CCS**	***z***	**CCS**	***z***	**CCS**
13	36.4 ± 0.2	15	45.6 ± 0.1	15	45.8 ± 0.3	17	51.4 ± 0.8	17	51.8 ± 0.3
13	37.9 ± 0.6	16	46.1 ± 0.3	16	45.8 ± 0.1	18	53.0 ± 0.4	18	52.8 ± 0.6
14	37.1 ± 0.2	17	46.5 ± 0.6	17	45.8 ± 0.2	19	53.9 ± 0.7	19	54.2 ± 0.2
14	38.8 ± 0.3	18	47.3 ±0.2			20	57.1 ± 0.2	20	55.1 ± 0.2
15	39.3 ± 0.6								
16	40.8 ± 0.6								

1:1 Skp:tBamA		2:1 Skp:OmpT		2:1 Skp:OmpF		2:1 Skp:tBamA	
z	**CCS**	***z***	**CCS**	***z***	**CCS**	***z***	**CCS**
18	54.2 ± 0.3	22	71.7 ± 0.04	22	71.2 ± 0.5	22	72.8 ± 0.2
19	55.6 ± 0.8	23	72.7 ± 0.03	23	72.1 ± 0.9	23	74.1 ± 0.1
20	57.4 ± 0.6	24	74.4 ± 0.02	24	73.4 ± 0.4	24	75.5 ± 0.5
		25	76.2 ± 0.4	25	75.1 ± 0.1	25	75.7 ± 0.3
				26	76.9 ± 0.6	26	78.20 ± 0.1

Table 4.3: Experimentally determined modal collision cross-sections (CCSs) (nm^2) for the observed charge states (*z*) of Skp and Skp:OMP complexes. CCS values are shown as mean ± standard deviation of three independent measurements.

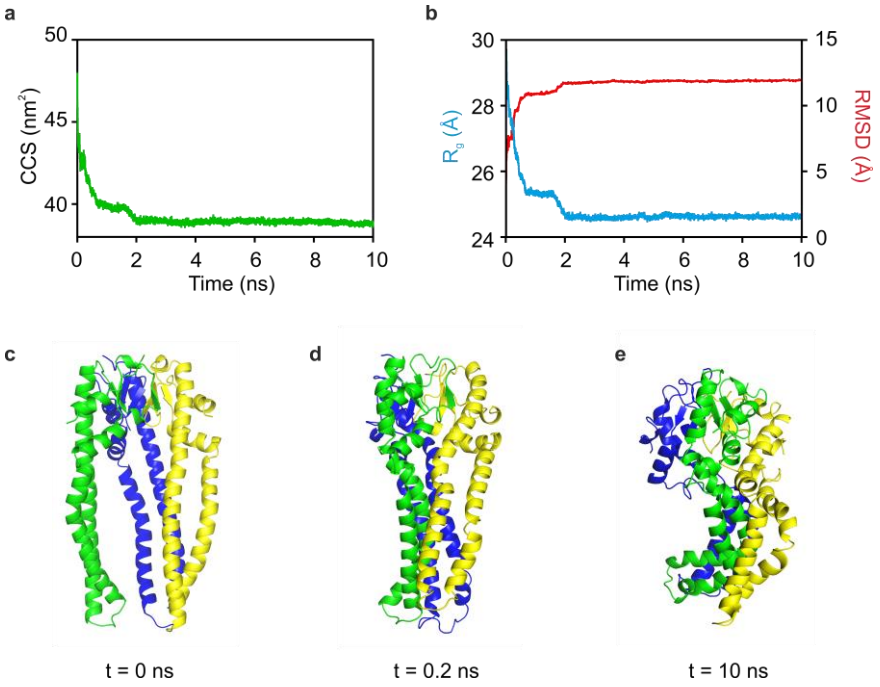


Figure 4.8: *In vacuo* molecular dynamics simulations of Skp. (a) Theoretical CCS, and (b) backbone RMSD (red line) and radius of gyration (R_g) (light blue line) calculated for the initial 10 ns of the 100 ns simulation in the absence of solvent. (c) The starting model Skp structure used for the MD simulation (PDB: 1U2M[150], with missing residues in chains B and C modelled from chain A). Skp subunits are coloured green, blue and yellow. (d,e) Structures of Skp after a simulation time of (d) 0.2 ns and (e) 10 ns. The CCS of the collapsed Skp structure after a simulation time of 100 ns simulation (37.3 ± 1.9 nm^2, Table 4.4) agrees favourably with the modal CCS of Skp at the lowest observed charge state (37.9 ± 0.6 nm^2).

4.5 Modelling of larger OMPs in complex with two copies of Skp gives structures with different architectures

To investigate the architecture(s) of Skp complexes with OmpT, tBamA or OmpF, four different models were constructed and the theoretical CCSs of the resulting structures compared with the experimentally determined CCS. As a starting point, encapsulated tOmpA was modelled as a sphere with a radius of 20 Å (Figure 4.9a) which has a volume of 33,500 $Å^3$, consistent with previous estimates from the crystal stucture[150], and with NMR data for the 8-stranded OmpX in complex with Skp[171]. It was assumed that the amino acid density for non-native OMPs bound to Skp is independent of the mass of the OMP studied, consistent with the MS data presented above (Figure 4.7h). Therefore, to model a 16-stranded OMP, a spherical volume of ~67,000 $Å^3$ was assumed giving a radius of ~25 Å, and generated a Skp model with its three subunits surrounding a sphere of this size (Figure 4.9b). The

theoretical CCS of the resulting structure (50.4 nm^2) (orange dotted line (iii) in Figure 4.7g) is in good agreement with the measured CCS values for 1:1 Skp:OmpT/OmpF/tBamA complexes (51.4, 51.8 and 54.2 nm^2, respectively) (Table 4.3). The results suggest, therefore, that 1:1 Skp:OMP complexes with larger OMPs, involve an expanded Skp cavity.

Next, models were generated of the 2:1 Skp:OMP complexes with theoretical CCS values consistent with those measured by IMS. It was theorised that Skp could arrange in a side-by-side configuration, in either a parallel or antiparallel arrangement (Figure 4.9c,d), with the OMP represented by a capsule with a cylinder height of 37 Å and cap radii of 20 Å. The theoretical CCSs of these assemblies were determined to be 79.1 and 78.2 nm^2, respectively (green dotted line (iv) in Figure 4.7g). Alternatively, a model in which the OMP substrate (represented by a sphere with a radius of ~25 Å) may be encapsulated by two interlocked copies of Skp (Figure 4.9e) was considered, which results in a complex with a theoretical CCS of 73.5 nm^2 (pink dotted line (v) in Figure 4.7g). All three values are in good agreement with the measured CCS values for 2:1 Skp:OmpT/OmpF/tBamA complexes (71.7, 71.2 and 72.8 nm^2, respectively) (Table 4.3).

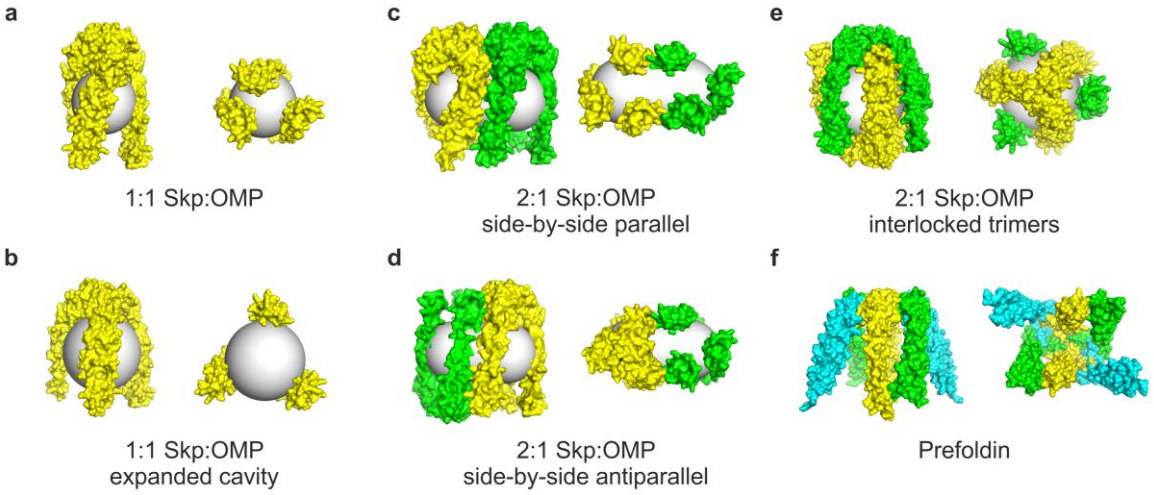


Figure 4.9: Possible architectures of Skp:OMP complexes. Side (left) and bottom (right) view surface representations of models of (a) Skp (yellow) (based on the published crystal structure (PDB: 1U2M[150]) with missing residues modelled by molecular replacement), bound to an 8-stranded OMP represented by a grey sphere of radius 20 Å, (b) Skp with an expanded central cavity surrounding a 16-stranded OMP represented by a grey sphere of radius 25 Å, (c,d) 2:1 Skp:OMP structures (Skp coloured in yellow and green) arranged side-by-side in a (c) parallel or (d) antiparallel arrangement, with the 16-stranded OMP substrate represented by a grey capsule with cylinder height of 37 Å and cap radii of 20 Å, (e) 2:1 Skp:OMP complex with an interlocked trimer architecture with the 16-stranded OMP represented by a grey sphere of radius 25 Å, and (f) the hexameric eukaryotic chaperone prefoldin (PDB: 1FXK[174]), with chains A, B and C of the crystal structure shown in green, cyan and yellow, respectively.

4.6 Investigation of Skp:OMP complex models by molecular dynamics simulations

To model these Skp:OMP complexes further, and aid their visualisation, a series of molecular dynamics simulations were performed. A simulation of apo-Skp in explicit solvent over 100 ns demonstrated that the individual subunits are highly dynamic and flexible. Each subunit undergoes a transition to an 'open' state, in which subunit helices splay from the central axis, resulting in an expanded central cavity, consistent with previous MD studies[156,172] (Figure 4.10). The average radius of gyration (R_g) of Skp from the simulation (31.5 Å) is in good agreement with published SANS data (~33 Å)[172]. This R_g value is ~10 % higher than predicted from the crystal structure, indicating that dynamic motions of the Skp subunits observed in the simulation likely reflect those observed in solution. In the two Skp crystal structures solved to date[149,150], the lower section of one of the subunits is

unresolved, indicating flexibility, and the angles with which the Skp subunits extend away from the 'body' domain are different for each of the subunits in the two structures[149,150]. The individual subunits in the crystal structure of the heterohexameric eukaryotic chaperone prefoldin (Figure 4.9f) also make different angles with respect to the multimerisation domain, which has also been suggested to indicate conformational flexibility that may be functionally relevant[174].

Next, models were generated of tOmpA (8-stranded) and tBamA (16-stranded) alone in an unfolded, extended conformation, and their behaviour simulated in solvent (mimicking the situation in which OMPs are diluted from 8 M urea). In each case, the OMPs collapse rapidly to an approximately globular form (Figure 4.10d,e).

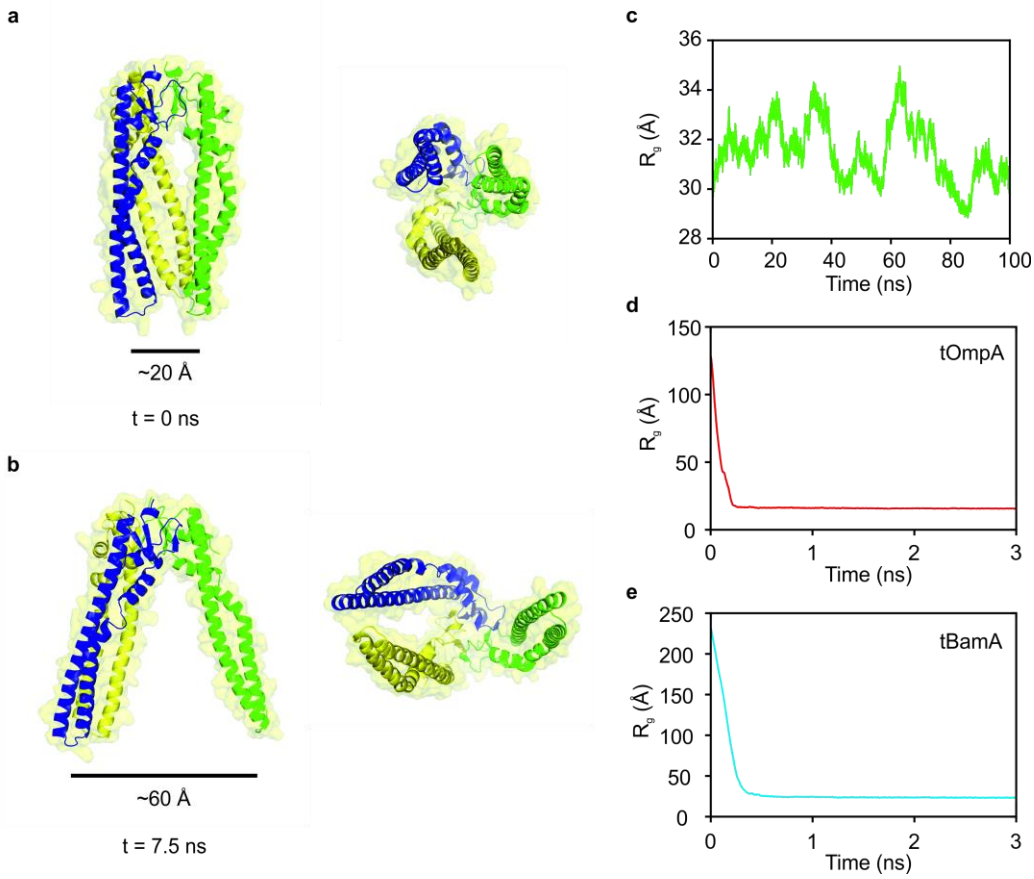


Figure 4.10: Molecular dynamics simulations of Skp, tOmpA and tBamA in solvent. (a) Starting structure used for explicit solvent MD simulation of Skp (PDB: 1U2M[150], with missing residues in chains B and C modelled from chain A), shown from the side (left) and bottom (right). (b) Example structure of Skp in an 'open' conformation (t = 7.5 ns), shown from the side (left) and bottom (right). Skp subunits in (a) and (b) are coloured green, blue and yellow. (c) Radius of gyration of Skp over the course a MD simulation in explicit solvent. Structural collapse of initially extended chains of (d) tOmpA and (e) tBamA simulated with an implicit solvent model (Section 2.19).

A model of the 1:1 Skp:tOmpA complex was then generated by placing the collapsed tOmpA structure within the cavity of Skp in an 'open' conformation from the simulation of apo-Skp (Figure 4.10b) and relaxing the resulting structure *in vacuo*. In the simulation, the subunits of Skp collapse rapidly around the tOmpA substrate resulting in a structure with a CCS value (43.7 ± 1.2 nm^2) in excellent agreement with that measured by ESI-IMS-MS (45.6 ± 0.1 nm^2), suggesting that, at least in the gas phase, Skp "clamps" around the substrate (Figure 4.11a,b). To provide insight into the potential atomistic structure of the 2:1 Skp:OMP complexes observed using ESI-MS (Figure 4.5), and to aid visualisation of the size of Skp relative to its 16-stranded OMP clients, a model for the gas-phase 2:1 Skp:tBamA

complex was created by placing the collapsed tBamA structure in the hydrophobic cavity formed by two copies of Skp in their 'open' conformations. A side-by-side parallel orientation (Figure 4.9c) was chosen based on the striking resemblance of this model to the structure the eukaryotic prefoldin chaperone[174] (Figure 4.9f). The size of the collapsed tBamA model clearly exceeds the maximal dimensions of the cavity of a single Skp observed in an 'open' conformation (Figure 4.11c). Simulation of the 2:1 Skp:tBamA complex *in vacuo* showed that the Skp subunits also rapidly "clamp" around the tBamA substrate creating a complex with a CCS value (74.4 ± 1.4 nm^2) again in good agreement with the IMS data (72.8 ± 0.2 nm^2) (Figure 4.11d). Therefore, the CCS data obtained from both experiment and simulation (Table 4.4) are consistent with a model in which multivalent Skp binding is necessary to sequester OMPs that exceed the dimensions of the Skp cavity, to prevent their aggregation.

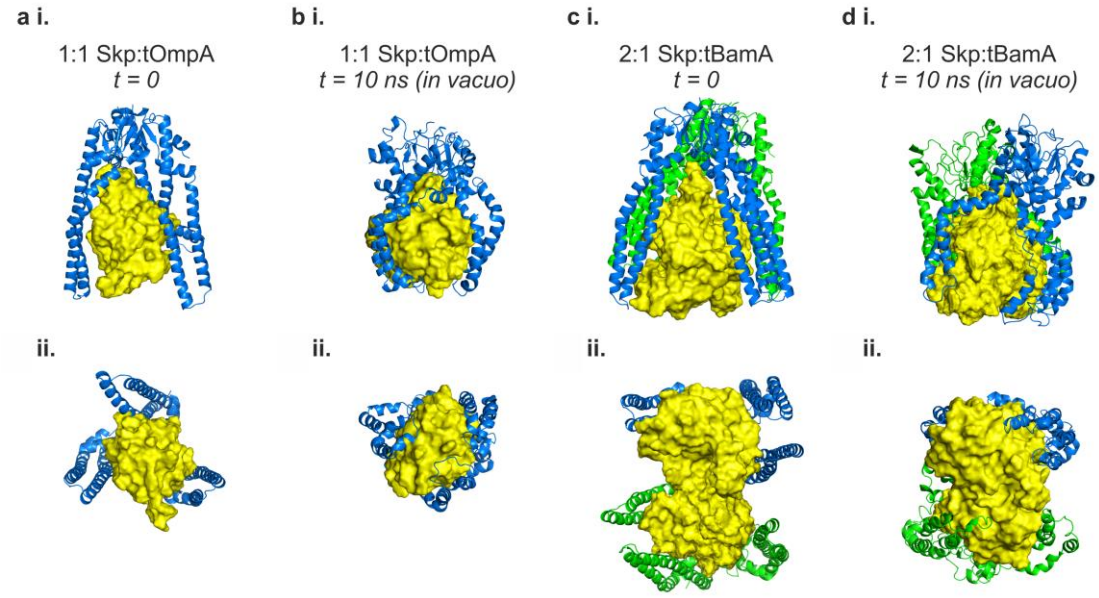


Figure 4.11: *In vacuo* molecular dynamics simulations of 1:1 and 2:1 Skp-OMP complexes. (a) Starting model of a 1:1 Skp:tOmpA complex used for MD simulations, and (b) the structure obtained after 10 ns of *in vacuo* simulation. (c) Starting model of a 2:1 Skp:tBamA complex used for MD simulations, with the two copies of Skp arranged in a side-by-side parallel orientation (Figure 4.9c), and (d) the structure obtained after 10 ns of *in vacuo* simulation. Views from the (i) side and (ii) bottom are shown. Skp (green/blue) is shown in cartoon representation. OMPs (yellow) are shown in surface representation

To provide evidence that a similar "clamping" motion of Skp around its OMP substrates could occur in solution, analogous MD simulations of Skp:tOmpA and 2:1

Skp:tBamA complexes were performed in explicit solvent. In these simulations Skp subunits are also observed wrapping around their OMP substrates (Figure 4.12a,b), with ‘clamping’ movements similar to those observed in the gas-phase simulations (Figure 4.11). The complexes formed are stable over 100 ns (Figure 4.12c,d) and have larger calculated CCS values (56.5 ± 0.3 nm^2 and 101.2 ± 6.0 nm^2 for Skp-tOmpA and 2:1 Skp-tBamA, respectively) than those following gas-phase simulation (43.7 ± 1.2 nm^2 and 74.4 ± 1.4 nm^2 for Skp-tOmpA and 2:1 Skp-tBamA, respectively). These data are consistent with a model in which the subunits of Skp are dynamic, resulting in expansion of the hydrophobic cavity that allows entry of substrates of varying sizes, prior to the Skp subunits ‘wrapping’ around the sequestered client to protect it from aggregration until folding into the bilayer can take place.

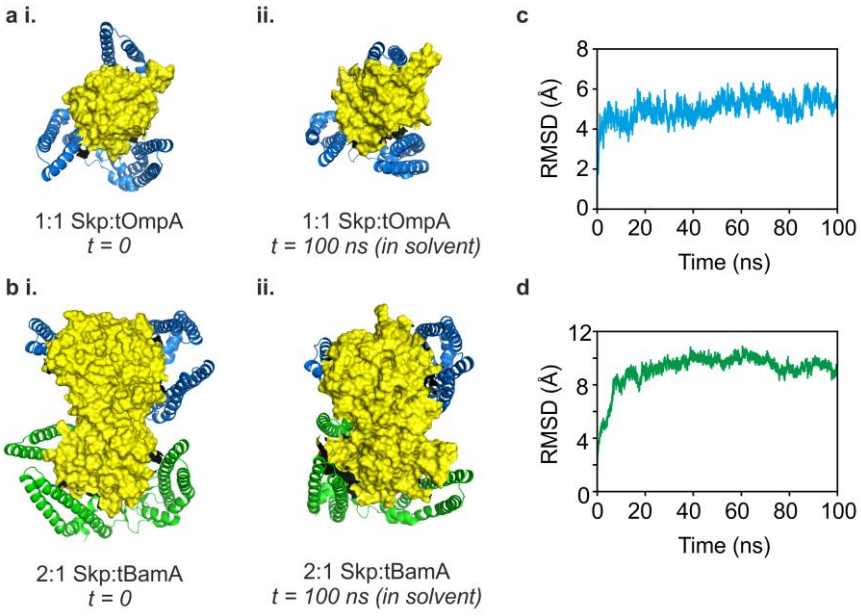


Figure 4.12: Molecular dynamics simulations of 1:1 and 2:1 Skp:OMP complexes in explicit solvent. (a) (i) Starting model of a 1:1 Skp:tOmpA complex used for MD simulations, and (a) (ii) the structure obtained after 100 ns of simulation in explicit solvent. (b) (i) Starting model of a 2:1 Skp:tBamA complex used for MD simulations, with the two copies of Skp arranged in a side-by-side parallel orientation (Figure 4.9c), and (b) (ii) the structure obtained after 100 ns of simulation in explicit solvent. (c,d) Backbone RMSDs calculated for the 100 ns simulations of (c) 1:1 Skp:tOmpA, and (d) 2:1 Skp:tBamA in explicit solvent, demonstrating that the complexes are stable over this timescale.

Protein(s)	Calculated CCS from simulation (nm^2)	CCS measured by IMS-MS (nm^2)
tOmpA alone	19.0 ± 0.06	-
tBamA alone	37.9 ± 0.9	-
1:1 Skp-tOmpA (*in vacuo*)	43.7 ± 1.2	45.6 ± 0.1
2:1 Skp:tBamA (*in vacuo*)	74.4 ± 1.4	72.8 ± 0.2
Skp (*in vacuo*)	37.3 ± 1.9	37.9 ± 0.6
1:1 Skp-tOmpA (H_2O)	56.5 ± 0.3	-
2:1 Skp:tBamA (H_2O)	101.2 ± 6.0	-

Table 4.4: Comparison of CCS values from MD simulations and IMS-MS measurements. CCS values from IMS-MS are shown as the mean ± standard deviation of the modal value of the lowest charge state ion from three independent measurements. CCS values from the simulations are the mean ± standard deviation for the endpoint structures at 100 ns from three independent simulations. For globular assemblies, the error between measured CCS values and those computed from atomic coordinates has been estimated to be below 3%[438].

4.7 Discussion

Major advances in the understanding of the roles of molecular chaperones and folding catalysts involved in OMP biogenesis have been made in recent years[17], yet the molecular details of how OMPs are bound by molecular chaperones, transported across the periplasm and assembled into the outer membrane, all without using the energy of ATP binding/hydrolysis, remain unclear[10,14,16]. In this chapter, new insights have been provided into how Skp is able to chaperone its broad array of OMP clients, including substrates which are too large to be accommodated within its hydrophobic cavity. The results show that Skp utilises subunit dynamics to expand the size of its client binding cavity and demonstrate that Skp can function as a multivalent OMP chaperone in order to sequester and prevent aggregation of its larger OMP clients. Further, ESI-IMS-MS was used to gain structural insight into the 1:1 and 2:1 Skp:OMP complexes identified. Using kinetic refolding and ESI-IMS-MS data, combined with MD simulations, models consistent with the experimental results have been presented, in which Skp sequesters larger OMPs by binding in a multivalent arrangement (side-by-side parallel or anti-parallel, and/or via an interlocking structure) (Figure 4.9c-e). Interestingly, the parallel side-by-side model (Figure 4.9c) bears a striking resemblance to the structure of the non-homologous chaperone prefoldin (Figure 4.9f)[174]. The precise orientation(s) of Skp molecules in these multivalent complexes cannot be extracted from the experimentally obtained CCS values obtained. Indeed, the models of 2:1 Skp:OMP complexes generated (Figure 4.9c-e) are not mutually exclusive and given the dynamic nature of Skp and the wider range of CCS values obtained for Skp in complex with larger OMPs (Figure 4.4d-f) it is possible many different configurations exist in solution. Additional information from further studies, for example using cross-linking experiments followed by MS/MS, will be required to obtain a more complete structural picture. Nevertheless, the kinetic, MS and MD results in this chapter indicate that the ability of Skp to chaperone substrates ranging from 35-43 kDa in size requires both subunit dynamics and its ability to function as a multivalent chaperone. Skp has been shown *in vivo* to interact with much larger substrates than those investigated here (19-43 kDa), including BtuB (66 kDa), FhuA (79 kDa) and LptD (87 kDa) which form β-barrels composed of 22, 22, and 26 β-strands, respectively[155,178]. It is likely these proteins also form multivalent complexes with Skp, and indeed recent experiments on a Skp:FhuA complex, using analytic SEC, are consistent with a greater than 1:1 Skp:OMP stoichiometry[97].

The results presented have implications for understanding how OMPs are chaperoned by Skp in the periplasm, including the mechanisms of substrate binding and release. The atomistic MD trajectories of apo-Skp show that it exists in a wide range of open conformations with large differences in the area of the cavity entrance formed between the tips of its three subunits (Figure 4.10a-c). Such conformational flexibility has been implicated in the mechanisms of other ATP-independent chaperones[168,442]. Therefore, in rescuing aggregation-prone proteins[145] Skp could be thought of as analogous to a pair of 'callipers', sampling open conformations prior to capture of its client, allowing it to adjust the volume of its central cavity. In this model once the substrate has entered the Skp cavity, the Skp subunits 'clamp down' to protect the exposed hydrophobic surfaces of the protein, in a mechanism similar to those of other chaperones such as Trigger Factor or Hsp90[443], as observed in the simulations of Skp:tOmpA and Skp:tBamA complexes (Figure 4.12). For those substrates which are too large to be accommodated within the Skp substrate cavity, additional copies of Skp recognise and engulf sections of the substrate not already encapsulated. *In vivo* cross-linking evidence suggests that Skp can interact with OMPs as they emerge from the SecYEG translocon[148]. Thus it is possible that during translocation of larger OMPs, the substrate is fed directly into the cavity of Skp and the chaperone's maximum binding capacity would then be reached before the complete polypeptide chain is translocated. Subsequent polypeptide chain emerging from the translocon would then be bound by a second or more Skp(s), ensuring sequestration of the entire polypeptide sequence so that periplasmic aggregation is prevented.

Recent equilibrium sedimentation experiments of Skp in the absence of substrate have demonstrated a dynamic equilibrium between folded subunit monomers and trimers at physiological concentrations[163]. Therefore, a possible alternative *in vivo* pathway to Skp-OMP complex formation may involve sequential binding of monomer subunits to OMP substrates, with Skp trimerisation linked to (and indeed driven by) subunit binding to substrate. Sandlin *et al*. raise the possibility that more diverse species of Skp could form around an unfolded OMP[163]. However, for the complexes studied here only Skp:OMP complexes containing monomeric Skp subunits in multiples of three are observed (Figure 4.3), suggesting that either the trimeric unit is the OMP binding species, or that complexes between OMPs and non-trimeric Skp subunits are unstable in the gas-phase.

It has been proposed that transient exposure of the C-terminal OMP targeting sequence[129,444] (Section 1.2.2), and its recognition by the BAM complex[136], triggers substrate release from Skp[171]. The space between Skp subunits (~25 Å in the crystal structure[150]) and/or the inherent dynamics of the complex may facilitate the transient solvent-exposure of regions of the OMP substrate, permitting β-signal recognition. Consistent with this, a study found that the presence of BamA in DUPC liposomes relieves the *in vitro* folding inhibition of OmpA by Skp[299]. Interestingly, in all 2:1 Skp:OMP models proposed here there remains a substantial distance (>20 Å) between the Skp subunits, which would permit exposure of sections of the substrate polypeptide required for BAM signalling and/or membrane insertion. The release of the OMP from Skp is likely to be driven by the increased thermodynamic stability of the folded OMP relative to the chaperone-bound state[445], and for OMPs that are bound to more than one Skp, it is possible that individual copies of Skp are released sequentially in a process driven by the free energy of OMP folding. Kinetic modelling of the flux of OMPs through the periplasm suggested that the population of Skp-OMP complexes is dynamic[124]. Association rates (k_{on}) between Skp and its OMP clients are diffusion limited, and dissociation rates (k_{off}) are on the millisecond time-scale, much shorter than the time taken for an OMP to cross the periplasm (~2 min)[123,124]. The authors suggest that rapid binding and unbinding of Skp to its OMP clients may promote conformations which favour folding, and disfavour aggregation[124]. Single molecule force spectroscopy (SMFS) experiments in which FhuA was refolded in the presence or absence of Skp or SurA also suggested that periplasmic chaperones have a role in modulating the folding pathway to promote folding[97].

Chaperones utilise two general strategies to protect substrates from misfolding and aggregation. In the first, substrates are chaperoned by sequential binding and release of exposed hydrophobic surfaces along an extended polypeptide chain. This 'beads on a string' model is typified by chaperones such as the Hsp70s and Trigger Factor[446,447]. Alternatively, aberrant interactions may be prevented by sequestration of the substrate from the cellular environment within an enclosed space, as is the case for the 'cage-like' chaperonins such as GroEL/ES and TriC[446,447]. The data presented here suggest that Skp operates with a 'hybrid' mechanism, employing both of these strategies to bind and encapsulate its OMP clients, thereby preventing their aggregation and facilitating their delivery to the OM.

In summary, the combined results in this chapter provide a new understanding of how Skp is able to bind, chaperone and release substrates that vary dramatically in size, with expansion of the binding cage and/or formation of multivalent complexes allowing the chaperone to adapt to the demands of its clients. It also introduces native-MS as a new method for studying membrane protein-chaperone complexes. This technique has exciting prospects for examining the structure and dynamics of other MP-chaperone systems, particularly when, as here, it is used in combination with MD simulations. The following chapter considers the next stage in the OMP biogenesis pathway, when OMPs are released by periplasmic chaperones, such as Skp, and subsequently interact with the BamA folding catalyst.

5 Effects of periplasmic chaperones and membrane thickness on BamA catalysed OMP folding

5.1 Introduction

In the previous chapter the interactions and effects on *in vitro* folding of a single folding factor (the Skp chaperone) with a number of substrate OMPs was investigated. Here, work is described which takes the complementary approach of considering a single substrate, tOmpA, and examining the influence of multiple folding factors on its folding kinetics. Since this thesis commenced, several SDS-PAGE based studies have shown that BamA alone is able to accelerate OMP folding *in vitro*[299,300,302]. This chapter will begin by recapitulating this result using fluorescent spectroscopy. It then examines the effects of the periplasmic chaperones Skp and SurA on BamA-catalysed folding reactions. Finally, the effects of the hydrophobic thickness of the bilayer environment on BamA activity are analysed.

The molecular details of how OMPs are transported across the periplasm and delivered to the BAM complex remain unresolved[10]. Genetic studies suggest that the two major OMP chaperones in *E. coli*, Skp and SurA, operate in parallel pathways[145,147]. However, while Skp has been cross-linked at the inner membrane[148], no direct *in vivo* cross-linking of Skp to BAM components has been reported. By contrast, SurA is readily cross-linked to BAM components[145,187,188], supporting the notion that Skp and SurA may cooperate in a sequential pathway in which Skp interacts early in the pathway and hands over substrates to SurA for delivery to BAM[68,176]. An alternative pathway in which Skp delivers substrates directly to the OM is supported by *in vitro* data which have shown that the highly positively charged Skp can deliver OMPs directly to negatively charged bilayers[152,153], and recent data which suggest that Skp can deliver substrates into membranes *in vivo*[154]. By systematically examining the effects on tOmpA folding of combinations of chaperones and BamA, the data in this chapter suggest that Skp may directly deliver substrates to BAM, and provides evidence against a sequential model for the Skp and SurA chaperones.

Many studies have shown that OMPs are able to spontaneously fold and insert into lipid bilayers[298], and that the folded state is highly stable[448]. However, *in vivo* the

kinetics of spontaneous folding into the membrane environment of the OM is too slow to be biologically relevant and therefore a dedicated catalyst, the BAM complex is required[15]. Therefore, the role of the BAM complex *in vivo* is to modulate the kinetics of an already thermodynamically favourable process[445]. So what are the factors that increase the kinetic barriers to OMP folding? *In vitro* evidence points to a role for the physicochemical properties of the lipids themselves, as OMP folding kinetics and efficiency *in vitro* are influenced by lipid curvature, the degree of saturation in acyl chain lengths, the nature of lipid head groups, and the hydrophobic thickness of the membrane[322]. BamA, in the absence of the other BAM lipoproteins, was first shown to accelerate OmpA folding in DLPC ($di_{C12:0}$PC) bilayers[299]. Subsequently, it was demonstrated that the presence of prefolded BamA in DDPC ($di_{C10:0}$PC) liposomes containing a mole percentage of 20% DDPE ($di_{C10:0}$PE) or DDPG ($di_{C10:0}$PG) leads to a much greater difference in OMP folding kinetics between the catalysed and uncatalysed folding reactions[300]. Therefore, it was proposed that one function of BamA is to overcome the kinetic barrier to folding imposed by native lipid head groups[300].

Another kinetic barrier to OMP folding is imposed by the hydrophobic thickness of the membrane, and it as been suggested that the decreased thickness of the OM compared to the IM may be one mechanism by which OMPs are sorted to the correct membrane *in vivo*[53,336]. The results in Chapter 3 showed that the OMP PagP folds fastest in thinner bilayers, which follows the same trend as that observed for other OMPs[313]. The acyl chain lengths of lipids in the OM are mostly longer (C14-C18[34]) than commonly used in *in vitro* studies, and accordingly, most OMPs cannot be folded directly into liposomes composed of *E. coli* polar lipid extract[300,313]. For those OMPs which are able to fold into these native lipids, folding does not occur on biologically relevant timescales[300,313]. Therefore, it is likely that one strategy by which BAM aids OMP folding *in vivo* is to locally thin the bilayer. In support of this, the crystal structure of BamA revealed an asymmetric β-barrel, with a thinner aromatic girdle on the side of the barrel closest to the β1-β16 seam (Figure 5.1)[11]. By contrast, OmpA, BtuB and the Omp85 family member FhaC have a similar hydrophobic thickness all the way around the barrel (Figure 5.1).

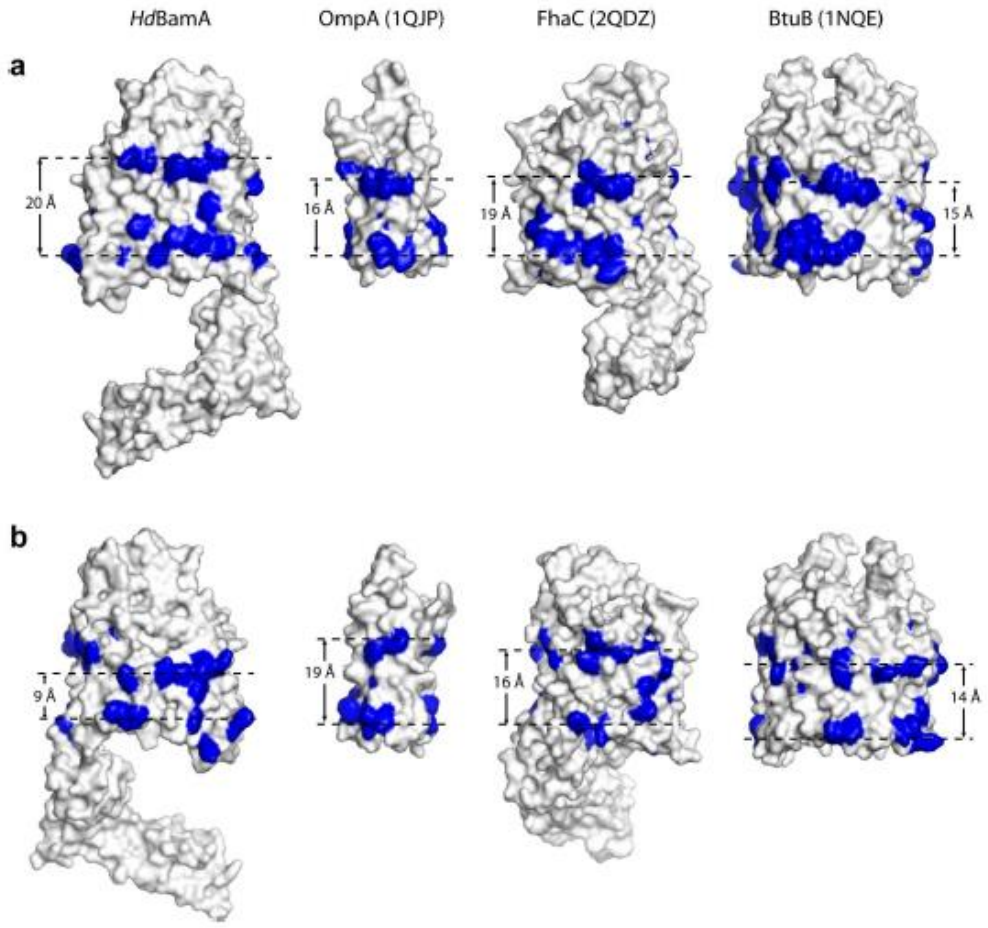


Figure 5.1: The BamA β-barrel has a reduced hydrophobic thickness on one side close to the β1 and β16 interface. (a) The location of aromatic girdle residues (Trp, Tyr and Phe, highlighted in blue) in BamA (*H. ducreyi*, PDB: 4K3C[11]*)*, OmpA (*E. coli*, PDB: 1QJP[303]), FhaC (*B. pertussis*, PDB: 2QDZ[449]) and BtuB (*E. coli*, PDB: 1NQE[450]), (b) the same structures rotated 180º around the y-axis. Image reproduced from Noinaj *et al*. (2013)[11].

Further, a simulation of the β-barrel domain of BamA from *N. gonorrhoeae* in a DMPE bilayer, exhibited dramatic membrane thinning of 16 Å close to β16[11]. However, in recent simulations of full-length *E. coli* BamA in an atomistic outer membrane, thinning of only a few angstroms was observed, and this was not always localised to the β1-β16 interface[262]. Additionally, a similar extent of membrane thinning was observed in simulations of *E. coli* OmpLA in an all-atom native membrane[451]. Membrane thinning of lipid bilayers by clusters of BamA molecules was proposed to have been observed by negative stain EM[452]. However, very few EM images were shown of the notches and distortions in BamA-containing

liposomes suggested to be caused by BamA, and no quantitative analysis was presented[452].

Related to the membrane thinning hypothesis is the idea that BAM may act as a 'disruptase' creating areas of local defects within the membrane to enhance folding rates[445]. A study on the permeation of small ions (protons and potassium ions) through bilayers composed of unsaturated PC lipids concluded that the diffusion of ions is aided by the creation of transient pores, but only if the membrane is sufficiently thin (acyl chain lengths of <C18 and <C22 were required for the permeation via pores of potassium ions and protons, respectively)[453]. Further, atomistic simulations showed that the free energy penalty for pore formation was much lower in shorter DLPC bilayers (C12) than those composed of DMPC (C14) or DPPC (C16)[454]. Direct evidence that membrane defects accelerate OMP folding was recently provided by experiments in which the kinetics of folding of tOmpA into DMPC LUVs was monitored at, or close to, the transition temperature (T_m) of the bilayer[365]. AFM studies have shown that at the DMPC T_m (24 ºC) liquid and gel phase domains coexist[455,456], thus creating defects in the membrane and increasing its permeability[365]. The kinetics of tOmpA were substantially faster (>10-fold) at the T_m than at a temperature 2 ºC higher[365]. It was proposed that this effect is due to the hydrophobic mismatch between lipids at the boundaries of these two phases creating transient defects and pores, which OMPs utilise to aid insertion and folding[365]. These hypothesised 'membrane-based' mechanisms of OMP folding catalysis are not mutually exclusive with other proposed models of BAM function. For example, they do not provide evidence for or against the currently favoured 'BamA-budding' and 'BamA-assisted'[422] models (Section 1.4.7), which are both consistent with a role for BAM in altering membrane properties to aid folding[445].

In this study it was decided to examine the effects of the central component BamA in assisting OMP folding in the absence of the other BAM lipoproteins for several reasons: (1) BamA is the central platform for the other BAM components[16]; (2) it has been shown that BamA can be refolded *in vitro* from inclusion bodies[129,313]; (3) there is evidence of direct interaction of BamA with substrates *in vitro*[129,444]; (3) a catalytic effect on OMP folding has been demonstrated for BamA in the absence of other BAM subunits[299,300,302]; and (4) BamA is the only BAM complex member for which homologs have been found in all known Gram-negative bacterial genomes[224,457]. BamB-E have varying degrees of conservation across Gram-negative bacteria[224]; no homologs of BAM lipoproteins have been found in *Thermus thermophilus*[224,243], or in

the cyanobacterium *Thermosynechococcus elongates*[224,244]. Interestingly, the BamA homologue from *T. thermophilus* appears to be able to aid assembly of one of its substrates (TtoA) *in vitro* in the absence of other factors[444,458].

In this chapter a systematic and stepwise strategy is taken towards gaining understanding of the function of BamA using kinetic assays. The results presented for BamA, in combination with different chaperones, and in liposomes of different composition, also lay a foundation for future work involving the remaining BAM lipoproteins.

In the work presented in this chapter the N-terminal cysteine tOmpA mutant used in microscale thermophoresis (MST) experiments was mutated, expressed and purified by Tom Crosskey (University of Leeds). The temperature denaturation of BamA was performed by Anna Higgins (University of Leeds).

5.2 BamA has a catalytic effect on OMP folding in DUPC liposomes

Cold SDS-PAGE based studies have demonstrated that BamA alone can increase the kinetics of OmpA folding in DLPC LUVs[299], and of tOmpA, OmpA, OmpX, and OmpLA folding in LUVs composed an 80:20 DDPC:DDPE mix[300,302]. Therefore, whether BamA could also accelerate the folding of tOmpA into DUPC liposomes was first investigated, using the fluorescence assay developed in Chapters 3 and 4. To examine the effects of BamA on tOmpA folding *in vitro* three separate BamA constructs were produced: (1) the soluble POTRA domains of BamA (residues 21-424), (2) the β-barrel domain of BamA (tBamA) (residues 425-810), and (3) full-length BamA (residues 21-810) (Figure 1.15). Initially, the POTRA domains were cloned into pET11a, however this construct could not be expressed despite trials in BL21(DE3), BL21(DE3)Star, BL21-CodonPlus(DE3)-RIPL, BL21(DE3) BLR(pLysS), and Rosetta-2(DE3)pLysS expression strains (data not shown). The addition of an N-terminal His-tag allowed expression, but this construct was unable to be bound by Ni^{2+} beads or His-Trap columns, despite numerous attempts (data not shown). It is likely that self-association of the POTRA domains obscured the His-tag, consistent with previous observations of *in vitro* self-association of full-length BamA[129]. Finally, the BamA POTRA domains were cloned and expressed as a fusion with maltose-binding protein (MBP) (Section 2.4), and pure protein obtained following removal of MBP by TEV cleavage (Section 2.4, Figure 5.2a, upper). To compare the effects of BamA on OMP folding in the presence of chaperones, His-tagged constructs of Skp

and SurA were expressed and purified (Sections 2.6 and 2.7, respectively, Figure 5.2b,c, upper). Far-UV CD spectra indicated the BamA POTRA domains, Skp and SurA have native-like secondary structure under the conditions used in kinetic experiments (Figure 5.2, lower).

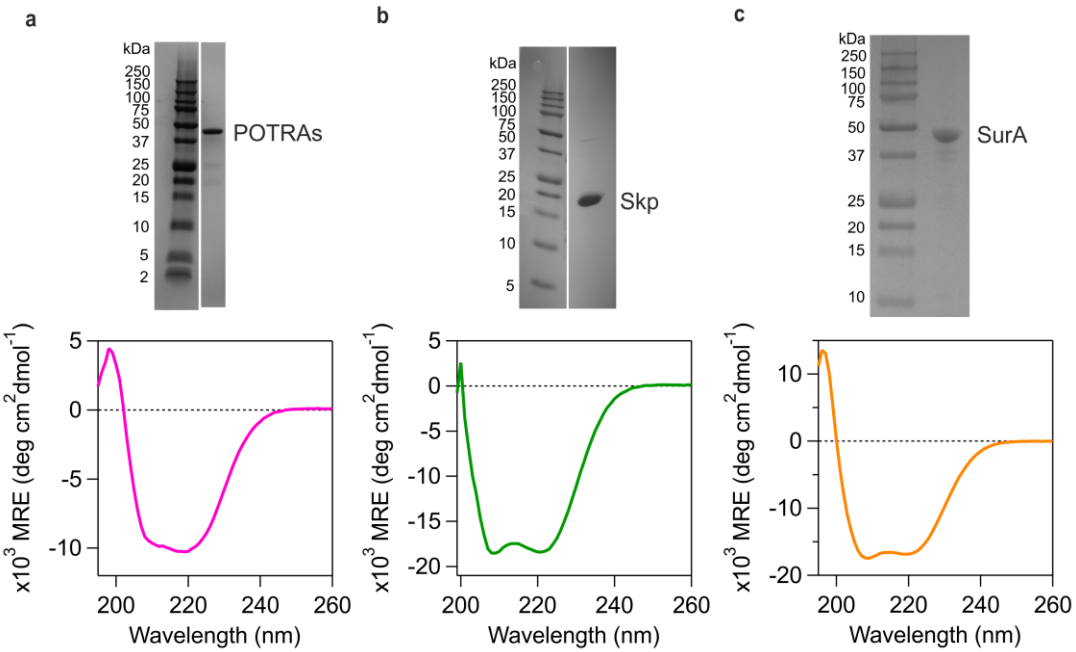


Figure 5.2: Purification of BamA POTRA domains, Skp and SurA. SDS-PAGE purification gels (top) and CD spectra (bottom) for (a) BamA POTRA domains 1-5, (b) Skp, and (c) SurA. CD spectra where acquired in 50 mM glycine-NaOH, pH 9.5, at 25 ºC. Protein concentrations in CD samples for BamA POTRA domains, Skp and SurA were 5 μM, 3 μM (trimer equivalent), and 5 μM, respectively. CD spectra of Skp and SurA are of constructs lacking a His-tag.

To ensure that any effect on tOmpA folding kinetics by the presence of BamA was specific to BamA and not simply due the presence of a prefolded OMP in the bilayer, full-length OmpA was also expressed and purified for use as a control. OmpA was selected for several reasons: (1) OmpA, like BamA, has a periplasmic domain that makes up ~50% of the protein; (2) OmpA has been shown to fold into liposomes in an orientation in which the soluble domain faces outwards[297], as has also been demonstrated for BamA[299]; and (3) OmpA and BamA have similar theoretical pIs (5.59 and 4.87 for OmpA and BamA, respectively[364]), and (4) OmpA and BamA have been shown to fold in PC containing LUVs with similar efficiencies[313]. To verify that BamA, tBamA, OmpA, and the substrate tOmpA, are able to fold into DUPC liposomes under the experimental conditions used, fluorescence emission spectra

were acquired for each OMP either unfolded in 8 M urea, or in 0.24 M urea following a folding reaction in the presence of DUPC LUVs (LPR 3200:1) (Figure 5.3). The spectra of each OMP following folding exhibit characteristic decreases in λ_{max} and increases in fluorescence intensity compared with spectra acquired for the OMPs in an unfolded state, indicating that a substantial fraction of each OMP was successfully folded. To confirm and quantify this result, SDS-PAGE band shift assays were performed for BamA, tBamA and OmpA (Figure 5.4). The results show that BamA and OmpA fold with similar efficiencies (~80%) while tBamA has a slightly lower folding efficiency (~70 %), as quantified by densitometry.

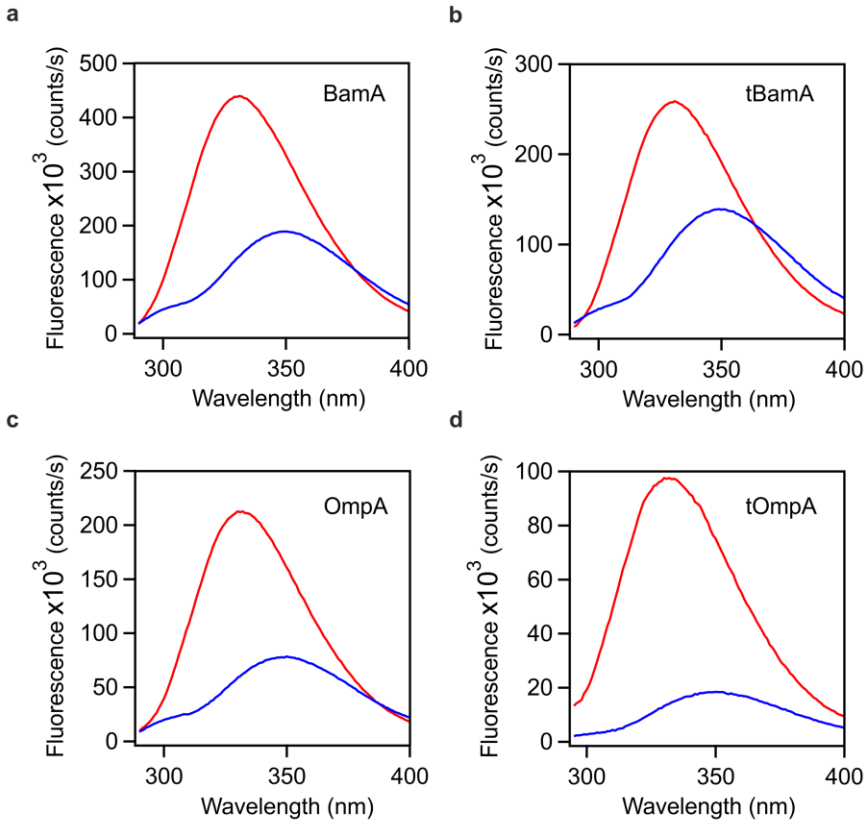


Figure 5.3: Fluorescence emission spectra of folded and unfolded OMPs. (a) BamA, (b) tBamA, (c) OmpA, and (d) tOmpA. Folded samples (red) contained 0.24 M urea, 1.28 mM DUPC (LPR 3200:1), 50 mM glycine-NaOH, pH 9.5. Samples were folded for >1.5 h. Unfolded samples (blue) contained 8 M urea, 50 mM glycine-NaOH, pH 9.5. Fluorescence was excited at 280 nm and emission measured between 290-400 nm. The OMP concentrations were 0.8 μM in (a-c) and 0.4 μM in (d), and all spectra were acquired at 25 ºC.

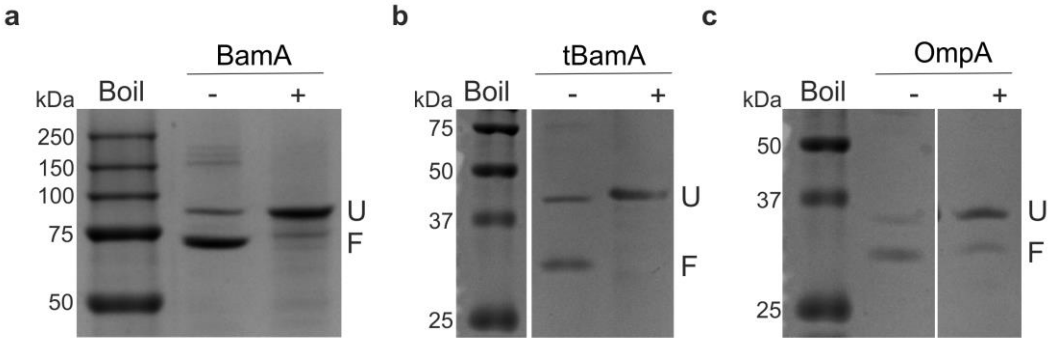


Figure 5.4: Substantial fractions of BamA, tBamA and OmpA are folded in DUPC liposomes. Semi-native SDS-PAGE band shift assays for (a) BamA, (b) tBamA, and (c) OmpA. Samples were folded for >1.5 h at 25 ºC and contained 0.24 M urea, DUPC liposomes (LPR 1600:1), in 50 mM glycine-NaOH, pH 9.5. Samples were run with or without prior boiling, indicated by '+' and '-' symbols, respectively.

To further characterise and compare the folding behaviour of BamA and OmpA in DUPC liposomes, SDS-PAGE band shift assays for folding reactions in increasing concentrations of urea (0.24 – 8 M) were performed (Figure 5.5a-c). The results show that BamA and OmpA exhibit similar folding efficiencies when diluted to different final urea concentrations from an unfolded state in 9 M. The urea concentration at which 50% of molecules are folded (apparent C_m) was ~3.5 M for both BamA and OmpA (Figure 5.5c). The apparent C_m is slightly higher than that obtained for OmpA in DLPC liposomes (~2.5 M) using a fluorescence-based assay[416]. By contrast, the unfolding behaviour for BamA and OmpA when denatured by heat is markedly different (Figure 5.5d). Fluorescence based temperature ramp experiments showed that while folded OmpA is resistant to thermal denaturation at temperatures up to 95 ºC, BamA begins to unfold at ~70 ºC with a denaturation midpoint (apparent T_m) of ~75 ºC. This apparent instability of BamA compared with OmpA may be due to the incomplete hydrogen bonding between the first and last β-strands of the BamA barrel[11,287] (Figure 1.15), potentially making the structure more unstable at higher temperatures.

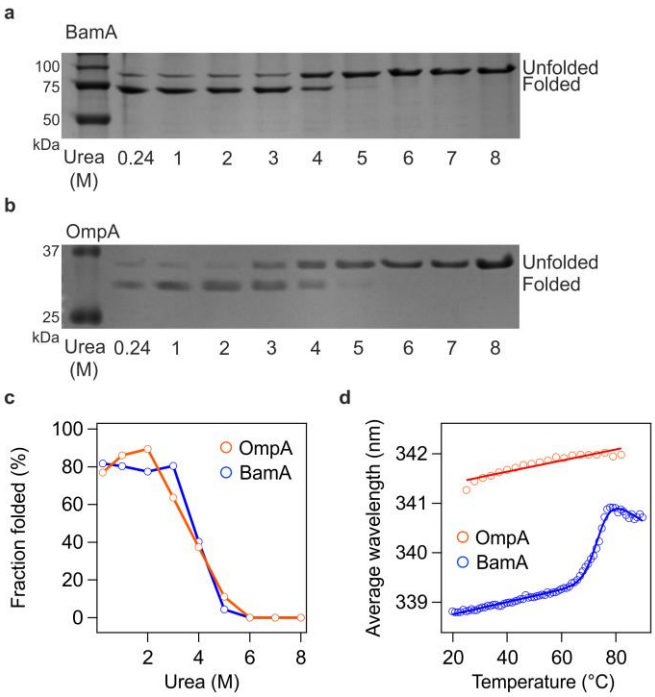


Figure 5.5: Effects of urea on BamA and OmpA folding efficiency in DUPC liposomes. Cold SDS-PAGE band shift assays in DUPC liposomes for concentrations of urea between 0.24 M and 8 M urea for (a) BamA, and (b) OmpA. (c) Quantification of gels in (a) and (b) by densitometry. Samples contained 1.5 μM BamA or 4 μM OmpA, in 50 mM glycine-NaOH, pH 9.5, and were folded overnight at 25 ºC. (d) Thermal denaturation of BamA and OmpA prefolded overnight in DUPC liposomes. Samples contained 0.8 μM OMP, 0.24 M urea, 50 mM glycine-NaOH, pH 9.5, and were folded overnight at 25 ºC prior to heat denaturation. Data for OmpA were fitted to a linear model, while data for BamA was fitted to a two-state thermal denaturation model (Section 2.16).

Having gathered these reagents the effects of a two-fold molar excess of each folding factor on tOmpA folding kinetics was next analysed systematically. OmpA prefolded in liposomes marginally lowered the folding rate constant compared with tOmpA alone (Figure 5.6a,b), while preincubation of tOmpA with BamA POTRAs had no effect on tOmpA folding kinetics (Figure 5.6c). A slightly higher folding rate constant was observed in the presence of prefolded tBamA (Figure 5.6d), and this was the same when tOmpA was preincubated with BamA POTRAs and then added to liposomes containing prefolded tBamA (Figure 5.6e). However, when tOmpA was folded in the presence of prefolded BamA, the kinetics of tOmpA folding were increased, and the folding behaviour was altered, with transients requiring an additional exponential term for adequate fitting (Figure 5.6f). The rate constant for the faster phase (k_1) was ~2-fold faster than that measured for tOmpA alone. These results are summarised in Table 5.1.

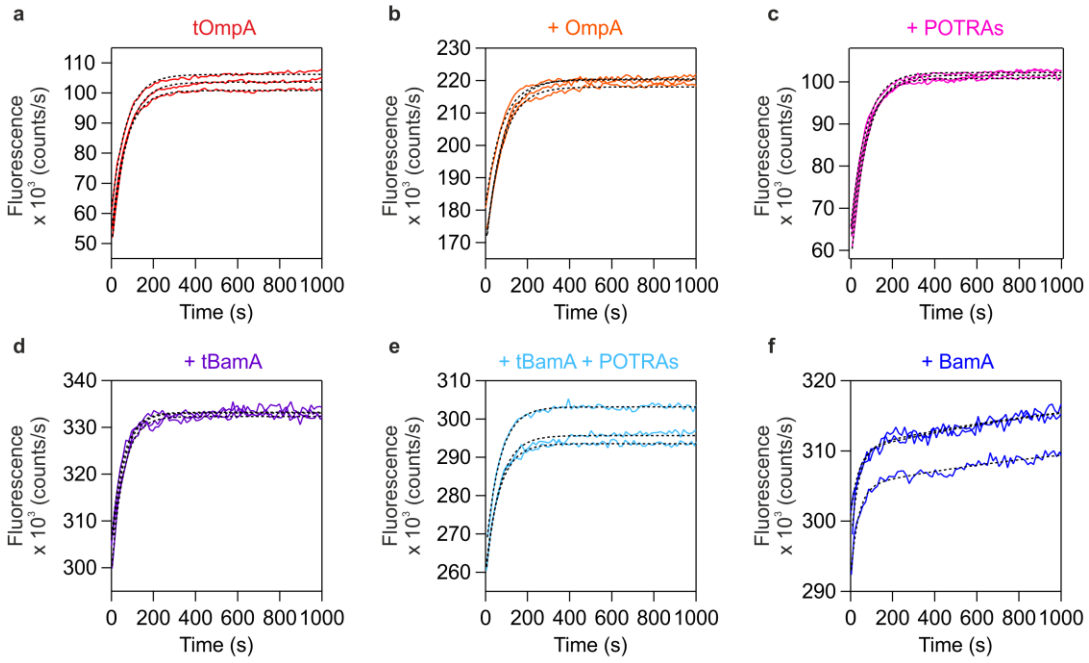


Figure 5.6: Effects of BamA and its constituent domains on the folding kinetics of tOmpA in DUPC liposomes. Kinetic folding traces for (a) tOmpA alone, and with (b) prefolded OmpA, (c) preincubation with BamA POTRA domains, (d) prefolded tBamA, (e) prefolded tBamA and preincubation with BamA POTRA domains, and (f) prefolded BamA (full-length). Samples contained 0.4 μM tOmpA, 1.28 mM DUPC (LPR 3200:1), 0.24 M urea, 50 mM glycine-NaOH, pH 9.5. A two-fold molar excess (0.8 μM) of OmpA, POTRAs, tBamA and BamA was used. A minimum of three transients are shown in each panel. Fits to a single exponential, (a-e), and double exponential function, (f), are indicated by dashed black lines (Table 5.1).

Folding reaction	k_1 ($x10^{-3}$ s^{-1})	k_2 ($x10^{-3}$ s^{-1})
tOmpA alone	14.9 ± 0.3	N/A
tOmpA + OmpA	14.1 ± 0.8	N/A
tOmpA + POTRAs	15.9 ± 0.1	N/A
tOmpA + tBamA*	20.1 ± 0.2	N/A
tOmpA + tBamA + POTRAs	16.9 ± 0.9	N/A
tOmpA + BamA*	33.6 ± 6.0	0.8 ± 0.1

Table 5.1: Measured rate constants for tOmpA folding into DUPC liposomes in the presence of BamA constructs or OmpA. Data are shown as the mean ± the standard error of the mean (s.e.m.) of the rate constants obtained from separate folding experiments, each using independently prepared batches of liposomes. For each batch of liposomes at least three folding transients were fitted globally to obtain the rate constants shown. Data are from three independent experiments, except for conditions indicated by an asterisk for which the data are from two experiments. N/A: The kinetic traces for the condition were adequately described by a single exponential.

5.3 Release and folding of tOmpA from its complex with Skp is dependent on the β-barrel domain of BamA

The results in Chapter 4 demonstrated that tOmpA is prevented from folding into DUPC liposomes when preincubated with a two-fold molar excess of Skp, at least on a timescale of 2 h (Figure 4.3a). Here, experiments were carried out to determine whether BamA could release tOmpA from Skp, as this may recapitulate an *in vivo* assembly pathway (Figure 5.7). When Skp-tOmpA was added to liposomes containing the prefolded OmpA control, similarly to Skp-tOmpA alone, no folding was observed (Figure 5.7a,b). By contrast, on addition of Skp-tOmpA to liposomes containing prefolded BamA, a folding transient is observed (Figure 5.7c), with a folding rate constant ~15-fold lower than for tOmpA alone (1.1 ± 0.2 and 14.9 ± 0.3 x 10^{-3} s^{-1} for tOmpA folding from Skp in the presence of BamA, and tOmpA alone, respectively). Therefore, release and folding of tOmpA from Skp is specifically dependent on BamA. This dramatic result is consistent with previous work in which the retardation of OmpA folding into 80:20 DLPC:DLPE liposomes by Skp was relieved by the presence of prefolded BamA[299].

Next, the effect of the N-terminal and C-terminal domains of BamA on Skp-tOmpA complexes were investigated separately to determine which part of BamA is responsible for the observed release of tOmpA from Skp. No folding was observed when premixed Skp and tOmpA were added to free POTRA domains in the presence of liposomes (Figure 5.7d), therefore the soluble domains of BamA alone are unable to release tOmpA from Skp. By contrast, when tOmpA preincubated with Skp was added to liposomes which contained prefolded tBamA a folding transient was observed (Figure 5.7e,f), indicating that the BamA β-barrel domain is necessary and sufficient for substrate release from Skp. However, the tOmpA folding rate constant on release from Skp by tBamA was ~three-fold lower than for full-length BamA (0.4 ± 0.1 and 1.1 ± 0.2 x 10^{-3} s^{-1} for tOmpA folding from Skp in the presence of tBamA and BamA, respectively (Table 5.2)). Therefore, the most efficient release of tOmpA from Skp in the presence of prefolded BamA depends upon the connection between the BamA barrel and the POTRA domains at the membrane.

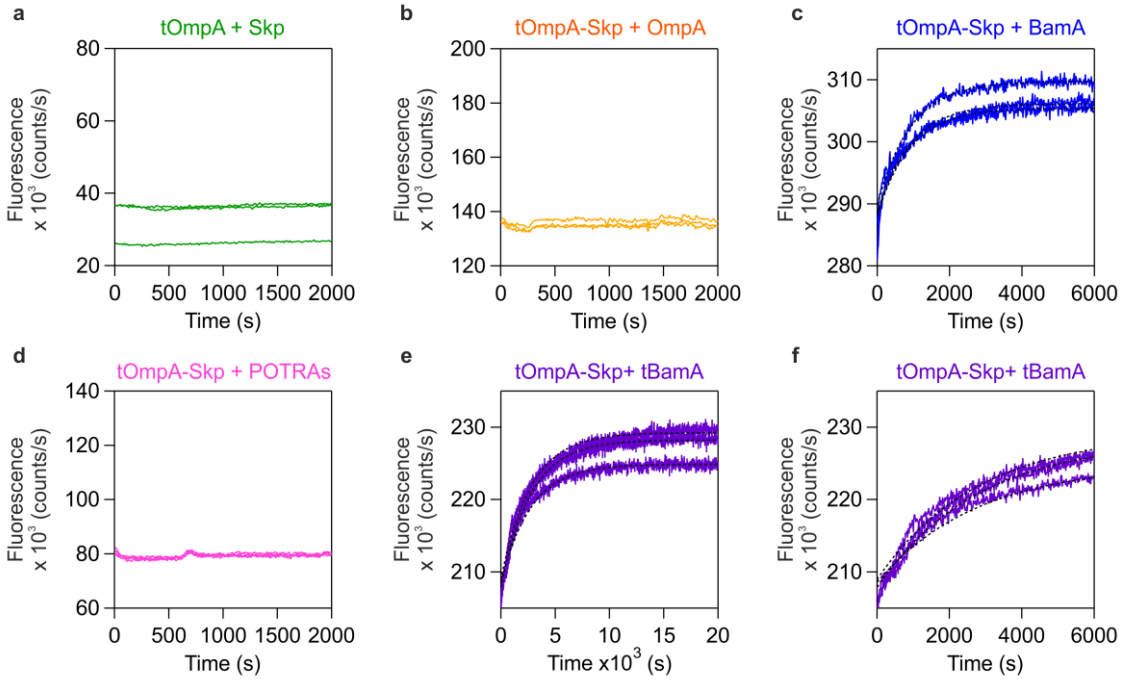


Figure 5.7: The β-barrel domain of BamA facilitates release and folding of tOmpA from its complex with Skp. Kinetic folding traces for (a) tOmpA-Skp alone, and with (b) prefolded OmpA, (c) prefolded BamA (full-length), (d) BamA POTRA domains, (e) prefolded tBamA, and (f) prefolded tBamA, first 6000 seconds of data in (e) for comparison with data for BamA in (c). Samples contained 0.4 μM tOmpA, 1.28 mM DUPC (LPR 3200:1), 0.24 M urea, 50 mM glycine-NaOH, pH 9.5. A two-fold molar excess (0.8 μM) of Skp, OmpA, POTRAs, tBamA and BamA was used. A minimum of three transients are shown in each panel. Global fits to a single exponential are indicated by dashed black lines (Table 5.2).

Folding reaction	k_1 (x10^{-3} s^{-1})	k_2 (x10^{-3} s^{-1})
tOmpA alone	14.9 ± 0.3	N/A
tOmpA-Skp	No folding	No folding
tOmpA-Skp + OmpA	No folding	No folding
tOmpA-Skp + BamA	1.1 ± 0.2	N/A
tOmpA-Skp + POTRAs	No folding	No folding
tOmpA-Skp + tBamA*	0.4 ± 0.1	N/A
tOmpA-Skp + SurA	No folding	No folding
tOmpA + SurA	12.7 ± 0.5	N/A

Table 5.2: Measured rate constants for tOmpA folding into DUPC liposomes from its complex with Skp in the presence of BamA constructs, OmpA or SurA. Data are shown as the mean ± the standard error of the mean (s.e.m.) of the rate constants obtained from separate folding experiments, each using independently prepared batches of liposomes. For each batch of liposomes at least three folding transients were fitted globally to obtain the rate constants shown. Data are from three independent experiments, except for conditions indicated by an asterisk for which the data are from two experiments. N/A: The kinetic traces for the condition were adequately described by a single exponential.

5.4 SurA cannot release tOmpA from its complex with Skp

Few studies have investigated the effects of SurA on *in vitro* OMP folding. SurA has been shown to increase the *in vitro* folding efficiency of OmpT[233,301] and BamA[283] in the presence BAM-containing native lipid proteoliposomes. In the latter case, SurA could be functionally replaced with urea, suggesting that its role could be that of a passive solubility agent in these experiments[283]. A study investigating the effects of chaperones on *in vitro* PagP folding in the absence of BAM components found SurA had no effect on *in* vitro PagP folding efficiency or yield[14,152]. Here, the effect of SurA on tOmpA folding in the presence and absence of Skp was investigated. First, kinetic experiments were performed in which tOmpA was preincubated with a two-fold molar excess of SurA, in the absence of Skp. These experiments showed that SurA had little effect on the folding rate constant (Figure 5.8a, Table 5.2), consistent with previous results for PagP[152]. Next, tOmpA was preincubated with a two-fold excess of Skp, prior to mixing with liposomes in the presence of a two-fold excess of

SurA (Figure 5.8b). No folding was observed, demonstrating that SurA is unable to release tOmpA from its complex with Skp to allow folding.

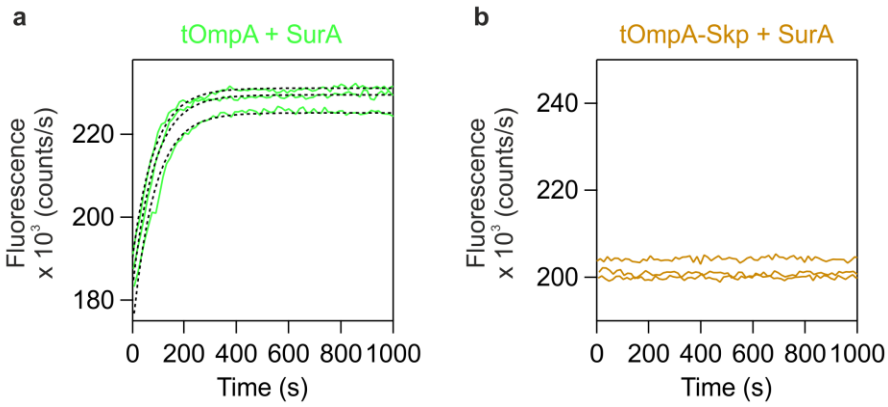


Figure 5.8: SurA is unable to release tOmpA from its complex with Skp. Kinetic folding traces for (a) tOmpA preincubated with SurA, and (b) tOmpA preincubated with Skp, then added to liposomes in the presence of SurA. Samples contained 0.4 μM tOmpA, 1.28 mM DUPC (LPR 3200:1), 0.24 M urea, 50 mM glycine-NaOH, pH 9.5. A two-fold molar excess (0.8 μM) of Skp and SurA, was used. A minimum of three transients are shown in each panel. Global fits to a single exponential in (a) are indicated by dashed black lines (Table 5.2).

Interestingly, the addition of SurA had no effect on the folding of tOmpA, with or without Skp, yet the SurA protein purified was pure (Figure 5.2c, upper) and far-UV CD indicated that it was correctly folded to the native state (Figure 5.2c, lower). To verify that SurA interacts with tOmpA under the experimental conditions employed, an assay for SurA binding to full-length OMPs was developed using microscale thermopheresis (MST) (Section 2.20). In this technique one of the binding partners is fluorescently labelled and kept at a constant low (nM) concentration, and its movement in a temperature gradient monitored in the presence of different concentrations of its potential binding partner[459]. An N-terminal cysteine mutant of tOmpA was produced and labelled via maleimide chemistry with AlexaFluor 488 dye (Section 2.20.1) for use in the MST assay. The results demonstrate that this technique works very well for OMP binding studies, likely due to the relatively low OMP concentrations (50 nM) that can be used compared with other binding analysis techniques such as ITC (in which normally ~10 μM is required[460,461]). The MST data in Figure 5.9 demonstrate conclusively that SurA binds tOmpA under the experimental conditions used in the kinetic assays above, with the observed binding affinity (K_d: 1.4 ± 0.4 μM) in agreement with literature values[186,197,198].

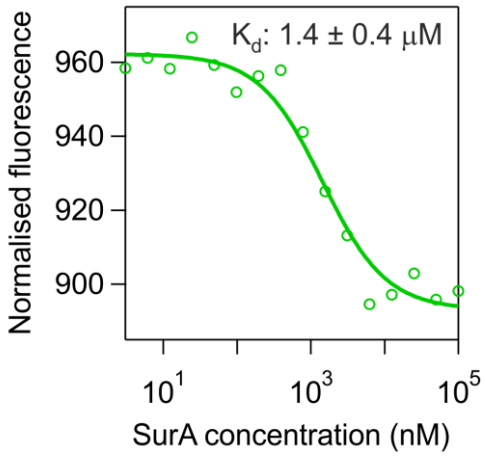
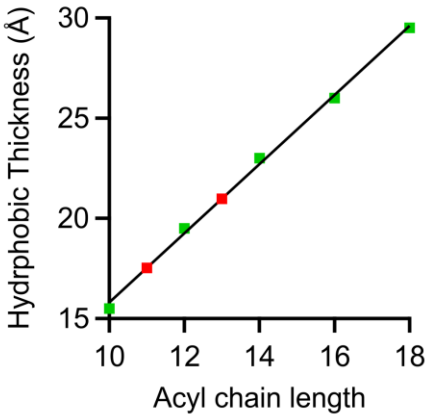


Figure 5.9: Binding of SurA to tOmpA measured by microscale thermopheresis. Samples contained 50 nM tOmpA labelled with AlexaFluor 488, 0.003-100 μM SurA, in 0.24 M urea, 50 mM glycine-NaOH, pH 9.5. Data were fitted to a quadratic tight binding equation (Section 2.20.3). Error indicated is the error in the fit to this model.

5.5 The catalytic effect of BamA is dependent on membrane thickness

Next, to investigate the hypothesis that the mismatch between the hydrophobic thickness of the membrane and the BamA barrel domain is important in the BamA mechanism, the effect of BamA on the folding of tOmpA in bilayers with different hydrophobic thicknesses was compared. PC lipids with different acyl chain lengths, DLPC (C12), DTPC (C13) and DMPC (C14), were selected. While the exact hydrophobic thickness of the OM is not known[53,462], the six acyl chains of LPS molecules contain predominantly myristoyl (C14) acyl chains[53,56], while phospholipids in the inner leaflet are mostly composed of lipids with C16 and C18 acyl chain lengths[34,39]. X-ray scattering experiments have shown that the hydrophobic thickness of PC bilayers is linearly dependent on acyl chain length[463], with DMPC bilayers most closely matching the expected hydrophobic thickness of the OM (Figure 5.10), as judged by the average hydrophobic thickness of a set of 24 OMP structures[464]. OMPs have an average hydrophobic thickness of 23.7 ± 1.3 Å[464], while DMPC bilayers have a hydrophobic thickness of 23.0 ± 1 Å[463].

a



b

Chain length	Hydrophobic thickness ± 1 (Å)
11	17.5
12	19.5
13	21.0
14	23.0

Figure 5.10: Hydrophobic thickness of PC bilayers varies linearly with acyl chain length. (a) Plot of hydrophobic thicknesses of PC bilayers composed of lipids with different acyl chain lengths obtained using X-ray scattering data from Lewis and Engelman (1983)[463]. Literature values are shown as green squares, and calculated values for DLPC and DTPC based on a line of best fit for the data (black line) are shown as red squares. (b) Table of hydrophobic thicknesses of the PC bilayers used in the current study from data in (a).

To verify that any observed differences in BamA behaviour in different lipids were not due to differences in folding yield, semi-native SDS-PAGE band shift assays[129] were performed to assess the fraction of folded BamA in DLPC, DTPC and DMPC liposomes (Figure 5.11a, Table 5.3). The results show that BamA folds with similar efficiencies in each case (62.7 ± 2.0, 57.3 ± 2.9 and 51.7 ±1.7 %, for DLPC, DTPC and DMPC, respectively), with a slight trend of decreasing folding yield with increasing acyl chain length, consistent with previous results[313] (Figure 5.11b). Far-UV CD data also indicated successful BamA folding into liposomes of each lipid type with a similar folding efficiency (Figure 5.11c). These data are consistent with previous CD data for BamA from *Thermus thermophilus*[243] and imply that both the BamA barrel and POTRA domains have been successfully folded to the native state.

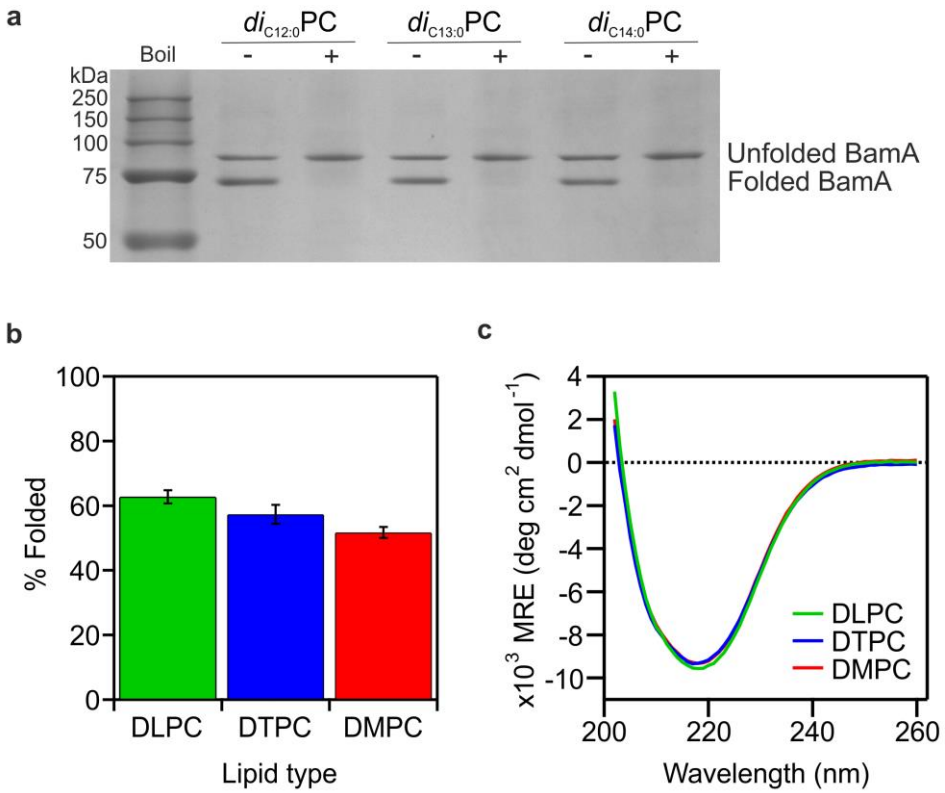


Figure 5.11: BamA folds with similar efficiencies into DLPC, DTPC and DMPC liposomes. (a) Semi-native SDS-PAGE band shift assay of BamA folding efficiency in DLPC (C12), DTPC (C13) and DMPC (C14) liposomes. Samples were run with or without prior boiling, indicated by '+' and '-' symbols, respectively. (b) Quantification of data in (a) by densitometry. Error bars represent the standard deviation from three independent experiments. (c) CD spectra of BamA in DLPC, DTPC and DMPC LUVs. Samples for semi-native SDS-PAGE contained 0.8 μM BamA 1.28 mM lipids (LPR 1600:1), 0.24 M urea, 50 mM glycine-NaOH, pH 9.5. Samples for CD contained 1.5 μM BamA, 1.2 mM lipids (LPR 800:1), 0.24 M urea, 50 mM glycine-NaOH, pH 9.5. All samples were folded overnight at 25 ºC.

Lipid type	Acyl chain length	BamA folding efficiency (%)
DLPC	C12	62.7 ± 2.0
DTPC	C13	57.3 ± 2.9
DMPC	C14	51.7 ± 1.7

Table 5.3: BamA folding efficiencies in DLPC, DTPC and DMPC liposomes. Data shown are the mean ± the standard deviation of the folding efficiencies obtained from three separate folding experiments, each using independently prepared batches of liposomes.

Next, experiments were performed in which tOmpA folding was monitored by fluorescence spectroscopy into DLPC, DTPC or DMPC liposomes, in the presence or absence of prefolded BamA in the bilayer. Folding of tOmpA alone in each of the three lipid types was characterised by a lag phase in the transients, giving a sigmoidal shape suggestive of the presence of folding intermediates (Figure 5.12a,c,e)[365]. Despite the small change in lipid structure, the addition of a single methylene group to the acyl chain has a dramatic effect on the folding of tOmpA, such that folding was much slower (~six-fold) in DTPC than in DLPC LUVs (Table 5.4). Similarly, tOmpA folding into DMPC liposomes was substantially slower (~seven-fold) than in DTPC liposomes (Table 5.4). The addition of a two-fold molar excess of prefolded BamA to the bilayer prior to folding of tOmpA removed the lag phase observed in the absence of BamA (Figure 5.12b,d,f). The data for tOmpA folding into DLPC or DTPC LUVs in the presence of BamA could be fitted to a single exponential, but required a double exponential function to fit the data for folding reactions in DMPC (Figure 5.13a,b,c).

As tOmpA exhibited nonexponential kinetics in the absence of BamA, the kinetics of folding were compared using the time taken to achieve 50% of the total observed fluorescence change (t_{50}) (Table 5.4, Section 2.14.2). In DLPC liposomes a ~two-fold decrease in t_{50} occurred in the presence of BamA, compared with in its absence. By contrast, in DTPC liposomes a ~six-fold decrease in t_{50} was observed. Dramatically, the presence of BamA in DMPC liposomes led to a ~ten-fold decrease in t_{50} (Figure 5.13d,e). Thus, as the hydrophobic thickness of the bilayer increases, and is closer to that of the OM[464], the folding rate enhancement of tOmpA mediated by BamA increases substantially.

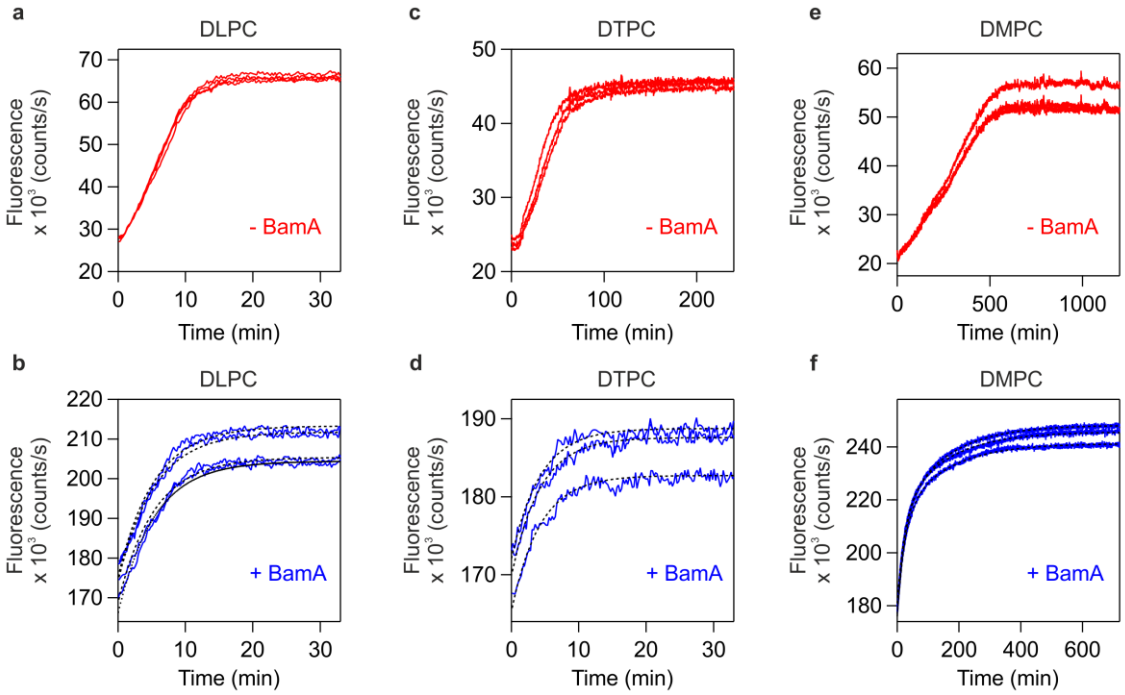


Figure 5.12: Example raw kinetic traces of tOmpA folding in the presence or absence of BamA in DLPC, DTPC and DMPC liposomes. Folding kinetics of tOmpA alone (upper) and in the presence of BamA (lower) into liposomes composed of (a) DLPC, (b) DTPC, or (c) DMPC liposomes. Samples contained 0.4 μM tOmpA, 1.28 mM DUPC (LPR 3200:1), 0.24 M urea, 50 mM glycine-NaOH, pH 9.5. A two-fold molar excess (0.8 μM) of BamA was used. For samples containing BamA, exponential fits to the data are shown as black dotted lines. Four transients for each of three liposome batches (12 in total) were used for calculation of the average T_{50} value for each condition. These experiments were performed at 30 ºC to be well above the T_m for DMPC (24 ºC)[365].

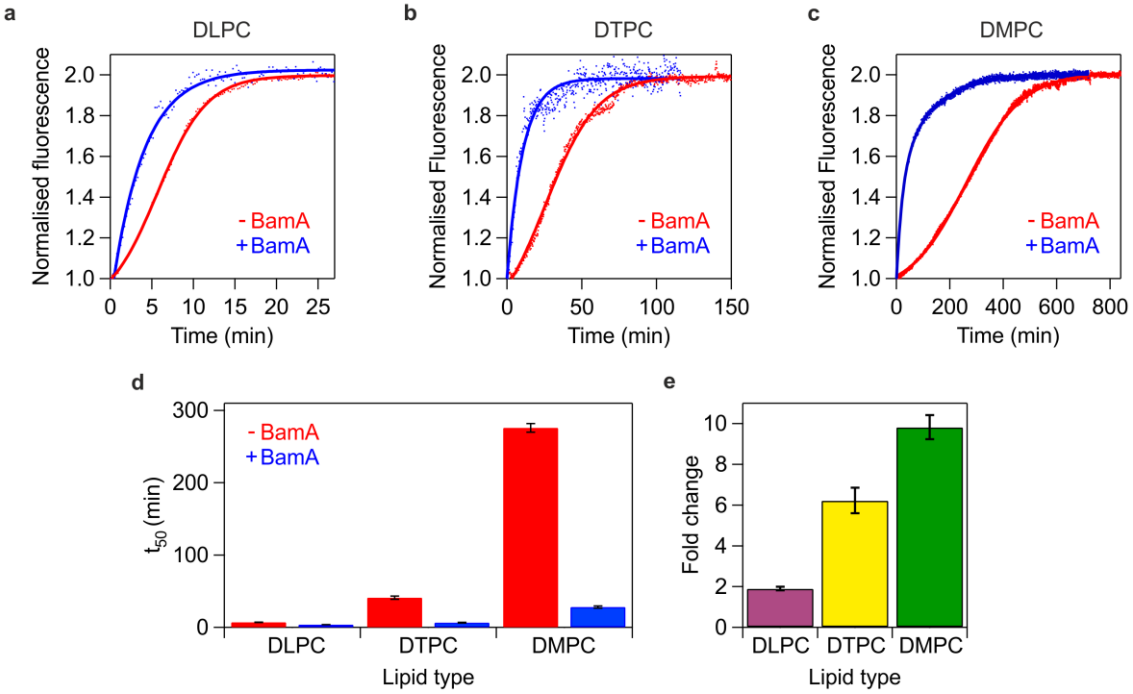


Figure 5.13: BamA accelerates tOmpA folding more effectively in liposomes with longer acyl chain lengths. Comparison of tOmpA folding in the presence and absence of BamA in LUVs composed of (a) DLPC, (b) DTPC, or (c) DMPC. Fits to a sigmoidal function (-BamA) or an exponential function (+BamA) are shown to guide the eye. (d) Comparison of t_{50} values for tOmpA folding into DLPC, DTPC or DMPC LUVs in the presence or absence of BamA. (e) Fold change in t_{50} values between tOmpA folding into DLPC, DTPC or DMPC LUVs in the presence or absence of BamA. Samples contained 0.4 μM tOmpA, 1.28 mM DUPC (LPR 3200:1), 0.24 M urea, 50 mM glycine-NaOH, pH 9.5, at 30 ºC. A two-fold molar excess (0.8 μM) of BamA was used.

Lipid type	Folding reaction	t_{50} value (min)	Fold change (+/- BamA)
DLPC	tOmpA alone	7.0 ± 0.3	1.9 ± 0.1
	tOmpA + BamA	3.7 ± 0.1	
DTPC	tOmpA alone	41.0 ± 2.2	6.2 ± 0.6
	tOmpA + BamA	6.6 ± 0.6	
DMPC	tOmpA alone	276.0 ± 6.0	9.8 ± 0.6
	tOmpA + BamA	28.1 ± 1.6	

Table 5.4: Measured t_{50} values for tOmpA folding into DLPC, DTPC or DMPC liposomes in the presence or absence of BamA. Data are shown as the mean ± the standard error of the mean (s.e.m.) of the t_{50} values obtained from three separate folding experiments, each using independently prepared batches of liposomes. Fold change errors are the propagated s.e.m. from folding reactions in the presence or absence of BamA (Section 2.14.2).

One explanation for these results is that BamA function is impaired by the increase in hydrophobic mismatch between the side of its barrel domain opposite to the β1-β16 seam and the bilayer as it is placed in thinner membranes. Membrane proteins as well as lipids may adjust to ensure the optimum hydrophobic matching[465], and it could be that deformations in BamA in shorter chain lipids impair its catalytic activity. However, available evidence suggests that to prevent the exposure of hydrophobic surfaces to hydrophilic environments, lipids more readily adapt to OMPs than vice versa, due to the high stablility of β-barrel OMP structures[11,292,331]. Therefore, it is likely that the increase in hydrophobic mismatch between BamA barrel in the β1-β16 seam region and the membrane as the chain length is lengthened from C12-C13-C14 leads to local increases in lipid disorder, which may be the origin of the increased BamA catalytic effect in longer chain LUVs. The results thus suggest that the hydrophobic mismatch between the BamA and the OM plays an important role in the overall BAM catalytic mechanism.

5.6 Discussion

While an increasing number of OMP folding studies have been carried out in the presence of folding factors[298], a systematic study of the combined effects of Skp, SurA and BamA on the kinetics of OMP folding was lacking. The results in this chapter confirmed that BamA alone can enhance the rate of folding of OMPs

(specifically tOmpA) *in vitro* (Figure 5.6) and showed that this effect was fairly modest (~two-fold) in short chain (C11) lipids, in the absence of additional factors (Table 5.1). However, when the selected substrate tOmpA is preincubated with Skp (at a 2:1 Skp:tOmpA molar ratio), folding is prevented on the experimental timescale used. The results then showed that BamA is able to release tOmpA from Skp, allowing it to fold (Figure 5.7). Experiments involving the two halves of BamA separately expressed and purified demonstrated that the BamA barrel domain (tBamA) is predominantly responsible for this activity (Figure 5.7). Interestingly tOmpA release from Skp by tBamA was much less efficient than that observed for full-length BamA. These differences are unlikely to be explained by these constructs folding with different orientations in the bilayer, i.e. with the periplasmic side of the barrel facing out from or in to liposomes in different ratios for BamA and tBamA. Trypsin digest experiments demonstrated that full-length BamA and a tBamA construct comprising residues 404-810 both exhibit orientated folding into PC liposomes with the periplasmic side of the barrel facing out from liposomes[299]. Additionally, the results are consistent with previous experiments in DLPC LUVs which showed that tBamA enhanced the folding rate constant of OmpA less than full-length BamA[299]. Further, a construct consisting of the BamA barrel and POTRA 5 (BamAΔP1-4) was less able than wild-type BamA to accelerate the folding kinetics of tOmpA and OmpX in 80:20 DDPC:DDPE liposomes[300]. A two-fold difference in folding efficiencies between BamA and BamAΔP1-4 was reported, but the difference in apparent rates of the folding reactions they assisted was five-fold, leading the authors to speculate that BamAΔP1-4 has a reduced specific activity compared with wild-type BamA[300]. This study concluded that the catalytic activity, in terms of reduction of the kinetic barrier imposed by PE headgroups, was located at the membrane and required the β-barrel domain and POTRA domain 5[300]. Subsequent NMR studies have shown that in a BamAΔP1-4 construct POTRA 5 remains in an unfolded state in the absence of its N-terminal four POTRA domains[466,467], therefore is unlikely contribute to activity. The results presented here show that POTRA 5 is not required for tBamA to assist tOmpA folding at the membrane.

One explanation for the observed differences between the kinetics of tOmpA release from Skp by BamA and tBamA could be differences in BamA and tBamA folding efficiencies in DUPC under the conditions employed. However, as shown in Figure 5.4, when BamA and tBamA are added at a concentration of 0.8 μM they fold with similar efficiencies (~80% and ~70% for BamA and tBamA, respectively) (Figure

5.4a,b). The concentration difference between folded BamA and tBamA in the membrane (0.64 - 0.56 = 0.08 μM or ~10%) is unlikely to explain the ~four-fold difference in tOmpA folding rate. Therefore, the results suggest that the POTRA domains contribute to the efficient release of tOmpA from Skp when attached to the BamA β-barrel domain. Recent atomistic molecular dynamics simulations of BamA in a native outer membrane showed that the POTRA domains are highly dynamic when attached to the BamA barrel and were able to interact with the membrane independently of the barrel domain[262]. Insertion into the membrane of the two tryptophan residues in POTRA domain 3 was observed in a conformation which was consistent with the known arrangement of BamB and BamD in the complex[262]. Consistent with these results, in the recent cryo-EM structure of the BAM complex these two tryptophan residues are observed inserted in the detergent micelle (Figure 1.27a)[232]. This may be a stabilising interaction to anchor POTRA 3 during the BAM catalytic cycle (Section 6). However, it is possible the POTRA domains, as well as the BamA barrel may be involved in modulation of the membrane to aid catalysis, and attachment of the POTRA domains close to the membrane may facilitate this.

The POTRA domains may provide an additional electronegative surface close to the membrane. A key factor in the release of tOmpA from Skp by tBamA may be the electrostatic interaction between the positively charged tips of the Skp subunits and the electronegative patches at the periplasmic side of the BamA barrel (Figure 5.14a). It has been proposed that the POTRA domains may be instrumental in the release of substrates from their complex with Skp, due to the interaction of the highly basic Skp (pI ~9.5) with negatively charged patches on the POTRA domains (Figure 5.14b)[383]. Therefore, although the data in Figure 5.7 shows that the POTRA domains are insufficient on their own to release tOmpA from Skp, the electrostatics of the POTRA domains may still play an important role in substrate release from Skp in the context of full-length BamA.

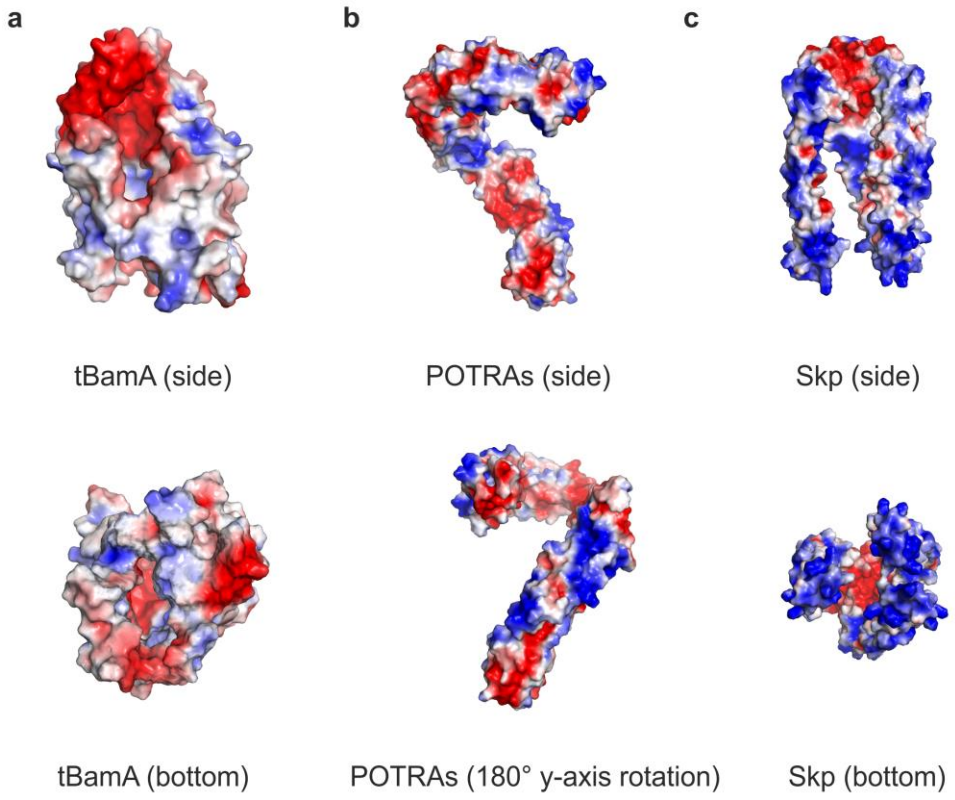


Figure 5.14: The BamA barrel and POTRA domains have electronegative regions which may be involved in release of substrates from Skp. Electrostatic surface representations of (a) tBamA, (b) the BamA POTRA domains, and (c) Skp. Regions of blue and red represent areas of electropositive and electronegative surface potential, respectively. Structure of BamA from the cryo-EM structure of the *E. coli* BAM complex (PDB: 5JLO[232]). Structure of Skp from PDB: 1U2M[150,] with missing residues from chains B and C modeled from chain A using MODELLER[191]. Electrostatic surface representations (-2 kT/e to +2 kT/e) were created using the APBS plugin for PyMOL[151].

A further possibility is that the presence of both the membrane and soluble BamA domains allows cooperative behaviour between them. Several studies have implicated functional cooperation between the periplasmic and barrel domains of BamA. Genetically, substitutions in the BamA barrel can restore viability to the conditional synthetic lethal phenotype of a *ΔBamBΔBamE* strain[468], with four of the six rescuing mutations located in the L6 loop. Another study also found that mutations in BamA L6 could ablate the SDS-sensitivity of a mutation (E373A) in BamA POTRA 5[469]. Further, domain swap experiments in which the *E. coli* BamA barrel or POTRA domains were individually substituted with those from other species showed that efficient OMP assembly depends upon precise alignment of the two domains[254].

Whether the release of OMP substrates to BAM via Skp occurs *in vivo* is still under question, and evidence of a genetic interaction is lacking[445]. However, *in vitro* studies have shown that Skp can deliver substrates to negatively charged bilayers[152,153]. Further, recent data in which induction of the Cpx envelope stress response[470] caused downregulation of Skp, thus preventing toxicity due to OMP mislocalisation in the IM, suggested that Skp is able to fold OMPs directly into membranes *in vivo*[154]. While SurA has been cross-linked to BamA *in vivo*[145,187], Skp has only been cross-linked *in vivo* to BAM components indirectly via the passenger domain of the EspP autotransporters[111].

There is a discussion in the literature as to whether or not Skp and SurA act in sequential[68,227] or parallel pathways[145,147]. The kinetic data presented here for SurA suggest that this chaperone cannot release tOmpA from Skp which is consistent with recent competitive affinity assays for the 22-stranded substrate FhuA[97]. The low micromolar K_d obtained from MST experiments (Figure 5.9) is consistent with previous results for SurA binding to peptides[186,197,198], and with that obtained for SurA binding to the full-length OMPs, OmpF and OmpG using a peptide competition assay[197]. However, tighter binding (K_d of ~0.1 μM) was observed for SurA binding to the 16-stranded OmpC in fluorescence-based experiments[162]. The differences in affinity for OMP substrates between Skp and SurA (low nanomolar and low micromolar for Skp[160,162,336] and SurA[186,197,198], respectively) also appears to argue against a handover mechanism. However, transport proteins participate in an inherently non-equilibrium process, due to flux through the system. Thus, although Skp and SurA are present in approximately the same quantities *in vivo* (~6600 and ~3900 molecules/cell for Skp and SurA, respectively[83]) the amount of free chaperone at any point will be dependent upon the flux through the pathway. If the rate of depletion of SurA-OMP complexes by substrate handover to the BamA POTRA domains is very rapid, the pool of available SurA may be much higher than that of Skp.

The results show that the BamA POTRA domains free in solution have no effect on *in vitro* tOmpA folding kinetics. POTRA domains are proposed to chaperone OMP substrates en route to the membrane[247,256]. Possibly, the edges of POTRA β-strands weakly bind the β-strands of incoming OMP substrates by β-augmentation allowing processive sliding motions towards the BamA barrel[12,471]. However, no direct binding of a full-length OMP to the POTRA domains has yet been reported. Here, POTRA-

tOmpA binding was unable to be demonstrated by MST (data not shown), in sharp contrast to the binding data obtained for SurA-tOmpA (Figure 5.9).

The work described in Section 5.5 demonstrates the importance of hydrophobic mismatch between the membrane environment and the BamA barrel in the BamA mechanism. This may be due to an increase in bilayer defects and/or disorder in the membrane[11,365]. The effects of increasing the dynamics and fluidity of membranes on OMP folding was recently dramatically demonstrated in experiments in which ~100% folding efficiency for OmpX and OmpA could be achieved by 'heat shock' of lipid membranes at 70 ºC[472]. The results in this chapter also highlight the importance of lipid:protein interactions in OMP folding[383], and support the view that BamA is a very unusual kind of catalyst in that it has two substrates, both proteins and lipids[300].

In summary, a systematic approach was used in this chapter to investigate the effects of periplasmic chaperones and lipid chain length on BamA-assisted tOmpA folding *in vitro*. The results obtained, together with a consideration of currently available literature, suggest a new 'barrel elongation' model for the assembly of OMPs in Gram-negative bacteria by the BAM complex, which is proposed in the following chapter.

6 Conclusions and final thoughts

Unlike the spontaneous protein folding observed in dilute solutions[473], protein folding in the cell is complicated by the high concentration of other proteins with which aberrant interactions can be made[474]. This can lead to aggregation, the accumulation of toxic species and cell death, therefore cells expend considerable effort to maintain unfolded proteins in a folding-competent state[446]. In the case of OMPs, folding is complicated by the fact that the site of synthesis in the cytosol is far from the location of their final folded state in the OM[68]. Therefore, a network of folding factors is important in ensuring OMP proteostasis, which becomes particularly important under stress conditions[10,18]. The work presented in this thesis has investigated OMP biogenesis by purifying the major folding factors involved, and trying to recreate aspects of the pathway *in vitro*. In this final chapter the major findings are summarised, and in conclusion, a new model is presented for the mechanism of OMP assembly by the BAM complex, which is consistent with the results obtained here, as well as with the currently available literature.

In this work, the mechanisms by which chaperones aid OMPs on their perilous journey across the aqueous periplasmic space has been investigated. The first assay to compare the kinetics of OMP folding between different OMPs using fluorescence spectroscopy was developed (Chapter 3). This was then used, together with IMS-MS and MD simulations, to compare the interaction of Skp with OMPs of different sizes (8- to 16-stranded). It was shown that Skp is able to function as a multivalent chaperone with mechanistic similarities to both the 'beads on a string' chaperones, such as Trigger Factor, and 'Anfinsen cage' chaperones, such as the chaperonins[173,446] (Chapter 4).

Evidence has been provided that SurA does not release OMPs from Skp, which argues against a sequential handover mechanism for these chaperones (Chapter 5). This is consistent with the notion that, rather than due to a specific interaction between Skp and SurA, the movement of substrates between the Skp and SurA chaperones depends upon (1) the relative k_{on} and k_{off} rates between the chaperones and unfolded OMPs, and (2) the relative concentrations of Skp and SurA (Chapter 5). This is consistent with genetic studies which suggest that Skp and SurA function in partially redundant parallel pathways, with SurA as the main pathway and the Skp/DegP more important in times of stress[145]. Recent kinetic modelling of the flux of OMPs through the periplasm also supports this hypothesis[124]. Therefore, it

appears there is no predefined sequential chaperone pathway; instead OMP proteostasis is controlled by the kinetics and thermodynamics of interactions between unfolded OMPs and chaperones in the periplasm[17]. Given that kinetic modelling suggests that OMPs make hundreds of separate interactions with chaperones on their way across the periplasm[124], conceptually thinking about OMPs following particular chaperone 'pathways' may be misleading.

The results in Chapter 5 suggest that Skp may deliver substrates to the BAM complex, as has been previously suggested[171,299,383]. Kinetic experiments showed that prefolded BamA can specifically release tOmpA from its complex with Skp, suggesting that this may be a relevant *in vivo* pathway (Chapter 5). There is no *in vivo* cross-linking data to support this, but it is striking that tOmpA is efficiently released from Skp by prefolded BamA, whereas prefolded OmpA has no effect. The presence of Skp in an *in vitro* reconstitution of autotransporter assembly assay, which included the BAM complex and SurA, slowed the kinetics of EspP folding[239]. However, this may be due to Skp binding to the EspP soluble passenger domain slowing its precession across the membrane[111].

Further, the results in Chapter 5 suggest that the BamA POTRA domains and the BamA β-barrel act cooperatively, as release and folding of tOmpA from its complex with Skp was more efficient (~3-fold) with prefolded BamA than with prefolded tBamA (Section 5.3). This is consistent with *in vivo* data which showed genetic interactions between BamA POTRA 5 and extracellular loop 6 (L6) in BamA[469]. The results are also consistent with previous *in vitro* studies, in the absence of Skp, which showed faster OMP folding in the presence of prefolded BamA compared with tBamA[299] or a BamA construct lacking POTRA domains 1 to 4 (BamAΔP1-4)[300]. Further, Cys mutations designed to cross-link BamA POTRA 5 to the BamA barrel in the 'lateral open' conformation were lethal, but could be rescued by the addition of reducing agent[221].

How might the cooperativity observed between the BamA barrel and POTRA domains function in the context of the full BAM complex? A striking feature of the now available structures of the complete[221,222,232] and near-complete[221,223] BAM complex is that the POTRA domains and BAM lipoproteins form a ring structure in the periplasm (Figure 1.25). Where protein complexes have ring-like structures these can be involved in cooperative behaviour in which movements are propagated

around the ring[475,476]. Examples of this include the GroEL chaperonin, which displays cooperativity of ATP binding and hydrolysis around the heptameric subunits in each of its two rings[446], and also the AAA (ATPases Associated with diverse cellular Activities) unfolding and degradation machine ClpX[477]. Thus, I propose a model in which the binding of nascent OMPs to BamD, and possibly simultaneously to BamA POTRA 5, causes a conformational change which is propagated around the periplasmic ring to the BamA β-barrel domain. In the recent cryo-EM model of the BAM complex, BamD exhibits a rigid body hinge movement between TPRs 3 and 4 when compared with other BAM structures[232]. Binding of an OMP could trigger this movement of the N-terminal three BamD TPR domains away from the C-terminus of BamD. This movement would then be propagated to POTRAs 1 and 2 through their interaction with the BamD N-terminal TPR domains. This interaction, between the BamD N-terminal TPR domains and the BamA N-terminal region, is structurally slightly different in all BAM structures, suggesting it is a dynamic region[232]. The BamD-POTRAs interaction in turn would lead to 'concertina-like' movements in the POTRA domains, with the major POTRA hinge motion, the narrowing of the angle between POTRAs 2 and 3, aided by BamB. Finally, the conformational change would be transmitted through BamA POTRA 5, to the BamA β-barrel, ensuring that BamA barrel β1 and β16 are in the 'lateral open' state, which is assumed here to be the 'acceptor' state (Figure 6.1). The two consecutive tryptophans in BamA POTRA 3 are observed inserted into the detergent micelle in the recent cryo-EM structure of the BAM complex in a 'lateral open' state[232] (Figure 1.27a), as seen previously in MD simulations of BamA alone in a native outer membrane[262]. These two residues are in a similar location in the two structures of the BAM complex in the 'lateral closed' state[221,222]. Therefore, it may be that the role of this loop is to anchor POTRA 3 in position to assist the transfer of conformational movement around the periplasmic ring between POTRAs 1-2, and POTRA 5. This model provides an explanation as to why BamD is an essential protein; it is required to transmit the information that a nascent OMP has entered the BAM periplasmic ring to trigger the BamA β-barrel domain into an 'active' state, ready for insertion and folding. The recently observed loop in BamD, which makes contact with the detergent micelle in the cryo-EM structure (Figure 1.27b)[232], may not be involved in signalling to the membrane. Rather, its importance could be due to stabilising the C-terminal BamD TPRs (4-5) to allow the N-terminal TPRs (1-3) to make the rigid body movement which is propagated around the ring. Note that in this model the minor conformational changes in BamD eventually lead to a major allosteric conformational change in the BamA barrel, due to mechanical propagation around

the ring. It may be that despite the major structural differences observed in the two states so far captured in structural studies of the BAM complex[221-223,232], there is not a large energetic barrier between them.

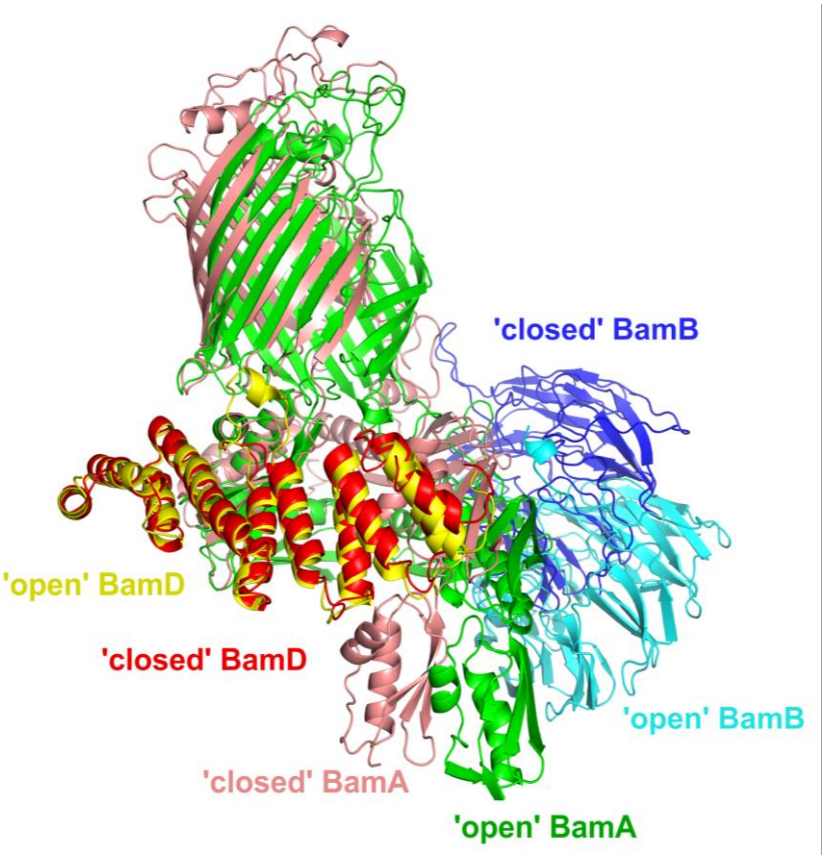


Figure 6.1: Movements in the N-terminal TPR domains (1-3) of BamD propagate around the periplasmic ring to the BamA barrel in the 'barrel elongation' model. BAM in a 'lateral closed' conformation (PDB: 5D0O[221]) and the 'lateral open' cryo-EM structure (PDB: 5LJ0[232]) are shown aligned on TPRs 4-5 of BamD (residues 160-241). The small movement between the N-terminal TPRs 1-3 of BamD in the two structures (a maximum displacement of 6 Å[232] is correlated with larger rearrangements in the BamA POTRA domains, the position of BamB, and the open/closed status of the BamA barrel. BamA, BamB and BamD in the 'lateral open' structure are shown in green, light blue, and yellow, respectively. BamA, BamB and BamD in the 'lateral closed' structure are shown in pink, dark blue, and red, respectively. For clarity, BamC and BamE are omitted. Image created with PyMOL.

This model may also explain why the C-terminal sequence of OMPs is so important for assembly *in vivo*[128,478]. It may initiate the conformational change in BamD which places BamA into the 'active' ('lateral open') state. The structural similarity of BamD to the TPR domains of PEX5, which binds peroxisomal targeting sequences[281], first suggested that BamD may be involved in recognition of OMP targeting signals[270].

BamD can be cross-linked to peptides containing OMP targeting sequences *in vitro*[235], and peptides containing the β-signal of BamA were shown to bind BamD in an *in vitro* pull-down assay[136]. Bacterial genetic studies[282] and *in vivo* cross-linking of EspP[112], and LptD[479], also suggest that BamD binds substrates separately from BamA (Section 1.4.4).

The trigger for completion of the reaction cycle may be the release of the C-terminal OMP sequence from BamD, allowing restoration of the 'resting' ('lateral closed') state. BamE may assist in this return to the initial state; genetic evidence suggests that BamE may be involved in returning BamA to the ground state through regulation of BamD[276,469]. In this model, the 'scissor-like' motion of the β1-β6 strands of the BamA barrel through ~65º, suggested by the structures of BamA in the 'lateral open' and 'inward open' states (Figure 6.2b), may be key to breaking the interaction between β1 of the BamA barrel and the final β-strand (β1) of the nascent OMP. This would allow the barrel of the nascent OMP to close and complete folding, and the restablishment of the interaction between BamA β1 and β16. Note that in this model the OMP 'targeting sequence' has an additional role other than targeting, and could be thought of as an 'activating sequence'. The role of BamC in this proposed mechanism is unclear. However, given that its N-terminal region covers a wide surface area of BamD, its role seems likely to consist of assisting and/or regulating the function of BamD. However, the model does provide a clear explanation as to why OMP biogenesis is so impaired in *bamB* null mutants[125,264]. BamB is required to regulate the hinge movement at POTRAs 2 and 3 for efficient propagation of conformational changes around the ring from BamD. This is also suggested by the available crystallographic structures of BAM. All complete structures are in the 'inward open' ('lateral closed') state[221,222], while those lacking BamB are in the 'lateral open' state[221,223].

Given that *ΔbamB* and *ΔsurA* mutants are phenotypically identical[123], and *ΔbamBΔsurA* mutants are non-viable[264], this strongly implicates SurA in also assisting in the movements in the POTRA domains required to promote the BamA barrel to the 'active' state. This most likely occurs through the known interaction of SurA with BamA POTRA 1[187].

The results in Chapter 5 also show that BamA functions much more effectively as an *in vitro* OMP folding catalyst when the hydrophobic thickness of the membrane

matches that of the OM. Therefore, local alteration of the lipid environment is clearly important in the BAM mechanism. These results are consistent with the 'BamA-assisted' model of OMP assembly, in which BAM 'primes' the membrane for insertion and aids trafficking to the thinned/destabilised area of the membrane[422]. However, the results are also consistent with the 'BamA-budding model', in which, in addition to membrane destabilisation, BamA creates a 'super-barrel' or 'hybrid barrel' with incoming OMP substrates[248,263,422]. Also, these two models are not mutually exclusive and a 'hybrid' model in which both mechanisms are correct is possible. The results in Chapter 5 strongly support the BamA-assisted model. By contrast, the 'hybrid barrel' appears less likely for the following reasons: (1) the 'hybrid barrel' model requires the making and breaking of large number of hydrogen bonds, each time a new β-hairpin is inserted, which seems energetically implausible; (2) β16 of the BamA barrel is a short β-strand, yet the BAM complex has to function with substrates with very different numbers of residues per β-strand (as determined by the number of β-strands in the barrel and the shear number)[480,481]; and (3) the model seems unnecessarily complex. Therefore, an alternative 'barrel elongation' model for the BAM complex mechanism is proposed here. In this model BamA is proposed to assist folding by affecting the local membrane environment due to the hydrophobic mismatch between the OM and side of the BamA barrel at the β1-β16 interface, as suggested previously[11,300,365]. In the 'barrel elongation' model the 'lateral open' state is the OMP acceptor state (see above), with folding initiated by templating of the C-terminal strand of the nascent OMP by β1 of the BamA barrel. Thus, the BamA barrel is extended ('elongated') from the β1 strand by β-augmentation, but without a β-augmentation interaction with BamA β16. Once this initial interaction is formed, successive β-strands are templated onto the barrel, possibly in β-hairpin units[97], in a similar manner to the amyloid formation reactions observed for peptides with alternating polar and non-polar residues[114].

The word 'elongation' is deliberately chosen as a reference to the elongation phase of amyloid aggregation reactions. OMP assembly is viewed as a β-sheet templating reaction similar to seeded amyloid aggregation reactions[482]. OMP folding via the BAM complex is therefore seen as an example of a highly controlled seeded aggregation reaction. *In vitro* studies of the kinetics of amyloid aggregation reactions have shown that the kinetics of fibril assembly are much faster if they are 'seeded' with small amounts of pre-formed amyloid fibres[425,482]. In the 'barrel elongation' model, a similar templating reaction to initiate the formation and association of

successive β-strands, combined with membrane priming, is the origin of the BAM catalytic effect.

One question for this model is how is the stability of the BamA β-barrel maintained if β16 is not involved in the formation of a hybrid barrel? In the structures of BamA alone[11] and of the BAM complex in the 'lateral-open' conformation[221,223,232], the BamA barrel can remain stably folded with very few (as low a 2) hydrogen bonds between β1 and β16[11]. Therefore, it is conceivable that the BamA barrel could remain stably folded without hydrogen bonds being sequentially formed and broken between β16 and the incoming substrate β-strands. MD studies of the BamA barrel in DMPE bilayers also suggest that the BamA barrel can exhibit transient opening between β1 and β16 and still remain stably folded in the membrane[11]. Omp85 family members have a well-conserved VRGF/Y motif present in L6[483]. Mutation of the RGF residues in this motif to alanine all affect OMP assembly *in vivo*, with R661, the most conserved residue, the most critical[484]. The functional role of this motif is unknown. One reason why the VRGF loop is so-well conserved in BamA homologues could be that it is required to stabilise the BamA barrel as nascent OMPs are elongated from BamA β1 (Figure 6.2a), although experimental evidence for such a role is lacking.

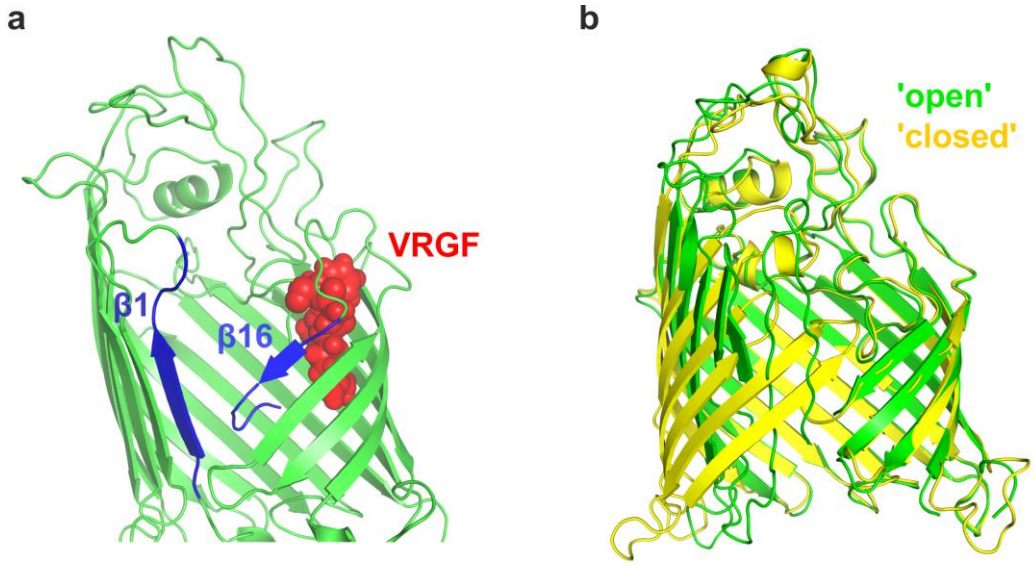


Figure 6.2: The conserved VRGF motif maintains BamA barrel stability during catalysis in the 'barrel elongation' model. (a) The BamA barrel in the 'lateral open' conformation (PDB: 5EVK[223]). BamA barrel strands β1 and β16 are highlighted in blue. In the model proposed in this chapter, the VRGF motif (red) stabilises the BamA barrel while the β-strands of substrates are elongated by templating from the β1 strand of the BamA barrel. (b) Alignment of BamA barrel structures in the 'lateral open' (green) (PDB: 5D0Q[221]) and 'lateral closed' (yellow) (PDB: 5D0O[221]) states. The opening of the barrel is proposed to allow the BamA barrel β1 strand to be available for templating the elongation reaction of the nascent OMP via β-augmentation. The ~65º movement of strands β1-β6 to return the BamA barrel to the closed state is proposed to sever the interaction between BamA β1 and the final β-strand of the nascent OMP. This allows completion of nascent OMP folding by the association of its first and last β-strands, and completes the BAM reaction cycle. Images created with PyMOL.

How does the 'barrel elongation' model fit with the known role of the BAM complex in autotransporter assembly?[107] As in the 'hybrid barrel' model, in the 'barrel elongation' model the BamA barrel also laterally opens, and therefore allows the secretion of the N-terminal passenger domains of autotransporters[248]. The trimeric autotransporters, such as YadA, which have 12-stranded β-barrels with each subunit contributing four β-strands, are also BAM-dependent[478]. The 'barrel elongation' model is consistent with their assembly, as incoming subunits could be involved in a sequential elongation reaction with the first subunit templated by β1 of BamA, β1 of the second subunit templated by β4 of the first subunit, and β1 of the third subunit templated by β4 of the second subunit. The model does not require that the β-strands which form the β-barrel of the nascent substrate enter the lumen of the BamA barrel, as proposed in the 'hybrid barrel' model[263]. However, to allow

secretion of autotransporters passenger domains, initiation of templating of the nascent β-barrel by the BamA β1 strand must be coupled to movements which prevent POTRA 5 from blocking access to the BamA barrel lumen. This allows a route across the OM for autotransporter passenger domains, as well as extracellular OMP loops and domains, via the BamA 'exit pore' above the β1-β16 lateral gate[263].

When the final β-strand of the nascent OMP is elongated the reaction cycle must complete to allow BAM to catalyse the folding of its next substrate. This requires removal of interaction between the final β-strand of the nascent OMP and β1 of BamA, to allow the first β-strand to the substrate to interact with its last β-strand to close the substrate OMP barrel, and allow dissociation from BamA. In the 'barrel elongation' model, this is caused by the scissor-like movement in the BamA barrel, linked to the 30º turn of the periplasmic ring (Figure 6.2b). The trigger for this movement could be the dissociation of the C-terminal sequence of the nascent OMP from BamD, and possibly BamA POTRA 5, assisted by BamE (see above). This trigger would then be propagated around the periplasmic ring to cause the conformational change in the BamA barrel required to complete the reaction cycle. Thus BAM is returned to the 'ground state', in the reverse process to that which mediated the transition to the 'acceptor state'. The displacement of the final β-strand of the nascent OMP from β1 of BamA may occur via a zip-in-zip-out mechanism similar to that proposed for donor strand exchange (DSE) in Gram-negative pilus assembly[485]. Note that if BamA β16 is not involved in the formation of a hybrid barrel, we would expect that it exhibits little structural change between the 'lateral closed' and 'lateral open' states, which is exactly what is observed (Figure 6.2b). One prediction from the mechanism outlined above is that the N- and C-termini of all OMPs in Gram-negative bacteria should be located in the periplasm, therefore restricting the number of β-strands to multiples of two; this is in agreement with all structural data to date[481,486]. Therefore, the 'barrel elongation' model of the BAM mechanism involves intricate coordinated cooperative movements around the periplasmic ring. Another role for the ring structure, aside from coordinating the conformation of the BamA barrel, may be to provide a sequestered environment while templating occurs, to protect elongating OMPs from aberrant interactions. This would help to ensure intramolecular elongation rather than intermolecular aggregation.

Consistent with the 'barrel elongation' model, as well as preventing aggregation[487], the role of chaperones is to localise unfolded OMPs to BAM through either non-specific interactions in the case of Skp[152,153], or specific interactions with the BamA POTRA domains in the case of SurA[187]. Chaperones may also modulate the folding pathway by biasing the unfolded ensemble towards conformations that lead to productive folding[97]. In the case of SurA, this may include keeping the nascent OMP in an extended conformation to favour the formation of β-hairpins, and prevent hydrophobic collapse, which would require subsequent structural rearrangement. Consistent with the 'barrel elongation' model, SurA is a potent substoichiometric inhibitor of Aβ40 amyloid formation *in vitro* (unpublished observations, B. Schiffrin, K. L. Stewart, and J. R. Humes (University of Leeds)). Further, Aβ42 oligomers were recently reported to form membrane pores with a β-barrel structure reminiscent of OMPs[488]. SurA may also be involved in modulating the conformation of BamA via its interaction with BamA POTRA 1 (see above), providing an additional reason for the severity of the *ΔsurA* phenotype[145,147,180].

At the present time we only have structures of BAM in essentially two conformations. While these structures must represent free energy minima, they do not exclude the possibility that other conformational changes can occur. Watching BAM in action will require techniques which monitor dynamic changes, such as NMR, AFM and FRET experiments, as well MD simulations[489]. Cross-linking studies during BAM-mediated folding are also likely to be vital in unpicking the conformational changes that occur during the BAM reaction cycle[490-492]. A promising future area of research will be the use of protection strategies, such as HDX[318], coupled with either NMR or MS, or labelling strategies, such as FPOP (Fast Photochemical Oxidation of Proteins)[493], to examine how the folding pathway of OMPs is altered by the presence of chaperones and the BAM complex, in residue-specific detail. All these techniques will be useful in providing evidence for or against the 'barrel elongation' model proposed here. A structure of a stalled intermediate along the BAM-mediated OMP assembly pathway would be very useful in differentiating between different models for the BAM mechanism. Stalled intermediates have been important in elucidating the mechanism of pilus assembly by the FimD usher[494], and efforts to obtain the structure of a BAM complex stalled intermediate are a logical next step[479].

There is much still to be learned about the role that the membrane environment plays in modulating the folding of OMPs, but its importance has been highlighted recently by work which has shown that the trimeric porin OmpF has specific binding sites for LPS, which are essential for its biogenesis *in vivo*[61]. Specific interactions between MPs and lipids are increasingly recognised to have biological significance and studies examining the functional significance of lipid:protein interactions will likely gain importance in the OMP folding field. Recently, native mass spectrometry has emerged as a useful tool to identify the lipids bound to membrane proteins and their complexes[285,495], and may prove invaluable in ascertaining whether BAM selectively binds particular lipids.

Further, progress in the field is likely to come from the improvement of computational techniques. Vast amounts of data on protein sequences, structures and dynamics has been acquired in recent years. The full potential for mining this data is yet to be realised, but improvements have been made recently in computational tools to predict the topologies and structures of OMPs[63,66,496], as well as their thermodynamic stabilities[66].

Greater understanding of OMP biogenesis pathways will lead to strategies for their inhibition. Complete inhibition of BAM complex function may not be required for drugs to be useful. Often the lack of effectiveness of an antibiotic is due to the low permeability of the Gram-negative OM[49]. Lowering the levels of functional BAM leads to an increase in outer membrane permeability, facilitating the entry of other antimicrobial drugs which find it difficult to cross the OM in normal circumstances[497]. Indeed bacterial genetic studies have used a strategy of reducing the permeability of the OM to toxic substances with mutations or deletions of proteins involved in OMP biogenesis, to gain insight into the process by looking for suppressor mutations[498]. Strategies which target BAM function to attenuate virulence without killing cells, may put pathogens under less selection pressure to evolve resistance[274,275]. If cooperativity around the periplasmic ring is important in the BAM mechanism, as proposed in the 'barrel elongation' model, this implies a multitude of potential areas in the complex which could be targeted by rational, structure-based drug design (SBDD)[499].

Overall, the results presented in this thesis provide valuable new insights into OMP folding mechanisms, and in particular the roles of cellular folding factors in OMP folding *in vitro*. The work in this thesis, together with consideration of the research

literature, has led to the new model for OMP assembly by Gram-negative bacteria proposed above. The beauty of modern molecular life science research is that there is an abundance of data from many different techniques, which provide different clues into biological processes. In this chapter I have attempted to collate all the available clues from genetic, biochemical and biophysical studies to provide a view of how the final stages of OMP biogenesis may occur. This offers testable hypotheses for future studies, which are likely to yield many more surprises. We are living through a golden age of stunning discoveries in the OM biogenesis field, and we can no doubt look forward to many more exciting results over the next few years.

References

1. Leewenhoeck, A. V. Observations, Communicated to the Publisher by Mr. Antony van Leewenhoeck, in a Dutch Letter of the 9th of Octob. 1676. Here English'd: concerning Little Animals by Him Observed in Rain-Well-Sea. and Snow Water; as Also in Water Wherein Pepper Had Lain Infused. *Philosophical Transactions of the Royal Society* (1676).
2. Wallin, E. & Heijne, von, G. Genome-wide analysis of integral membrane proteins from eubacterial, archaean, and eukaryotic organisms. *Protein Sci.* **7,** 1029–1038 (1998).
3. Overington, J. P., Al-Lazikani, B. & Hopkins, A. L. How many drug targets are there? *Nat Rev Drug Discov* **5,** 993–996 (2006).
4. Kennedy, S. J. Structures of membrane proteins. *J. Membr. Biol.* **42,** 265–279 (1978).
5. Wimley, W. C. & White, S. H. Experimentally determined hydrophobicity scale for proteins at membrane interfaces. *Nat Struct Mol Biol* **3,** 842–848 (1996).
6. Engelman, D. M. & Steitz, T. A. Identifying nonpolar transbilayer helices in amino acid sequences of membrane proteins. **15,** 321–353 (1986).
7. Kleffel, B., Garavito, R. M., Baumeister, W. & Rosenbusch, J. P. Secondary structure of a channel-forming protein: porin from E. coli outer membranes. *EMBO J.* **4,** 1589–1592 (1985).
8. Popot, J. L. & Engelman, D. M. Membrane protein folding and oligomerization: the two-stage model. *Biochemistry* **29,** 4031–4037 (1990).
9. Hagan, C. L., Silhavy, T. J. & Kahne, D. β-Barrel Membrane Protein Assembly by the Bam Complex. *Annu. Rev. Biochem.* **80,** 189–210 (2011).
10. Goemans, C., Denoncin, K. & Collet, J.-F. Folding mechanisms of periplasmic proteins. *BBA - Molecular Cell Research* **1843,** 1517–1528 (2014).
11. Noinaj, N., Kuszak, A. J., Gumbart, J. C., Lukacik, P., Chang, H., Easley, N. C., Lithgow, T. & Buchanan, S. K. Structural insight into the biogenesis of β-barrel membrane proteins. *Nature* **501,** 385–390 (2013).
12. Knowles, T. J., Scott-Tucker, A., Overduin, M. & Henderson, I. R. Membrane protein architects: the role of the BAM complex in outer membrane protein assembly. *Nat Rev Micro* **7,** 206–214 (2009).
13. Kim, K. H., Aulakh, S. & Paetzel, M. The bacterial outer membrane β-barrel assembly machinery. *Protein Science* **21,** 751–768 (2012).
14. McMorran, L. M., Brockwell, D. J. & Radford, S. E. Mechanistic studies of the biogenesis and folding of outer membrane proteins in vitro and in vivo: What have we learned to date? *Arch Biochem Biophys* **564,** 265–280 (2014).
15. Voulhoux, R. Role of a Highly Conserved Bacterial Protein in Outer Membrane Protein Assembly. *Science* **299,** 262–265 (2003).
16. Noinaj, N., Rollauer, S. E. & Buchanan, S. K. The β-barrel membrane protein insertase machinery from Gram-negative bacteria. *Current Opinion in Structural Biology* **31,** 35–42 (2015).
17. Plummer, A. M. & Fleming, K. G. From Chaperones to the Membrane with a BAM! *Trends in Biochemical Sciences* **In press,** (2016).
18. De Geyter, J., Tsirigotaki, A., Orfanoudaki, G., Zorzini, V., Economou, A. & Karamanou, S. Protein folding in the cell envelope of Escherichia coli. *Nature Microbiology* **1,** 16107 (2016).
19. Silhavy, T. J., Kahne, D. & Walker, S. The Bacterial Cell Envelope. *Cold Spring Harbor Perspectives in Biology* **2,** a000414–a000414 (2010).
20. Goodsell, D. S. Inside a living cell. *Trends in Biochemical Sciences* **16,**

203–206 (1991).
21. Silhavy, T. J., Ruiz, N. & Kahne, D. Advances in understanding bacterial outer-membrane biogenesis. *Nat Rev Micro* **4,** 57–66 (2006).
22. Du, D., Wang, Z., James, N. R., Voss, J. E., Klimont, E., Ohene-Agyei, T., Venter, H., Chiu, W. & Luisi, B. F. Structure of the AcrAB-TolC multidrug efflux pump. *Nature* **509,** 512–515 (2014).
23. Beveridge, T. J. & Graham, L. L. Surface layers of bacteria. *Microbiol. Rev.* **55,** 684–705 (1991).
24. Matias, V. R. F., Al-Amoudi, A., Dubochet, J. & Beveridge, T. J. Cryo-transmission electron microscopy of frozen-hydrated sections of Escherichia coli and Pseudomonas aeruginosa. *J. Bacteriol.* **185,** 6112–6118 (2003).
25. van Meer, G., Voelker, D. R. & Feigenson, G. W. Membrane lipids: where they are and how they behave. *Nat. Rev. Mol. Cell Biol.* **9,** 112–124 (2008).
26. Furse, S. & Scott, D. J. Three-Dimensional Distribution of Phospholipids in Gram Negative Bacteria. *Biochemistry* **55,** 4742–4747 (2016).
27. Shibuya, I., Miyazaki, C. & Ohta, A. Alteration of phospholipid composition by combined defects in phosphatidylserine and cardiolipin synthases and physiological consequences in Escherichia coli. *J. Bacteriol.* **161,** 1086–1092 (1985).
28. Tan, B. K., Bogdanov, M. & Zhao, J. Discovery of a cardiolipin synthase utilizing phosphatidylethanolamine and phosphatidylglycerol as substrates. *Proc. Natl. Acad. Sci. U.S.A.* **109,** 16504–16509 (2012).
29. DeChavigny, A., Heacock, P. N. & Dowhan, W. Sequence and inactivation of the pss gene of Escherichia coli. Phosphatidylethanolamine may not be essential for cell viability. *J. Biol. Chem.* **266,** 5323–5332 (1991).
30. Killian, J. A., Koorengevel, M. C., Bouwstra, J. A., Gooris, G., Dowhan, W. & de Kruijff, B. Effect of divalent cations on lipid organization of cardiolipin isolated from Escherichia coli strain AH930. *Biochim. Biophys. Acta* **1189,** 225–232 (1994).
31. Bogdanov, M., Mileykovskaya, E. & Dowhan, W. in *Lipids in Health and Disease* (eds. Quinn, P. J. & Wang, X.) **49,** 197–239 (Springer Netherlands, 2008).
32. Strandberg, E. & Ulrich, A. S. AMPs and OMPs: Is the folding and bilayer insertion of β-stranded outer membrane proteins governed by the same biophysical principles as for α-helical antimicrobial peptides? *BBA - Biomembranes* **1848,** 1944–1954 (2015).
33. Sprong, H., van der Sluijs, P. & van Meer, G. How proteins move lipids and lipids move proteins. *Nat. Rev. Mol. Cell Biol.* **2,** 504–513 (2001).
34. Lugtenberg, E. J. & Peters, R. Distribution of lipids in cytoplasmic and outer membranes of Escherichia coli K12. *Biochim. Biophys. Acta* **441,** 38–47 (1976).
35. Oliver, P. M., Crooks, J. A., Leidl, M., Yoon, E. J., Saghatelian, A. & Weibel, D. B. Localization of anionic phospholipids in Escherichia coli cells. *J. Bacteriol.* **196,** 3386–3398 (2014).
36. Vanounou, S., Parola, A. H. & Fishov, I. Phosphatidylethanolamine and phosphatidylglycerol are segregated into different domains in bacterial membrane. A study with pyrene-labelled phospholipids. *Mol. Microbiol.* **49,** 1067–1079 (2003).
37. Rassam, P., Copeland, N. A., Birkholz, O., Tóth, C., Chavent, M., Duncan, A. L., Cross, S. J., Housden, N. G., Kaminska, R., Seger, U., Quinn, D. M., Garrod, T. J., Sansom, M. S. P., Piehler, J., Baumann, C. G. & Kleanthous, C. Supramolecular assemblies underpin turnover of outer membrane proteins in bacteria. *Nature* **523,** 333–336 (2015).
38. Diedrich, D. L. & Cota-Robles, E. H. Heterogeneity in Lipid Composition of the Outer Membrane and Cytoplasmic Membrane of Pseudomonas BAL-31.

J. Bacteriol. **119,** 1006–1018 (1974).
39. Ishinaga, M., Kanamoto, R. & Kito, M. Distribution of phospholipid molecular species in outer and cytoplasmic membrane of Escherichia coli. *Journal of Biochemistry* **86,** 161–165 (1979).
40. Denks, K., Vogt, A., Sachelaru, I., Petriman, N.-A., Kudva, R. & Koch, H.-G. The Sec translocon mediated protein transport in prokaryotes and eukaryotes. *Molecular Membrane Biology* **31,** 58–84 (2014).
41. Luirink, J., Heijne, von, G., Houben, E. & de Gier, J.-W. Biogenesis of inner membrane proteins in Escherichia coli. *Annu. Rev. Microbiol.* **59,** 329–355 (2005).
42. Yamaguchi, K., Yu, F. & Inouye, M. A single amino acid determinant of the membrane localization of lipoproteins in E. coli. *Cell* **53,** 423–432 (1988).
43. Labischinski, H., Goodell, E. W., Goodell, A. & Hochberg, M. L. Direct proof of a 'more-than-single-layered' peptidoglycan architecture of Escherichia coli W7: a neutron small-angle scattering study. *J. Bacteriol.* **173,** 751–756 (1991).
44. Yao, X., Jericho, M., Pink, D. & Beveridge, T. Thickness and elasticity of gram-negative murein sacculi measured by atomic force microscopy. *J. Bacteriol.* **181,** 6865–6875 (1999).
45. Vollmer, W. & Seligman, S. J. Architecture of peptidoglycan: more data and more models. *Trends in Microbiology* **18,** 59–66 (2010).
46. Vazquez-Laslop, N., Lee, H., Hu, R. & Neyfakh, A. A. Molecular sieve mechanism of selective release of cytoplasmic proteins by osmotically shocked Escherichia coli. *J. Bacteriol.* **183,** 2399–2404 (2001).
47. Beveridge, T. J. Structures of Gram-Negative Cell Walls and Their Derived Membrane Vesicles. *J. Bacteriol.* **181,** 4725–4733 (1999).
48. Koch, A. L. & Woeste, S. Elasticity of the sacculus of Escherichia coli. *J. Bacteriol.* **174,** 4811–4819 (1992).
49. Nikaido, H. Molecular basis of bacterial outer membrane permeability revisited. *Microbiol. Mol. Biol. Rev.* **67,** 593–656 (2003).
50. Wülfing, C. & Plückthun, A. Protein folding in the periplasm of Escherichia coli. *Mol. Microbiol.* **12,** 685–692 (1994).
51. 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.* **106,** 8009–8014 (2009).
52. Thong, S., Ercan, B., Torta, F., Fong, Z. Y., Wong, H. Y. A., Wenk, M. R. & Chng, S.-S. Defining key roles for auxiliary proteins in an ABC transporter that maintains bacterial outer membrane lipid asymmetry. *eLife* **5,** R53 (2016).
53. Kleinschmidt, J. H. & Tamm, L. K. Secondary and tertiary structure formation of the β-barrel membrane protein OmpA is synchronized and depends on membrane thickness. *J. Mol. Biol.* **324,** 319–330 (2002).
54. Raetz, C. R. H. & Whitfield, C. Lipopolysaccharide Endotoxins. *Annu. Rev. Biochem.* **71,** 635–700 (2002).
55. Wang, L., Wang, Q. & Reeves, P. R. in *Endotoxins: Structure, Function and Recognition* **53,** 123–152 (Springer Netherlands, 2010).
56. Raetz, C. R. H. & Whitfield, C. Lipopolysaccharide Endotoxins. *Annu. Rev. Biochem.* **71,** 635–700 (2002).
57. Okuda, S., Sherman, D. J., Silhavy, T. J., Ruiz, N. & Kahne, D. Lipopolysaccharide transport and assembly at the outer membrane: the PEZ model. *Nat Rev Micro* **14,** 337–345 (2016).
58. Ruiz, N., Kahne, D. & Silhavy, T. J. Transport of lipopolysaccharide across the cell envelope: the long road of discovery. *Nat Rev Micro* **7,** 677–683 (2009).
59. Farris, C., Sanowar, S., Bader, M. W., Pfuetzner, R. & Miller, S. I.

Antimicrobial peptides activate the Rcs regulon through the outer membrane lipoprotein RcsF. *J. Bacteriol.* **192,** 4894–4903 (2010).
60. Konovalova, A., Mitchell, A. M. & Silhavy, T. J. A lipoprotein/β-barrel complex monitors lipopolysaccharide integrity transducing information across the outer membrane. *eLife* **5,** 5312 (2016).
61. Arunmanee, W., Pathania, M., Solovyova, A. S., Le Brun, A. P., Ridley, H., Baslé, A., van den Berg, B. & Lakey, J. H. Gram-negative trimeric porins have specific LPS binding sites that are essential for porin biogenesis. *Proc. Natl. Acad. Sci. U.S.A.* **113,** E5034–43 (2016).
62. Hagge, S. O., De Cock, H., Gutsmann, T., Beckers, F., Seydel, U. & Wiese, A. Pore formation and function of phosphoporin PhoE of Escherichia coli are determined by the core sugar moiety of lipopolysaccharide. *J. Biol. Chem.* **277,** 34247–34253 (2002).
63. Jackups, R., Jr & Liang, J. Interstrand Pairing Patterns in β-Barrel Membrane Proteins: The Positive-outside Rule, Aromatic Rescue, and Strand Registration Prediction. *J. Mol. Biol.* **354,** 979–993 (2005).
64. Heijne, von, G. Membrane protein structure prediction. Hydrophobicity analysis and the positive-inside rule. *J. Mol. Biol.* **225,** 487–494 (1992).
65. Seppala, S., Slusky, J. S., Lloris-Garcera, P., Rapp, M. & Heijne, von, G. Control of Membrane Protein Topology by a Single C-Terminal Residue. *Science* **328,** 1698–1700 (2010).
66. Lin, M., Gessmann, D., Naveed, H. & Liang, J. Outer Membrane Protein Folding and Topology from a Computational Transfer Free Energy Scale. *J. Am. Chem. Soc.* **138,** 2592–2601 (2016).
67. Tokuda, H. Biogenesis of Outer Membranes in Gram-Negative Bacteria. *Bioscience, Biotechnology, and Biochemistry* **73,** 465–473 (2009).
68. Bos, M. P., Robert, V. & Tommassen, J. Biogenesis of the Gram-Negative Bacterial Outer Membrane. *Annu. Rev. Microbiol.* **61,** 191–214 (2007).
69. Konovalova, A. & Silhavy, T. J. Outer membrane lipoprotein biogenesis: Lol is not the end. *Philos. Trans. R. Soc. Lond., B, Biol. Sci.* **370,** 20150030–11 (2015).
70. Braun, V. Covalent lipoprotein from the outer membrane of Escherichia coli. *Biochim. Biophys. Acta* **415,** 335–377 (1975).
71. Konovalova, A., Perlman, D. H., Cowles, C. E. & Silhavy, T. J. Transmembrane domain of surface-exposed outer membrane lipoprotein RcsF is threaded through the lumen of β-barrel proteins. *Proc. Natl. Acad. Sci. U.S.A.* **111,** E4350–8 (2014).
72. Webb, C. T., Selkrig, J., Perry, A. J., Noinaj, N., Buchanan, S. K. & Lithgow, T. Dynamic Association of BAM Complex Modules Includes Surface Exposure of the Lipoprotein BamC. *J. Mol. Biol.* **422,** 545–555 (2012).
73. Cowles, C. E., Li, Y., Semmelhack, M. F., Cristea, I. M. & Silhavy, T. J. The free and bound forms of Lpp occupy distinct subcellular locations in Escherichia coli. *Mol. Microbiol.* **79,** 1168–1181 (2011).
74. de Leij, L. & Witholt, B. Structural heterogeneity of the cytoplasmic and outer membranes of Escherichia coli. *Biochimica et Biophysica Acta (BBA) - Biomembranes* **471,** 92–104 (1977).
75. Dong, C., Beis, K., Nesper, J., Brunkan-LaMontagne, A. L., Clarke, B. R., Whitfield, C. & Naismith, J. H. Wza the translocon for E. coli capsular polysaccharides defines a new class of membrane protein. *Nature* **444,** 226–229 (2006).
76. Margulis, L. *On the origin of mitosing cells*. *J NIH Res* **5,** 65–72 (1967).
77. Eren, E., Murphy, M., Goguen, J. & van den Berg, B. An Active Site Water Network in the Plasminogen Activator Pla from Yersinia pestis. *Structure* **18,** 809–818 (2010).
78. Snijder, H. J., Ubarretxena-Belandia, I., Blaauw, M., Kalk, K. H., Verheij, H.

M., Egmond, M. R., Dekker, N. & Dijkstra, B. W. Structural evidence for dimerization-regulated activation of an integral membrane phospholipase. *Nature* **401,** 717–721 (1999).
79. Meyer, J. E., Hofnung, M. & Schulz, G. E. Structure of maltoporin from Salmonella typhimurium ligated with a nitrophenyl-maltotrioside. *J. Mol. Biol.* **266,** 761–775 (1997).
80. Koronakis, V., Sharff, A., Koronakis, E., Luisi, B. & Hughes, C. Crystal structure of the bacterial membrane protein TolC central to multidrug efflux and protein export. *Nature* **405,** 914–919 (2000).
81. Wimley, W. C. The versatile β-barrel membrane protein. *Current Opinion in Structural Biology* **13,** 404–411 (2003).
82. Jarosławski, S., Duquesne, K., Sturgis, J. N. & Scheuring, S. High-resolution architecture of the outer membrane of the Gram-negative bacteria Roseobacter denitrificans. *Mol. Microbiol.* **74,** 1211–1222 (2009).
83. Li, G.-W., Burkhardt, D., Gross, C. & Weissman, J. S. Quantifying Absolute Protein Synthesis Rates Reveals Principles Underlying Allocation of Cellular Resources. *Cell* **157,** 624–635 (2014).
84. Wilks, J. C. & Slonczewski, J. L. pH of the cytoplasm and periplasm of Escherichia coli: rapid measurement by green fluorescent protein fluorimetry. *J. Bacteriol.* **189,** 5601–5607 (2007).
85. Benz, R., Schmid, A. & Hancock, R. E. Ion selectivity of gram-negative bacterial porins. *J. Bacteriol.* **162,** 722–727 (1985).
86. Schirmer, T., Keller, T. A., Wang, Y. F. & Rosenbusch, J. P. Structural basis for sugar translocation through maltoporin channels at 3.1 A resolution. *Science* **267,** 512–514 (1995).
87. Wandersman, C. & Delepelaire, P. Bacterial Iron Sources: From Siderophores to Hemophores. *Annu. Rev. Microbiol.* **58,** 611–647 (2004).
88. Krewulak, K. D. & Vogel, H. J. TonB or not TonB: is that the question? *Biochem. Cell Biol.* **89,** 87–97 (2011).
89. Fairman, J. W., Noinaj, N. & Buchanan, S. K. The structural biology of β-barrel membrane proteins: a summary of recent reports. *Current Opinion in Structural Biology* **21,** 523–531 (2011).
90. Hiller, S., Garces, R. G., Malia, T. J., Orekhov, V. Y., Colombini, M. & Wagner, G. Solution structure of the integral human membrane protein VDAC-1 in detergent micelles. *Science* **321,** 1206–1210 (2008).
91. Bayrhuber, M., Meins, T., Habeck, M., Becker, S., Giller, K., Villinger, S., Vonrhein, C., Griesinger, C., Zweckstetter, M. & Zeth, K. Structure of the human voltage-dependent anion channel. *Proc. Natl. Acad. Sci. U.S.A.* **105,** 15370–15375 (2008).
92. Ujwal, R., Cascio, D., Colletier, J.-P., Faham, S., Zhang, J., Toro, L., Ping, P. & Abramson, J. The crystal structure of mouse VDAC1 at 2.3 A resolution reveals mechanistic insights into metabolite gating. *Proc. Natl. Acad. Sci. U.S.A.* **105,** 17742–17747 (2008).
93. Koebnik, R., Locher, K. P. & Van Gelder, P. Structure and function of bacterial outer membrane proteins: barrels in a nutshell. *Mol. Microbiol.* **37,** 239–253 (2000).
94. Schulz, G. E. β-barrel membrane proteins. *Current Opinion in Structural Biology* **10,** 443–447 (2000).
95. Sapra, K. T., Damaghi, M., Köster, S., Yildiz, Ö., Kühlbrandt, W. & Müller, D. J. One β Hairpin after the Other: Exploring Mechanical Unfolding Pathways of the Transmembrane β-Barrel Protein OmpG. *Angew. Chem. Int. Ed.* **48,** 8306–8308 (2009).
96. Thoma, J., Bosshart, P., Pfreundschuh, M. & Müller, D. J. Out but not in: the large transmembrane β-barrel protein FhuA unfolds but cannot refold via β-hairpins. *Structure* **20,** 2185–2190 (2012).

97. 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* **22,** 795–802 (2015).
98. Barnhart, M. M. & Chapman, M. R. Curli Biogenesis and Function. *Annu. Rev. Microbiol.* **60,** 131–147 (2006).
99. Goyal, P., Krasteva, P. V., Van Gerven, N., Gubellini, F., Van den Broeck, I., Troupiotis-Tsaïlaki, A., Jonckheere, W., Péhau-Arnaudet, G., Pinkner, J. S., Chapman, M. R., Hultgren, S. J., Howorka, S., Fronzes, R. & Remaut, H. Structural and mechanistic insights into the bacterial amyloid secretion channel CsgG. *Nature* **516,** 250–253 (2014).
100. Reichow, S. L., Korotkov, K. V., Hol, W. G. J. & Gonen, T. Structure of the cholera toxin secretion channel in its closed state. *Nature Structural & Molecular Biology* **17,** 1226–1232 (2010).
101. Huysmans, G. H. M., Guilvout, I. & Pugsley, A. P. Sequential steps in the assembly of the multimeric outer membrane secretin PulD. *J. Biol. Chem.* **288,** 30700–30707 (2013).
102. Huysmans, G. H. M., Guilvout, I., Chami, M., Nickerson, N. N. & Pugsley, A. P. Lipids assist the membrane insertion of a BAM-independent outer membrane protein. *Sci. Rep.* **5,** 15068 (2015).
103. Korotkov, K. V., Gonen, T. & Hol, W. G. J. Secretins: dynamic channels for protein transport across membranes. *Trends in Biochemical Sciences* **36,** 433–443 (2011).
104. Dunstan, R. A., Hay, I. D., Wilksch, J. J., Schittenhelm, R. B., Purcell, A. W., Clark, J., Costin, A., Ramm, G., Strugnell, R. A. & Lithgow, T. Assembly of the secretion pores GspD, Wza and CsgG into bacterial outer membranes does not require the Omp85 proteins BamA or TamA. *Mol. Microbiol.* **97,** 616–629 (2015).
105. Huysmans, G. H. M. Folding outer membrane proteins independently of the β-barrel assembly machinery: an assembly pathway for multimeric complexes? *Biochem. Soc. Trans.* **44,** 845–850 (2016).
106. Celik, N., Webb, C. T., Leyton, D. L., Holt, K. E., Heinz, E., Gorrell, R., Kwok, T., Naderer, T., Strugnell, R. A., Speed, T. P., Teasdale, R. D., Likić, V. A. & Lithgow, T. A Bioinformatic Strategy for the Detection, Classification and Analysis of Bacterial Autotransporters. *PLoS ONE* **7,** e43245–15 (2012).
107. Grijpstra, J., Arenas, J., Rutten, L. & Tommassen, J. Autotransporter secretion: varying on a theme. *Res. Microbiol.* **164,** 562–582 (2013).
108. Bernstein, H. D. Looks can be deceiving: recent insights into the mechanism of protein secretion by the autotransporter pathway. *Mol. Microbiol.* **97,** 205–215 (2015).
109. Bernstein, H. D. Are bacterial 'autotransporters' really transporters? *Trends in Microbiology* **15,** 441–447 (2007).
110. Rossiter, A. E., Leyton, D. L., Tveen-Jensen, K., Browning, D. F., Sevastsyanovich, Y., Knowles, T. J., Nichols, K. B., Cunningham, A. F., Overduin, M., Schembri, M. A. & Henderson, I. R. The essential β-barrel assembly machinery complex components BamD and BamA are required for autotransporter biogenesis. *J. Bacteriol.* **193,** 4250–4253 (2011).
111. 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. U.S.A.* **106,** 19120–19125 (2009).
112. Ieva, R., Tian, P., Peterson, J. H. & Bernstein, H. D. Sequential and spatially restricted interactions of assembly factors with an autotransporter β domain. *Proc. Natl. Acad. Sci. U.S.A.* **108,** E383–E391 (2011).
113. Selkrig, J., Mosbahi, K., Webb, C. T., Belousoff, M. J., Perry, A. J., Wells, T. J., Morris, F., Leyton, D. L., Totsika, M., Phan, M.-D., Celik, N., Kelly, M.,

Oates, C., Hartland, E. L., Robins-Browne, R. M., Ramarathinam, S. H., Purcell, A. W., Schembri, M. A., Strugnell, R. A., Henderson, I. R., Walker, D. & Lithgow, T. Discovery of an archetypal protein transport system in bacterial outer membranes. *Nature Structural & Molecular Biology* **19,** 506–10– S1 (2012).
114. West, M. W., Wang, W., Patterson, J., Mancias, J. D., Beasley, J. R. & Hecht, M. H. De novo amyloid proteins from designed combinatorial libraries. *Proc. Natl. Acad. Sci. U.S.A.* **96,** 11211–11216 (1999).
115. Beck, K., Wu, L. F., Brunner, J. & Müller, M. Discrimination between SRP- and SecA/SecB-dependent substrates involves selective recognition of nascent chains by SRP and trigger factor. *EMBO J.* **19,** 134–143 (2000).
116. D'Souza, V. & Summers, M. F. Structural basis for packaging the dimeric genome of Moloney murine leukaemia virus. *Nature* **431,** 586–590 (2004).
117. Bornemann, T., Holtkamp, W. & Wintermeyer, W. Interplay between trigger factor and other protein biogenesis factors on the ribosome. *Nat Commun* **5,** 4180 (2014).
118. Kumamoto, C. A. Escherichia coli SecB protein associates with exported protein precursors in vivo. *Proceedings of the National Academy of Sciences* **86,** 5320–5324 (1989).
119. Randall, L. L. & Hardy, S. J. S. SecB, one small chaperone in the complex milieu of the cell. *Cellular and Molecular Life Sciences (CMLS)* **59,** 1617–1623 (2002).
120. Watanabe, M. & Blobel, G. SecB functions as a cytosolic signal recognition factor for protein export in E. coli. *Cell* **58,** 695–705 (1989).
121. Hartl, F. U., Lecker, S., Schiebel, E., Hendrick, J. P. & Wickner, W. The binding cascade of SecB to SecA to SecYE mediates preprotein targeting to the E. coli plasma membrane. *Cell* **63,** 269–279 (1990).
122. Zimmer, J., Nam, Y. & Rapoport, T. A. Structure of a complex of the ATPase SecA and the protein-translocation channel. *Nature* **455,** 936–943 (2008).
123. 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.* **189,** 446–454 (2006).
124. Costello, S. M., Plummer, A. M., Fleming, P. J. & Fleming, K. G. Dynamic periplasmic chaperone reservoir facilitates biogenesis of outer membrane proteins. *Proc. Natl. Acad. Sci. U.S.A.* **113,** E4794–800 (2016).
125. Wu, T., Malinverni, J., Ruiz, N., Kim, S., Silhavy, T. J. & Kahne, D. Identification of a multicomponent complex required for outer membrane biogenesis in Escherichia coli. *Cell* **121,** 235–245 (2005).
126. Sklar, J. G., Wu, T., Gronenberg, L. S., Malinverni, J. C., Kahne, D. & Silhavy, T. J. Lipoprotein SmpA is a component of the YaeT complex that assembles outer membrane proteins in Escherichia coli. *Proceedings of the National Academy of Sciences* **104,** 6400–6405 (2007).
127. Struyvé, M., Moons, M. & Tommassen, J. Carboxy-terminal phenylalanine is essential for the correct assembly of a bacterial outer membrane protein. *J. Mol. Biol.* **218,** 141–148 (1991).
128. De Cock, H., Struyvé, M., Kleerebezem, M., van der Krift, T. & Tommassen, J. Role of the carboxy-terminal phenylalanine in the biogenesis of outer membrane protein PhoE of Escherichia coli K-12. *J. Mol. Biol.* **269,** 473–478 (1997).
129. Robert, V., Volokhina, E. B., Senf, F., Bos, M. P., Van Gelder, P. & Tommassen, J. Assembly factor Omp85 recognizes its outer membrane protein substrates by a species-specific C-terminal motif. *PLoS Biol.* **4,** e377 (2006).

130. Paramasivam, N., Habeck, M. & Linke, D. Is the C-terminal insertional signal in Gram-negative bacterial outer membrane proteins species-specific or not? *BMC Genomics* **13,** 510 (2012).
131. Mecsas, J., Rouviere, P. E., Erickson, J. W., Donohue, T. J. & Gross, C. A. The activity of sigma E, an Escherichia coli heat-inducible sigma-factor, is modulated by expression of outer membrane proteins. *Genes & Development* **7,** 2618–2628 (1993).
132. Walsh, N. P., Alba, B. M., Bose, B., Gross, C. A. & Sauer, R. T. OMP peptide signals initiate the envelope-stress response by activating DegS protease via relief of inhibition mediated by its PDZ domain. *Cell* **113,** 61–71 (2003).
133. Kutik, S., Stojanovski, D., Becker, L., Becker, T., Meinecke, M., Krüger, V., Prinz, C., Meisinger, C., Guiard, B., Wagner, R., Pfanner, N. & Wiedemann, N. Dissecting membrane insertion of mitochondrial beta-barrel proteins. *Cell* **132,** 1011–1024 (2008).
134. Walther, D. M., Papic, D., Bos, M. P., Tommassen, J. & Rapaport, D. Signals in bacterial beta-barrel proteins are functional in eukaryotic cells for targeting to and assembly in mitochondria. *Proc. Natl. Acad. Sci. U.S.A.* **106,** 2531–2536 (2009).
135. 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. *Molecular Biology and Evolution* **27,** 887–895 (2010).
136. 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. U.S.A.* **112,** 2011–2016 (2015).
137. Geyer, R., Galanos, C., Westphal, O. & Golecki, J. R. A lipopolysaccharide-binding cell-surface protein from Salmonella minnesota. Isolation, partial characterization and occurrence in different Enterobacteriaceae. *Eur. J. Biochem.* **98,** 27–38 (1979).
138. Holck, A., Lossius, I., Aasland, R. & Kleppe, K. Purification and characterization of the 17 K protein, a DNA-binding protein from Escherichia coli. *Biochim. Biophys. Acta* **914,** 49–54 (1987).
139. Koski, P., Rhen, M., Kantele, J. & Vaara, M. Isolation, cloning, and primary structure of a cationic 16-kDa outer membrane protein of Salmonella typhimurium. *J. Biol. Chem.* **264,** 18973–18980 (1989).
140. Thome, B. M., Hoffschulte, H. K., Schiltz, E. & Müller, M. A protein with sequence identity to Skp (FirA) supports protein translocation into plasma membrane vesicles of Escherichia coli. *FEBS Letters* **269,** 113–116 (1990).
141. Thome, B. M. & Müller, M. Skp is a periplasmic Escherichia coli protein requiring SecA and SecY for export. *Mol. Microbiol.* **5,** 2815–2821 (1991).
142. Aasland, R., Coleman, J., Holck, A. L., Smith, C. L., Raetz, C. R. & Kleppe, K. Identity of the 17-kilodalton protein, a DNA-binding protein from Escherichia coli, and the firA gene product. *J. Bacteriol.* **170,** 5916–5918 (1988).
143. Chen, R. & Henning, U. Aperiplasmic protein (Skp) of Escherichia coli selectively binds a class of outer membrane proteins. *Mol. Microbiol.* **19,** 1287–1294 (1996).
144. Schäfer, U., Beck, K. & Müller, M. Skp, a molecular chaperone of gram-negative bacteria, is required for the formation of soluble periplasmic intermediates of outer membrane proteins. *J. Biol. Chem.* **274,** 24567–24574 (1999).
145. 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 & development* **21,** 2473–2484 (2007).

146. Denoncin, K., Schwalm, J., Vertommen, D., Silhavy, T. J. & Collet, J.-F. Dissecting the Escherichia coliperiplasmic chaperone network using differential proteomics. *Proteomics* **12,** 1391–1401 (2012).
147. Rizzitello, A. E., Harper, J. R. & Silhavy, T. J. Genetic Evidence for Parallel Pathways of Chaperone Activity in the Periplasm of Escherichia coli. *J. Bacteriol.* **183,** 6794–6800 (2001).
148. Harms, N. The Early Interaction of the Outer Membrane Protein PhoE with the Periplasmic Chaperone Skp Occurs at the Cytoplasmic Membrane. *J. Biol. Chem.* **276,** 18804–18811 (2001).
149. Korndörfer, I. P., Dommel, M. K. & Skerra, A. Structure of the periplasmic chaperone Skp suggests functional similarity with cytosolic chaperones despite differing architecture. *Nat Struct Mol Biol* **11,** 1015–1020 (2004).
150. Walton, T. A. & Sousa, M. C. Crystal Structure of Skp, a Prefoldin-like Chaperone that Protects Soluble and Membrane Proteins from Aggregation. *Molecular Cell* **15,** 367–374 (2004).
151. Baker, N. A., Sept, D., Joseph, S., Holst, M. J. & McCammon, J. A. Electrostatics of nanosystems: application to microtubules and the ribosome. *Proc. Natl. Acad. Sci. U.S.A.* **98,** 10037–10041 (2001).
152. McMorran, L. M., Bartlett, A. I., Huysmans, G. H. M., Radford, S. E. & Brockwell, D. J. Dissecting the Effects of Periplasmic Chaperones on the In Vitro Folding of the Outer Membrane Protein PagP. *J. Mol. Biol.* **425,** 3178–3191 (2013).
153. Patel, G. J., Behrens-Kneip, S., Holst, O. & Kleinschmidt, J. H. The Periplasmic Chaperone Skp Facilitates Targeting, Insertion, and Folding of OmpA into Lipid Membranes with a Negative Membrane Surface Potential. *Biochemistry* **48,** 10235–10245 (2009).
154. Grabowicz, M., Koren, D. & Silhavy, T. J. The CpxQ sRNA Negatively Regulates Skp To Prevent Mistargeting of β-Barrel Outer Membrane Proteins into the Cytoplasmic Membrane. *mBio* **7,** e00312–16–8 (2016).
155. Jarchow, S., Lück, C., Görg, A. & Skerra, A. Identification of potential substrate proteins for the periplasmic Escherichia coli chaperone Skp. *Proteomics* **8,** 4987–4994 (2008).
156. 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* **7,** e46068–13 (2012).
157. Entzminger, K. C., Chang, C., Myhre, R. O., McCallum, K. C. & Maynard, J. A. The Skp Chaperone Helps Fold Soluble Proteins in Vitroby Inhibiting Aggregation. *Biochemistry* **51,** 4822–4834 (2012).
158. Bothmann, H. & Plückthun, A. Selection for a periplasmic factor improving phage display and functional periplasmic expression. *Nat. Biotechnol.* **16,** 376–380 (1998).
159. Ha, S. C., Pereira, J. H., Jeong, J. H., Huh, J. H. & Kim, S.-H. Purification of human transcription factors Nanog and Sox2, each in complex with Skp, an Escherichia coli periplasmic chaperone. *Protein Expression and Purification* **67,** 164–168 (2009).
160. 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.* **374,** 91–105 (2007).
161. Callon, M., Burmann, B. M. & Hiller, S. Structural mapping of a chaperone-substrate interaction surface. *Angew. Chem. Int. Ed. Engl.* **53,** 5069–5072 (2014).
162. Wu, S., Ge, X., Lv, Z., Zhi, Z., Chang, Z. & Zhao, X. S. Interaction between bacterial outer membrane proteins and periplasmic quality control factors: a kinetic partitioning mechanism. *Biochem. J.* **438,** 505–511 (2011).

163. Sandlin, C. W., Zaccai, N. R. & Fleming, K. G. Skp Trimer Formation Is Insensitive to Salts in the Physiological Range. *Biochemistry* **54,** 7059–7062 (2015).
164. Tan, A. E., Burgess, N. K., DeAndrade, D. S., Marold, J. D. & Fleming, K. G. Self-Association of Unfolded Outer Membrane Proteins. *Macromol. Biosci.* **10,** 763–767 (2010).
165. Ferguson, A. D., Hofmann, E., Coulton, J. W., Diederichs, K. & Welte, W. Siderophore-mediated iron transport: crystal structure of FhuA with bound lipopolysaccharide. *Science* **282,** 2215–2220 (1998).
166. 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.* **278,** 9092–9099 (2003).
167. Burmann, B. M., Holdbrook, D. A., Callon, M., Bond, P. J. & Hiller, S. Revisiting the Interaction between the Chaperone Skp and Lipopolysaccharide. *Biophysj* **108,** 1516–1526 (2015).
168. Burmann, B. M. & Hiller, S. Chaperones and chaperone-substrate complexes: Dynamic playgrounds for NMR spectroscopists. *Prog Nucl Magn Reson Spectrosc* **86-87,** 41–64 (2015).
169. Huang, C., Rossi, P., Saio, T. & Kalodimos, C. G. Structural basis for the antifolding activity of a molecular chaperone. *Nature* **537,** 202–206 (2016).
170. Walton, T. A., Sandoval, C. M., Fowler, C. A., Pardi, A. & Sousa, M. C. The cavity-chaperone Skp protects its substrate from aggregation but allows independent folding of substrate domains. *Proc. Natl. Acad. Sci. U.S.A.* **106,** 1772–1777 (2009).
171. 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* **20,** 1265–1272 (2013).
172. Zaccai, N. R., Sandlin, C. W., Hoopes, J. T., Curtis, J. E., Fleming, P. J., Fleming, K. G. & Krueger, S. Deuterium Labeling Together with Contrast Variation Small-Angle Neutron Scattering Suggests How Skp Captures and Releases Unfolded Outer Membrane Proteins. *Methods in Enzymology* **566,** 159–210 (2016).
173. Schiffrin, B., Calabrese, A. N., Devine, P. W. A., Harris, S. A., Ashcroft, A. E., Brockwell, D. J. & Radford, S. E. Skp is a multivalent chaperone of outer-membrane proteins. *Nat Struct Mol Biol* **23,** 786–793 (2016).
174. Siegert, R., Leroux, M. R., Scheufler, C., Hartl, F. U. & Moarefi, I. Structure of the molecular chaperone prefoldin: unique interaction of multiple coiled coil tentacles with unfolded proteins. *Cell* **103,** 621–632 (2000).
175. Martín-Benito, J., Boskovic, J., Gómez-Puertas, P., Carrascosa, J. L., Simons, C. T., Lewis, S. A., Bartolini, F., Cowan, N. J. & Valpuesta, J. M. Structure of eukaryotic prefoldin and of its complexes with unfolded actin and the cytosolic chaperonin CCT. *EMBO J.* **21,** 6377–6386 (2002).
176. Tommassen, J. Assembly of outer-membrane proteins in bacteria and mitochondria. *Microbiology* **156,** 2587–2596 (2010).
177. Bina, J. E. & Mekalanos, J. J. Vibrio cholerae tolC is required for bile resistance and colonization. *Infection and Immunity* **69,** 4681–4685 (2001).
178. Schwalm, J., Mahoney, T. F., Soltes, G. R. & Silhavy, T. J. Role for Skp in LptD Assembly in Escherichia coli. *J. Bacteriol.* **195,** 3734–3742 (2013).
179. Volokhina, E. B., Grijpstra, J., Stork, M., Schilders, I., Tommassen, J. & Bos, M. P. Role of the Periplasmic Chaperones Skp, SurA, and DegQ in Outer Membrane Protein Biogenesis in Neisseria meningitidis. *J. Bacteriol.* **193,** 1612–1621 (2011).
180. 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,** 2432–2443 (2009).

181. Tormo, A., Almirón, M. & Kolter, R. surA, an Escherichia coli gene essential for survival in stationary phase. *J. Bacteriol.* **172,** 4339–4347 (1990).
182. Lazar, S. W., Almirón, M., Tormo, A. & Kolter, R. Role of the Escherichia coli SurA protein in stationary-phase survival. *J. Bacteriol.* **180,** 5704–5711 (1998).
183. Rouviere, P. E. & Gross, C. A. SurA, a periplasmic protein with peptidyl-prolyl isomerase activity, participates in the assembly of outer membrane porins. *Genes & Development* **10,** 3170–3182 (1996).
184. Lazar, S. W. & Kolter, R. SurA assists the folding of Escherichia coli outer membrane proteins. *J. Bacteriol.* **178,** 1770–1773 (1996).
185. Behrens, S., Maier, R., De Cock, H., Schmid, F. X. & Gross, C. A. The SurA periplasmic PPIase lacking its parvulin domains functions in vivo and has chaperone activity. *EMBO J.* **20,** 285–294 (2001).
186. 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.* **280,** 23540–23548 (2005).
187. 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.* **77,** 1153–1171 (2010).
188. Wang, Y., Wang, R., Jin, F., Liu, Y., Yu, J., Fu, X. & Chang, Z. A Supercomplex Spanning the Inner and Outer Membranes Mediates the Biogenesis of β-Barrel Outer Membrane Proteins in Bacteria. *J. Biol. Chem.* **291,** 16720–16729 (2016).
189. Bitto, E. & McKay, D. B. Crystallographic structure of SurA, a molecular chaperone that facilitates folding of outer membrane porins. *Structure/Folding and Design* **10,** 1489–1498 (2002).
190. Hoffmann, A., Bukau, B. & Kramer, G. Structure and function of the molecular chaperone Trigger Factor. *BBA - Molecular Cell Research* **1803,** 650–661 (2010).
191. Fiser, A., Do, R. K. & Sali, A. Modeling of loops in protein structures. *Protein Sci.* **9,** 1753–1773 (2000).
192. Alcock, F. H., Grossmann, J. G., Gentle, I. E., Likić, V. A., Lithgow, T. & Tokatlidis, K. Conserved substrate binding by chaperones in the bacterial periplasm and the mitochondrial intermembrane space. *Biochem. J.* **409,** 377–387 (2008).
193. Xu, X., Wang, S., Hu, Y.-X. & McKay, D. B. The Periplasmic Bacterial Molecular Chaperone SurA Adapts its Structure to Bind Peptides in Different Conformations to Assert a Sequence Preference for Aromatic Residues. *J. Mol. Biol.* **373,** 367–381 (2007).
194. Ricci, D. P., Schwalm, J., Gonzales-Cope, M. & Silhavy, T. J. The Activity and Specificity of the Outer Membrane Protein Chaperone SurA Are Modulated by a Proline Isomerase Domain. *mBio* **4,** e00540–13–e00540–13 (2013).
195. Chai, Q., Ferrell, B., Zhong, M., Zhang, X., Ye, C. & Wei, Y. Diverse sequences are functional at the C-terminus of the E. coli periplasmic chaperone SurA. *Protein Engineering Design and Selection* **27,** 111–116 (2014).
196. Zhang, X. C. & Han, L. How does a β-barrel integral membrane protein insert into the membrane? *Protein & Cell* **7,** 471–477 (2016).
197. Bitto, E. & McKay, D. B. Binding of phage-display-selected peptides to the periplasmic chaperone protein SurA mimics binding of unfolded outer membrane proteins. *FEBS Letters* **568,** 94–98 (2004).
198. 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.* **278,** 49316–49322 (2003).
199. Saio, T., Guan, X., Rossi, P., Economou, A. & Kalodimos, C. G. Structural basis for protein antiaggregation activity of the trigger factor chaperone. *Science* **344,** 1250494–1250494 (2014).
200. 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.* **99,** 11020–11024 (2002).
201. Behrens-Kneip, S. The role of SurA factor in outer membrane protein transport and virulence. *Int. J. Med. Microbiol.* **300,** 421–428 (2010).
202. Obi, I. R. & Francis, M. S. Demarcating SurA Activities Required for Outer Membrane Targeting of Yersinia pseudotuberculosis Adhesins. *Infection and Immunity* **81,** 2296–2308 (2013).
203. Lipinska, B., Zylicz, M. & Georgopoulos, C. The HtrA (DegP) protein, essential for Escherichia coli survival at high temperatures, is an endopeptidase. *J. Bacteriol.* **172,** 1791–1797 (1990).
204. Spiess, C., Beil, A. & Ehrmann, M. A temperature-dependent switch from chaperone to protease in a widely conserved heat shock protein. *Cell* **97,** 339–347 (1999).
205. Krojer, T., Sawa, J., Schäfer, E., Saibil, H. R., Ehrmann, M. & Clausen, T. Structural basis for the regulated protease and chaperone function of DegP. *Nature* **453,** 885–890 (2008).
206. Kim, S. & Sauer, R. T. Cage assembly of DegP protease is not required for substrate-dependent regulation of proteolytic activity or high-temperature cell survival. *Proc. Natl. Acad. Sci. U.S.A.* **109,** 7263–7268 (2012).
207. Lyu, Z.-X. & Zhao, X. S. Periplasmic quality control in biogenesis of outer membrane proteins. *Biochem. Soc. Trans.* **43,** 133–138 (2015).
208. Missiakas, D., Betton, J. M. & Raina, S. New components of protein folding in extracytoplasmic compartments of Escherichia coli SurA, FkpA and Skp/OmpH. *Mol. Microbiol.* **21,** 871–884 (1996).
209. Ge, X., Lyu, Z.-X., Liu, Y., Wang, R., Zhao, X. S., Fu, X. & Chang, Z. Identification of FkpA as a key quality control factor for the biogenesis of outer membrane proteins under heat shock conditions. *J. Bacteriol.* **196,** 672–680 (2014).
210. Saul, F. A., Arié, J. P., Vulliez-le Normand, B., Kahn, R., Betton, J. M. & Bentley, G. A. Structural and Functional Studies of FkpA from Escherichia coli, a cis/trans Peptidyl-prolyl Isomerase with Chaperone Activity. *J. Mol. Biol.* **335,** 595–608 (2004).
211. Hu, K., Galius, V. & Pervushin, K. Structural Plasticity of Peptidyl−Prolyl Isomerase sFkpA Is a Key to Its Chaperone Function As Revealed by Solution NMR. *Biochemistry* **45,** 11983–11991 (2006).
212. Quan, S., Koldewey, P., Tapley, T., Kirsch, N., Ruane, K. M., Pfizenmaier, J., Shi, R., Hofmann, S., Foit, L., Ren, G., Jakob, U., Xu, Z., Cygler, M. & Bardwell, J. C. A. Genetic selection designed to stabilize proteins uncovers a chaperone called Spy. *Nat Struct Mol Biol* **18,** 262–269 (2011).
213. Liu, J. & Walsh, C. T. Peptidyl-prolyl cis-trans-isomerase from Escherichia coli: a periplasmic homolog of cyclophilin that is not inhibited by cyclosporin A. *Proc. Natl. Acad. Sci. U.S.A.* **87,** 4028–4032 (1990).
214. Kleerebezem, M., Heutink, M. & Tommassen, J. Characterization of an Escherichia coli rotA mutant, affected in periplasmic peptidyl-prolyl cis/trans isomerase. *Mol. Microbiol.* **18,** 313–320 (1995).
215. Dartigalongue, C. & Raina, S. A new heat-shock gene, ppiD, encodes a peptidyl-prolyl isomerase required for folding of outer membrane proteins in Escherichia coli. *EMBO J.* **17,** 3968–3980 (1998).
216. Justice, S. S., Hunstad, D. A., Harper, J. R., Duguay, A. R., Pinkner, J. S.,

Bann, J., Frieden, C., Silhavy, T. J. & Hultgren, S. J. Periplasmic peptidyl prolyl cis-trans isomerases are not essential for viability, but SurA is required for pilus biogenesis in Escherichia coli. *J. Bacteriol.* **187,** 7680–7686 (2005).
217. Stymest, K. H. & Klappa, P. The periplasmic peptidyl prolyl cis-trans isomerases PpiD and SurA have partially overlapping substrate specificities. *FEBS Journal* **275,** 3470–3479 (2008).
218. Mogensen, J. E. & Otzen, D. E. Interactions between folding factors and bacterial outer membrane proteins. *Mol. Microbiol.* **57,** 326–346 (2005).
219. 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,** 12245–12250 (2010).
220. Denoncin, K., Vertommen, D., Paek, E. & Collet, J.-F. The Protein-disulfide Isomerase DsbC Cooperates with SurA and DsbA in the Assembly of the Essential β-Barrel Protein LptD. *J. Biol. Chem.* **285,** 29425–29433 (2010).
221. Gu, Y., Li, H., Dong, H., Zeng, Y., Zhang, Z., Paterson, N. G., Stansfeld, P. J., Wang, Z., Zhang, Y., Wang, W. & Dong, C. Structural basis of outer membrane protein insertion by the BAM complex. *Nature* **531,** 64–69 (2016).
222. Han, L., Zheng, J., Wang, Y., Yang, X., Liu, Y., Sun, C., Cao, B., Zhou, H., Ni, D., Lou, J., Zhao, Y. & Huang, Y. Structure of the BAM complex and its implications for biogenesis of outer-membrane proteins. *Nat Struct Mol Biol* **23,** 192–196 (2016).
223. Bakelar, J., Buchanan, S. K. & Noinaj, N. The structure of the β-barrel assembly machinery complex. *Science* **351,** 180–186 (2016).
224. Webb, C. T., Heinz, E. & Lithgow, T. Evolution of the β-barrel assembly machinery. *Trends in Microbiology* **20,** 612–620 (2012).
225. Paschen, S. A., Waizenegger, T., Stan, T. & Preuss, M. Evolutionary conservation of biogenesis of β-barrel membrane proteins. *Nature* **426,** 862–866 (2003).
226. Gentle, I. E., Burri, L. & Lithgow, T. Molecular architecture and function of the Omp85 family of proteins. *Mol. Microbiol.* **58,** 1216–1225 (2005).
227. Walther, D. M., Rapaport, D. & Tommassen, J. Biogenesis of β-barrel membrane proteins in bacteria and eukaryotes: evolutionary conservation and divergence. *Cellular and Molecular Life Sciences (CMLS)* **66,** 2789–2804 (2009).
228. Chacinska, A., Koehler, C. M., Milenkovic, D., Lithgow, T. & Pfanner, N. Importing mitochondrial proteins: machineries and mechanisms. *Cell* **138,** 628–644 (2009).
229. Day, P. M., Potter, D. & Inoue, K. Evolution and targeting of Omp85 homologs in the chloroplast outer envelope membrane. *Front. Plant Sci.* **5,** 1241089–19 (2014).
230. Chen, Y.-L., Chen, L.-J. & Li, H.-M. Polypeptide Transport-Associated Domains of the Toc75 Channel Protein Are Located in the Intermembrane Space of Chloroplasts. *Plant Physiol.* **172,** 235–243 (2016).
231. Schleiff, E. & Soll, J. Membrane protein insertion: mixing eukaryotic and prokaryotic concepts. *EMBO Rep* **6,** 1023–1027 (2005).
232. Iadanza, M. G., Higgins, A. J., Schiffrin, B., Calabrese, A. N., Brockwell, D. J., Ashcroft, A. E., Radford, S. E. & Ranson, N. A. Lateral opening in the intact β-barrel assembly machinery captured by cryo-EM. *Nat Commun* **7,** 12865 (2016).
233. Hagan, C. L., Kim, S. & Kahne, D. Reconstitution of outer membrane protein assembly from purified components. *Science* **328,** 890–892 (2010).
234. Knowles, T. J., Browning, D. F., Jeeves, M., Maderbocus, R., Rajesh, S., Sridhar, P., Manoli, E., Emery, D., Sommer, U., Spencer, A., Leyton, D. L.,

Squire, D., Chaudhuri, R. R., Viant, M. R., Cunningham, A. F., Henderson, I. R. & Overduin, M. Structure and function of BamE within the outer membrane and the β-barrel assembly machine. *EMBO Rep* **12,** 123–128 (2011).
235. Albrecht, R. & Zeth, K. Structural Basis of Outer Membrane Protein Biogenesis in Bacteria. *J. Biol. Chem.* **286,** 27792–27803 (2011).
236. Kim, K. H., Kang, H.-S., Okon, M., Escobar-Cabrera, E., McIntosh, L. P. & Paetzel, M. Structural characterization of Escherichia coli BamE, a lipoprotein component of the β-barrel assembly machinery complex. *Biochemistry* **50,** 1081–1090 (2011).
237. Ricci, D. P. & Silhavy, T. J. The Bam machine: A molecular cooper. *BBA - Biomembranes* **1818,** 1067–1084 (2012).
238. Surana, N. K., Grass, S., Hardy, G. G., Li, H., Thanassi, D. G. & Geme, J. W. S. Evidence for conservation of architecture and physical properties of Omp85-like proteins throughout evolution. *Proc. Natl. Acad. Sci. U.S.A.* **101,** 14497–14502 (2004).
239. Roman-Hernandez, G., Peterson, J. H. & Bernstein, H. D. Reconstitution of bacterial autotransporter assembly using purified components. *eLife* **3,** 711–48 (2014).
240. 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).
241. Malinverni, J. C., Werner, J., Kim, S., Sklar, J. G., Kahne, D., Misra, R. & Silhavy, T. J. YfiO stabilizes the YaeT complex and is essential for outer membrane protein assembly in Escherichia coli. *Mol. Microbiol.* **61,** 151–164 (2006).
242. Jacob-Dubuisson, F., Villeret, V., Clantin, B., Delattre, A.-S. & Saint, N. First structural insights into the TpsB/Omp85 superfamily. *Biological Chemistry* **390,** 1–10 (2009).
243. Nesper, J., Brosig, A., Ringler, P., Patel, G. J., Muller, S. A., Kleinschmidt, J. H., Boos, W., Diederichs, K. & Welte, W. Omp85Tt from Thermus thermophilus HB27: an Ancestral Type of the Omp85 Protein Family. *J. Bacteriol.* **190,** 4568–4575 (2008).
244. Arnold, T., Zeth, K. & Linke, D. Omp85 from the Thermophilic Cyanobacterium Thermosynechococcus elongatus Differs from Proteobacterial Omp85 in Structure and Domain Composition. *J. Biol. Chem.* **285,** 18003–18015 (2010).
245. Sánchez-Pulido, L., Devos, D. & Genevrois, S. POTRA: a conserved domain in the FtsQ family and a class of β-barrel outer membrane proteins. *Trends Biochem Sci* **28,** 518–521 (2003).
246. Simmerman, R. F., Dave, A. M. & Bruce, B. D. *Structure and Function of POTRA Domains of Omp85/TPS Superfamily*. *International Review of Cell and Molecular Biology* **308,** 1–34 (Elsevier Inc., 2014).
247. Kim, S., Malinverni, J. C., Sliz, P., Silhavy, T. J., Harrison, S. C. & Kahne, D. Structure and Function of an Essential Component of the Outer Membrane Protein Assembly Machine. *Science* **317,** 961–964 (2007).
248. Gruss, F., Zähringer, F., Jakob, R. P., Burmann, B. M., Hiller, S. & Maier, T. The structural basis of autotransporter translocation by TamA. *Nat Struct Mol Biol* **20,** 1318–1320 (2013).
249. Shen, H.-H., Leyton, D. L., Shiota, T., Belousoff, M. J., Noinaj, N., Lu, J., Holt, S. A., Tan, K., Selkrig, J., Webb, C. T., Buchanan, S. K., Martin, L. L. & Lithgow, T. Reconstitution of a nanomachine driving the assembly of proteins into bacterial outer membranes. *Nat Commun* **5,** 5078 (2014).
250. Browning, D. F., Matthews, S. A., Rossiter, A. E., Sevastsyanovich, Y. R., Jeeves, M., Mason, J. L., Wells, T. J., Wardius, C. A., Knowles, T. J.,

Cunningham, A. F., Bavro, V. N., Overduin, M. & Henderson, I. R. Mutational and Topological Analysis of the Escherichia coli BamA Protein. *PLoS ONE* **8,** e84512–12 (2013).
251. Bos, M. P., Robert, V. & Tommassen, J. Functioning of outer membrane protein assembly factor Omp85 requires a single POTRA domain. *EMBO Rep* **8,** 1149–1154 (2007).
252. Volokhina, E. B., Grijpstra, J., Beckers, F., Lindh, E., Robert, V., Tommassen, J. & Bos, M. P. Species-specificity of the BamA component of the bacterial outer membrane protein-assembly machinery. *PLoS ONE* **8,** e85799 (2013).
253. Dekker, N. Outer-membrane phospholipase A: known structure, unknown biological function. *Mol. Microbiol.* **35,** 711–717 (2000).
254. Browning, D. F., Bavro, V. N., Mason, J. L., Sevastsyanovich, Y. R., Rossiter, A. E., Jeeves, M., Wells, T. J., Knowles, T. J., Cunningham, A. F., Donald, J. W., Palmer, T., Overduin, M. & Henderson, I. R. Cross-species chimeras reveal BamA POTRA and β-barrel domains must be fine-tuned for efficient OMP insertion. *Mol. Microbiol.* **97,** 646–659 (2015).
255. Stegmeier, J. F. Characterization of Pores Formed by YaeT (Omp85) from Escherichia coli. *Journal of Biochemistry* **140,** 275–283 (2006).
256. Knowles, T. J., Jeeves, M., Bobat, S., Dancea, F., McClelland, D., Palmer, T., Overduin, M. & Henderson, I. R. Fold and function of polypeptide transport-associated domains responsible for delivering unfolded proteins to membranes. *Mol. Microbiol.* **68,** 1216–1227 (2008).
257. Gatzeva-Topalova, P. Z., Walton, T. A. & Sousa, M. C. Crystal Structure of YaeT: Conformational Flexibility and Substrate Recognition. *Structure/Folding and Design* **16,** 1873–1881 (2008).
258. 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* **18,** 1492–1501 (2010).
259. Renault, M., Bos, M. P., Tommassen, J. & Baldus, M. Solid-State NMR on a Large Multidomain Integral Membrane Protein: The Outer Membrane Protein Assembly Factor BamA. *J. Am. Chem. Soc.* **133,** 4175–4177 (2011).
260. Jansen, K. B., Baker, S. L. & Sousa, M. C. Crystal Structure of BamB Bound to a Periplasmic Domain Fragment of BamA, the Central Component of the β-Barrel Assembly Machine. *J. Biol. Chem.* **290,** 2126–2136 (2015).
261. Chen, Z., Zhan, L. H., Hou, H. F., Gao, Z. Q., Xu, J. H., Dong, C. & Dong, Y. H. Structural basis for the interaction of BamB with the POTRA3-4 domains of BamA. *Acta Crystallogr D Struct Biol* **72,** 236–244 (2016).
262. Fleming, P. J., Patel, D. S., Wu, E. L., Qi, Y., Yeom, M. S., Sousa, M. C., Fleming, K. G. & Im, W. BamA POTRA Domain Interacts with a Native Lipid Membrane Surface. *Biophysj* **110,** 2698–2709 (2016).
263. 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* **22,** 1055–1062 (2014).
264. Onufryk, C., Crouch, M. L., Fang, F. C. & Gross, C. A. Characterization of Six Lipoproteins in the E Regulon. *J. Bacteriol.* **187,** 4552–4561 (2005).
265. Charlson, E. S., Werner, J. N. & Misra, R. Differential Effects of yfgL Mutation on Escherichia coli Outer Membrane Proteins and Lipopolysaccharide. *J. Bacteriol.* **188,** 7186–7194 (2006).
266. Palomino, C., Marín, E. & Fernández, L. Á. The fimbrial usher FimD follows the SurA-BamB pathway for its assembly in the outer membrane of Escherichia coli. *J. Bacteriol.* **193,** 5222–5230 (2011).
267. Heuck, A., Schleiffer, A. & Clausen, T. Augmenting β-Augmentation: Structural Basis of How BamB Binds BamA and May Support Folding of

Outer Membrane Proteins. *J. Mol. Biol.* **406,** 659–666 (2011).
268. Jansen, K. B., Baker, S. L. & Sousa, M. C. Crystal Structure of BamB from Pseudomonas aeruginosa and Functional Evaluation of Its Conserved Structural Features. *PLoS ONE* **7,** e49749–9 (2012).
269. O'Neil, P. K., Rollauer, S. E., Noinaj, N. & Buchanan, S. K. Fitting the Pieces of the β-Barrel Assembly Machinery Complex. *Biochemistry* **54,** 6303–6311 (2015).
270. Kim, K. H., Aulakh, S. & Paetzel, M. Crystal Structure of -Barrel Assembly Machinery BamCD Protein Complex. *J. Biol. Chem.* **286,** 39116–39121 (2011).
271. Dong, C., Yang, X., Hou, H. F., Shen, Y.-Q. & Dong, Y. H. Structure of Escherichia coli BamB and its interaction with POTRA domains of BamA. *Acta Crystallogr. D Biol. Crystallogr.* **68,** 1134–1139 (2012).
272. Noinaj, N., Fairman, J. W. & Buchanan, S. K. The Crystal Structure of BamB Suggests Interactions with BamA and Its Role within the BAM Complex. *J. Mol. Biol.* **407,** 248–260 (2011).
273. Kim, K. H. & Paetzel, M. Crystal Structure of Escherichia coli BamB, a Lipoprotein Component of the β-Barrel Assembly Machinery Complex. *J. Mol. Biol.* **406,** 667–678 (2011).
274. Hsieh, P.-F., Hsu, C.-R., Chen, C.-T., Lin, T.-L. & Wang, J.-T. The Klebsiella pneumoniaeYfgL (BamB) lipoprotein contributes to outer membrane protein biogenesis, type-1 fimbriae expression, anti-phagocytosis, and in vivovirulence. *Virulence* **7,** 587–601 (2016).
275. Krachler, A. M. BamB and outer membrane biogenesis - The Achilles' heel for targeting Klebsiella infections? *Virulence* **7,** 508–511 (2016).
276. Rigel, N. W., Schwalm, J., Ricci, D. P. & Silhavy, T. J. BamE Modulates the Escherichia coli Beta-Barrel Assembly Machine Component BamA. *J. Bacteriol.* **194,** 1002–1008 (2012).
277. Kim, K. H., Aulakh, S., Tan, W. & Paetzel, M. Crystallographic analysis of the C-terminal domain of the Escherichia coli lipoprotein BamC. *Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun.* **67,** 1350–1358 (2011).
278. Dong, C., Hou, H. F., Yang, X., Shen, Y.-Q. & Dong, Y. H. Structure of Escherichia coli BamD and its functional implications in outer membrane protein assembly. *Acta Crystallogr. D Biol. Crystallogr.* **68,** 95–101 (2012).
279. Sandoval, C. M., Baker, S. L., Jansen, K., Metzner, S. I. & Sousa, M. C. Crystal Structure of BamD: An Essential Component of the β-Barrel Assembly Machinery of Gram-Negative Bacteria. *J. Mol. Biol.* **409,** 348–357 (2011).
280. Scheufler, C., Brinker, A., Bourenkov, G., Pegoraro, S., Moroder, L., Bartunik, H., Hartl, F. U. & Moarefi, I. Structure of TPR Domain–Peptide Complexes. *Cell* **101,** 199–210 (2000).
281. Gatto, G. J., Geisbrecht, B. V. & Gould, S. J. Peroxisomal targeting signal-1 recognition by the TPR domains of human PEX5. *Nat Struct Mol Biol* **7,** 1091–1095 (2000).
282. 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,** 3487–3491 (2012).
283. Hagan, C. L., Westwood, D. B. & Kahne, D. Bam Lipoproteins Assemble BamA in Vitro. *Biochemistry* **52,** 6108–6113 (2013).
284. Endo, T., Kawano, S. & Yamano, K. BamE structure: the assembly of β-barrel proteins in the outer membranes of bacteria and mitochondria. *EMBO Rep* **12,** 94–95 (2011).
285. Laganowsky, A., Reading, E., Allison, T. M., Ulmschneider, M. B., Degiacomi, M. T., Baldwin, A. J. & Robinson, C. V. Membrane proteins bind

lipids selectively to modulate their structure and function. *Nature* **510,** 172–175 (2015).
286. Majdalani, N. & Gottesman, S. The Rcs phosphorelay: a complex signal transduction system. *Annu. Rev. Microbiol.* **59,** 379–405 (2005).
287. Ni, D., Wang, Y., Yang, X., Zhou, H., Hou, X., Cao, B., Lu, Z., Zhao, X., Yang, K. & Huang, Y. Structural and functional analysis of the β-barrel domain of BamA from Escherichia coli. *FASEB J.* **28,** 2677–2685 (2014).
288. Frauenfeld, J., Gumbart, J., van der Sluis, E. O., Funes, S., Gartmann, M., Beatrix, B., Mielke, T., Berninghausen, O., Becker, T., Schulten, K. & Beckmann, R. Cryo-EM structure of the ribosome–SecYE complex in the membrane environment. *Nature Structural & Molecular Biology* **18,** 614–621 (2011).
289. Hagn, F., Etzkorn, M., Raschle, T. & Wagner, G. Optimized Phospholipid Bilayer Nanodiscs Facilitate High-Resolution Structure Determination of Membrane Proteins. *J. Am. Chem. Soc.* **135,** 1919–1925 (2013).
290. Kleanthous, C., Rassam, P. & Baumann, C. G. Protein–protein interactions and the spatiotemporal dynamics of bacterial outer membrane proteins. *Current Opinion in Structural Biology* **35,** 109–115 (2015).
291. Stanley, A. M. & Fleming, K. G. The process of folding proteins into membranes: Challenges and progress. *Arch. Biochem. Biophys.* **469,** 46–66 (2008).
292. Bowie, J. U. Solving the membrane protein folding problem. *Nature* **438,** 581–589 (2005).
293. Shanmugavadivu, B., Apell, H.-J., Meins, T., Zeth, K. & Kleinschmidt, J. H. Correct Folding of the β-Barrel of the Human Membrane Protein VDAC Requires a Lipid Bilayer. *J. Mol. Biol.* **368,** 66–78 (2007).
294. White, S. H. & Wimley, W. C. Membrane protein folding and stability: physical principles. *Annu Rev Biophys Biomol Struct* **28,** 319–365 (1999).
295. Koebnik, R. In vivo membrane assembly of split variants of the E.coli outer membrane protein OmpA. *EMBO J.* **15,** 3529–3537 (1996).
296. Booth, P. J. & Curran, A. R. Membrane protein folding. *Current Opinion in Structural Biology* **9,** 115–121 (1999).
297. Surrey, T. & Jähnig, F. Refolding and oriented insertion of a membrane protein into a lipid bilayer. *Proceedings of the National Academy of Sciences* **89,** 7457–7461 (1992).
298. Popot, J. L. Folding membrane proteins in vitro: a table and some comments. *Arch Biochem Biophys* **564,** 314–326 (2014).
299. 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* **52,** 3974–3986 (2013).
300. Gessmann, D., Chung, Y. H., Danoff, E. J., Plummer, A. M., Sandlin, C. W., Zaccai, N. R. & Fleming, K. G. Outer membrane β-barrel protein folding is physically controlled by periplasmic lipid head groups and BamA. *Proc. Natl. Acad. Sci. U.S.A.* **111,** 5878–5883 (2014).
301. Hagan, C. L. & Kahne, D. The reconstituted Escherichia coli Bam complex catalyzes multiple rounds of β-barrel assembly. *Biochemistry* **50,** 7444–7446 (2011).
302. Plummer, A. M. & Fleming, K. G. BamA Alone Accelerates Outer Membrane Protein Folding In Vitro through a Catalytic Mechanism. *Biochemistry* **54,** 6009–6011 (2015).
303. Pautsch, A. & Schulz, G. E. High-resolution structure of the OmpA membrane domain. *J. Mol. Biol.* **298,** 273–282 (2000).
304. Ishida, H., Garcia-Herrero, A. & Vogel, H. J. The periplasmic domain of Escherichia coli outer membrane protein A can undergo a localized

temperature dependent structural transition. *Biochim. Biophys. Acta* **1838,** 3014–3024 (2014).
305. Tamm, L. K., Hong, H. & Liang, B. Folding and assembly of β-barrel membrane proteins. *Biochimica et Biophysica Acta (BBA) - Biomembranes* **1666,** 250–263 (2004).
306. Smith, S. G. J., Mahon, V., Lambert, M. A. & Fagan, R. P. A molecular Swiss army knife: OmpA structure, function and expression. *FEMS Microbiology Letters* **273,** 1–11 (2007).
307. Dornmair, K., Kiefer, H. & Jähnig, F. Refolding of an integral membrane protein. OmpA of Escherichia coli. *J. Biol. Chem.* **265,** 18907–18911 (1990).
308. Inouye, M. & Yee, M.-L. Homogeneity of Envelope Proteins of Escherichia coli Separated by Gel Electrophoresis in Sodium Dodecyl Sulfate. *J. Bacteriol.* **113,** 304–312 (1973).
309. Schweizer, M., Hindennach, I., Garten, W. & Henning, U. Major proteins of the Escherichia coli outer cell envelope membrane. Interaction of protein II with lipopolysaccharide. *Eur. J. Biochem.* **82,** 211–217 (1978).
310. Nakamura, K. & Mizushima, S. Effects of heating in dodecyl sulfate solution on the conformation and electrophoretic mobility of isolated major outer membrane proteins from Escherichia coli K-12. *Journal of Biochemistry* **80,** 1411–1422 (1976).
311. Surrey, T., Schmid, A. & Jähnig, F. Folding and membrane insertion of the trimeric beta-barrel protein OmpF. *Biochemistry* **35,** 2283–2288 (1996).
312. Kleinschmidt, J. H., Wiener, M. C. & Tamm, L. K. Outer membrane protein A of E. coli folds into detergent micelles, but not in the presence of monomeric detergent. *Protein Sci.* **8,** 2065–2071 (1999).
313. 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.* **283,** 26748–26758 (2008).
314. Kleinschmidt, J. H. & Tamm, L. K. Folding intermediates of a beta-barrel membrane protein. Kinetic evidence for a multi-step membrane insertion mechanism. *Biochemistry* **35,** 12993–13000 (1996).
315. Rodionova, N. A., Tatulian, S. A., Surrey, T. & Jaehnig, F. Characterization of two membrane-bound forms of OmpA. *Biochemistry* **34,** 1921–1929 (1995).
316. Kleinschmidt, J. H. & Tamm, L. K. Time-resolved distance determination by tryptophan fluorescence quenching: probing intermediates in membrane protein folding. *Biochemistry* **38,** 4996–5005 (1999).
317. Kleinschmidt, J. H., Blaauwen, den, T., Driessen, A. J. & Tamm, L. K. Outer membrane protein A of Escherichia coli inserts and folds into lipid bilayers by a concerted mechanism. *Biochemistry* **38,** 5006–5016 (1999).
318. Raschle, T., Rios Flores, P., Opitz, C., Müller, D. J. & Hiller, S. Monitoring Backbone Hydrogen-Bond Formation in β-Barrel Membrane Protein Folding. *Angew. Chem.* **128,** 6056–6059 (2016).
319. Pocanschi, C. L., Apell, H.-J., Puntervoll, P., Høgh, B., Jensen, H. B., Welte, W. & Kleinschmidt, J. H. The Major Outer Membrane Protein of Fusobacterium nucleatum (FomA) Folds and Inserts into Lipid Bilayers via Parallel Folding Pathways. *J. Mol. Biol.* **355,** 548–561 (2006).
320. Kleinschmidt, J. H. Folding kinetics of the outer membrane proteins OmpA and FomA into phospholipid bilayers. *Chemistry and Physics of Lipids* **141,** 30–47 (2006).
321. Andersen, K. K., Wang, H. & Otzen, D. E. A Kinetic Analysis of the Folding and Unfolding of OmpA in Urea and Guanidinium Chloride: Single and Parallel Pathways. *Biochemistry* **51,** 8371–8383 (2012).
322. Otzen, D. E. & Andersen, K. K. Folding of outer membrane proteins. *Arch Biochem Biophys* **531,** 34–43 (2013).

323. Huysmans, G. H. M., Radford, S. E., Baldwin, S. A. & Brockwell, D. J. Malleability of the Folding Mechanism of the Outer Membrane Protein PagP: Parallel Pathways and the Effect of Membrane Elasticity. *J. Mol. Biol.* **416,** 453–464 (2012).
324. Kleinschmidt, J. H., Bulieris, P. V., Qu, J., Dogterom, M. & Blaauwen, den, T. Association of Neighboring β-Strands of Outer Membrane Protein A in Lipid Bilayers Revealed by Site-Directed Fluorescence Quenching. *J. Mol. Biol.* **407,** 316–332 (2011).
325. Koebnik, R. Structural and functional roles of the surface-exposed loops of the beta-barrel membrane protein OmpA from Escherichia coli. *J. Bacteriol.* **181,** 3688–3694 (1999).
326. Koebnik, R. Membrane assembly of the Escherichia coli outer membrane protein OmpA: exploring sequence constraints on transmembrane beta-strands. *J. Mol. Biol.* **285,** 1801–1810 (1999).
327. Stapleton, J. A., Whitehead, T. A. & Nanda, V. Computational redesign of the lipid-facing surface of the outer membrane protein OmpA. *Proc. Natl. Acad. Sci. U.S.A.* **112,** 9632–9637 (2015).
328. De Cock, H. & Tommassen, J. Lipopolysaccharides and divalent cations are involved in the formation of an assembly-competent intermediate of outer-membrane protein PhoE of E.coli. *EMBO J.* **15,** 5567–5573 (1996).
329. De Cock, H., Schäfer, U., Potgeter, M., Demel, R., Müller, M. & Tommassen, J. Affinity of the periplasmic chaperone Skp of Escherichia coli for phospholipids, lipopolysaccharides and non-native outer membrane proteins. *European Journal of Biochemistry* **259,** 96–103 (1999).
330. Steeghs, L., Hartog, den, R., Boer, den, A., Zomer, B., Roholl, P. & van der Ley, P. Meningitis bacterium is viable without endotoxin. *Nature* **392,** 449–450 (1998).
331. Hong, H. & Tamm, L. K. Elastic coupling of integral membrane protein stability to lipid bilayer forces. *Proc. Natl. Acad. Sci. U.S.A.* **101,** 4065–4070 (2004).
332. 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.* **107,** 4099–4104 (2010).
333. Moon, C. P., Kwon, S. & Fleming, K. G. Overcoming Hysteresis to Attain Reversible Equilibrium Folding for Outer Membrane Phospholipase A in Phospholipid Bilayers. *J. Mol. Biol.* **413,** 484–494 (2011).
334. Nandi, L. G., Guerra, J. P. T. A., Bellettini, I. C., Machado, V. G. & Minatti, E. Properties of aqueous solutions of lentinan in the absence and presence of zwitterionic surfactants. *Carbohydr Polym* **98,** 1–7 (2013).
335. Moon, C. P. & Fleming, K. G. Side-chain hydrophobicity scale derived from transmembrane protein folding into lipid bilayers. *Proc. Natl. Acad. Sci. U.S.A.* **108,** 10174–10177 (2011).
336. Moon, C. P., Zaccai, N. R., Fleming, P. J., Gessmann, D. & Fleming, K. G. Membrane protein thermodynamic stability may serve as the energy sink for sorting in the periplasm. *Proc. Natl. Acad. Sci. U.S.A.* **110,** 4285–4290 (2013).
337. Pocanschi, C. L., Popot, J. L. & Kleinschmidt, J. H. Folding and stability of outer membrane protein A (OmpA) from Escherichia coli in an amphipathic polymer, amphipol A8-35. *Eur Biophys J* **42,** 103–118 (2013).
338. Popot, J. L., Althoff, T., Bagnard, D., Banères, J. L., Bazzacco, P., Billon-Denis, E., Catoire, L. J., Champeil, P., Charvolin, D., Cocco, M. J., Crémel, G., Dahmane, T., la Maza, de, L. M., Ebel, C., Gabel, F., Giusti, F., Gohon, Y., Goormaghtigh, E., Guittet, E., Kleinschmidt, J. H., Kühlbrandt, W., Le Bon, C., Martinez, K. L., Picard, M., Pucci, B., Sachs, J. N., Tribet, C., van Heijenoort, C., Wien, F., Zito, F. & Zoonens, M. Amphipols from A to Z.

Annu. Rev. Biophys. **40,** 379–408 (2011).
339. Bishop, R. E., Gibbons, H. S., Guina, T., Trent, M. S., Miller, S. I. & Raetz, C. R. H. Transfer of palmitate from phospholipids to lipid A in outer membranes of Gram-negative bacteria. *EMBO J.* **19,** 5071–5080 (2000).
340. Bishop, R. E. Structural biology of membrane-intrinsic β-barrel enzymes: Sentinels of the bacterial outer membrane. *Biochimica et Biophysica Acta (BBA) - Biomembranes* **1778,** 1881–1896 (2008).
341. Guo, L., Lim, K. B., Poduje, C. M., Daniel, M. & Gunn, J. S. Lipid A acylation and bacterial resistance against vertebrate antimicrobial peptides. *Cell* **95,** 189–198 (1998).
342. Bishop, R. E. The lipid A palmitoyltransferase PagP: molecular mechanisms and role in bacterial pathogenesis. *Mol. Microbiol.* **57,** 900–912 (2005).
343. Hwang, P. M., Choy, W.-Y., Lo, E. I., Chen, L., Forman-Kay, J. D., Raetz, C. R. H., Privé, G. G., Bishop, R. E. & Kay, L. E. Solution structure and dynamics of the outer membrane enzyme PagP by NMR. *Proc. Natl. Acad. Sci. U.S.A.* **99,** 13560–13565 (2002).
344. Ahn, V. E., Lo, E. I., Engel, C. K., Chen, L., Hwang, P. M., Kay, L. E., Bishop, R. E. & Privé, G. G. A hydrocarbon ruler measures palmitate in the enzymatic acylation of endotoxin. *EMBO J.* **23,** 2931–2941 (2004).
345. Cuesta-Seijo, J. A., Neale, C., Khan, M. A., Moktar, J., Tran, C. D., Bishop, R. E., Pomès, R. & Privé, G. G. PagP Crystallized from SDS/Cosolvent Reveals the Route for Phospholipid Access to the Hydrocarbon Ruler. *Structure/Folding and Design* **18,** 1210–1219 (2010).
346. Bannwarth, M. & Schulz, G. E. The expression of outer membrane proteins for crystallization. *Biochimica et Biophysica Acta (BBA) - Biomembranes* **1610,** 37–45 (2003).
347. Hwang, P. M., Pan, J. S. & Sykes, B. D. A PagP fusion protein system for the expression of intrinsically disordered proteins in Escherichia coli. *Protein Expression and Purification* **85,** 148–151 (2012).
348. Huysmans, G. H. M., Radford, S. E., Brockwell, D. J. & Baldwin, S. A. The N-terminal Helix Is a Post-assembly Clamp in the Bacterial Outer Membrane Protein PagP. *J. Mol. Biol.* **373,** 529–540 (2007).
349. Adil Khan, M., Neale, C., Michaux, C., Pomès, R., Privé, G. G., Woody, R. W. & Bishop, R. E. Gauging a Hydrocarbon Ruler by an Intrinsic Exciton Probe. *Biochemistry* **46,** 4565–4579 (2007).
350. Fersht, A. R., Matouschek, A. & Serrano, L. The folding of an enzyme: I. Theory of protein engineering analysis of stability and pathway of protein folding. *J. Mol. Biol.* **224,** 771–782 (1992).
351. Matouschek, A., Kellis, J. T., Serrano, L. & Fersht, A. R. Mapping the transition state and pathway of protein folding by protein engineering. *Nature* **340,** 122–126 (1989).
352. Bond, P. J. & Sansom, M. S. P. Insertion and assembly of membrane proteins via simulation. *J. Am. Chem. Soc.* **128,** 2697–2704 (2006).
353. Curnow, P. & Booth, P. J. Combined kinetic and thermodynamic analysis of alpha-helical membrane protein unfolding. *Proc. Natl. Acad. Sci. U.S.A.* **104,** 18970–18975 (2007).
354. Curnow, P. & Booth, P. J. The transition state for integral membrane protein folding. *Proc. Natl. Acad. Sci. U.S.A.* **106,** 773–778 (2009).
355. Kramer, R. A., Zandwijken, D., Egmond, M. R. & Dekker, N. In vitro folding, purification and characterization of Escherichia coli outer membrane protease ompT. *Eur. J. Biochem.* **267,** 885–893 (2000).
356. Jansen, C., Heutink, M., Tommassen, J. & De Cock, H. The assembly pathway of outer membrane protein PhoE of Escherichia coli. *Eur. J. Biochem.* **267,** 3792–3800 (2000).
357. Norell, D., Heuck, A., Tran-Thi, T.-A., Götzke, H., Jacob-Dubuisson, F.,

Clausen, T., Daley, D. O., Braun, V., Müller, M. & Fan, E. Versatile in vitro system to study translocation and functional integration of bacterial outer membrane proteins. *Nat Commun* **5,** 5396 (2014).
358. NEBasechanger. *httpnebasechanger.neb.com* at <http://nebasechanger.neb.com>
359. McMorran, L. M. Mechanisms of outer membrane protein folding: effects of lipid environment and periplasmic chaperones. PhD thesis, University of Leeds (2013). at <http://etheses.whiterose.ac.uk/5906/1/McMorran_LM_Astbury_PhD_2013.pdf>
360. Eurofins Genomics. pEX-K vector map, multiple cloning site and sequence. *www.operon.comproductsgene-synthesisimagespEX-KMap.pdf*
361. Crosskey, T. D. Investigating the folding and insertion of bacterial outer membrane proteins by atomic force microscopy. Master's thesis, University of Leeds (2015). (2015).
362. Schneider, C. A., Rasband, W. S. & Eliceiri, K. W. NIH Image to ImageJ: 25 years of image analysis. *Nature Methods* **9,** 671–675 (2012).
363. Wilkins, M. R., Gasteiger, E., Bairoch, A., Sanchez, J. C., Williams, K. L., Appel, R. D. & Hochstrasser, D. F. Protein identification and analysis tools in the ExPASy server. *Methods Mol. Biol.* **112,** 531–552 (1999).
364. The ExPASy Protparam Server. *httpweb.expasy.orgprotparam* at <http://web.expasy.org/protparam/>
365. Danoff, E. J. & Fleming, K. G. Membrane Defects Accelerate Outer Membrane β-Barrel Protein Folding. *Biochemistry* **54,** 97–99 (2015).
366. Welcome to Python.org. (2001). at <www.python.org>
367. Jackson, S. E. & Fersht, A. R. Folding of chymotrypsin inhibitor 2. 1. Evidence for a two-state transition. *Biochemistry* **30,** 10428–10435 (1991).
368. Nicholson, E. M. & Scholtz, J. M. Conformational stability of the Escherichia coli HPr protein: test of the linear extrapolation method and a thermodynamic characterization of cold denaturation. *Biochemistry* **35,** 11369–11378 (1996).
369. Morgner, N. & Robinson, C. V. Massign: An Assignment Strategy for Maximizing Information from the Mass Spectra of Heterogeneous Protein Assemblies. *Anal. Chem.* **84,** 2939–2948 (2012).
370. Ruotolo, B. T., Benesch, J. L. P., Sandercock, A. M., Hyung, S.-J. & Robinson, C. V. Ion mobility–mass spectrometry analysis of large protein complexes. *Nat Protoc* **3,** 1139–1152 (2008).
371. Bush, M. F., Hall, Z., Giles, K., Hoyes, J., Robinson, C. V. & Ruotolo, B. T. Collision Cross Sections of Proteins and Their Complexes: A Calibration Framework and Database for Gas-Phase Structural Biology. *Anal. Chem.* **82,** 9557–9565 (2010).
372. Smith, D. P., Knapman, T. W., Campuzano, I., Malham, R. W., Berryman, J. T., Radford, S. E. & Ashcroft, A. E. Deciphering drift time measurements from travelling wave ion mobility spectrometry-mass spectrometry studies. *Eur J Mass Spectrom (Chichester, Eng)* **15,** 113–130 (2009).
373. Benesch, J. L. P., Ruotolo, B. T., Simmons, D. A. & Robinson, C. V. Protein complexes in the gas phase: technology for structural genomics and proteomics. *Chem. Rev.* **107,** 3544–3567 (2007).
374. Schrödinger, LLC. The PyMOL Molecular Graphics System, Version 1.8. (2015).
375. Marklund, E. G., Degiacomi, M. T., Robinson, C. V., Baldwin, A. J. & Benesch, J. L. P. Collision Cross Sections for Structural Proteomics. *Structure/Folding and Design* **23,** 791–799 (2015).
376. Case, D. A., Berryman, J. T., Betz, R. M., Cerutti, D. S., Cheatham, T. E.,

III, Darden, T. A., Duke, R. E., Giese, T. J., Gohlke, H., Goetz, A. W., Homeyer, N., Izadi, S., Janowski, P., Kaus, J., Kovalenko, A., Lee, T. S., LeGrand, S., Li, P., Luchko, T., Luo, R., Madej, B., Merz, K. M., Monard, G., Needham, P., Nguyen, H., Nguyen, H. T., Omelyan, I., Onufriev, A., Roe, D. R., Roitberg, A., Salomon-Ferrer, R., Simmerling, C. L., Smith, W., Swails, J., Walker, R. C., Wang, J., Wolf, R. M., Wu, X., York, D. M. & Kollman, P. A. AMBER 15. *University of California, San Francisco* (2015).

377. Maier, J. A., Martinez, C., Kasavajhala, K., Wickstrom, L., Hauser, K. E. & Simmerling, C. ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB. *J Chem Theory Comput* **11,** 3696–3713 (2015).
378. Tsui, V. & Case, D. A. Theory and applications of the generalized Born solvation model in macromolecular simulations. *Biopolymers* **56,** 275–291 (2000).
379. Weiser, J., Shenkin, P. S. & Still, W. C. Approximate atomic surfaces from linear combinations of pairwise overlaps (LCPO). *Journal of Computational Chemistry* **20,** 217–230 (1999).
380. Irobalieva, R. N., Fogg, J. M., Catanese, D. J., Sutthibutpong, T., Chen, M., Barker, A. K., Ludtke, S. J., Harris, S. A., Schmid, M. F., Chiu, W. & Zechiedrich, L. Structural diversity of supercoiled DNA. *Nat Commun* **6,** 8440 (2015).
381. Humphrey, W., Dalke, A. & Schulten, K. VMD: visual molecular dynamics. *J Mol Graph* **14,** 33–8– 27–8 (1996).
382. Roe, D. R. & Cheatham, T. E., III. PTRAJ and CPPTRAJ: Software for Processing and Analysis of Molecular Dynamics Trajectory Data. *J Chem Theory Comput* **9,** 3084–3095 (2013).
383. Kleinschmidt, J. H. Folding of β-barrel membrane proteins in lipid bilayers — Unassisted and assisted folding and insertion. *BBA - Biomembranes* **1848,** 1927–1943 (2015).
384. Dewald, A. H., Hodges, J. C. & Columbus, L. Physical determinants of β-barrel membrane protein folding in lipid vesicles. *Biophys. J.* **100,** 2131–2140 (2011).
385. Buchanan, S. K. Beta-barrel proteins from bacterial outer membranes: structure, function and refolding. *Current Opinion in Structural Biology* **9,** 455–461 (1999).
386. Dong, H., Xiang, Q., Gu, Y., Wang, Z., Paterson, N. G., Stansfeld, P. J., He, C., Zhang, Y., Wang, W. & Dong, C. Structural basis for outer membrane lipopolysaccharide insertion. *Nature* **511,** 52–56 (2014).
387. Kleinschmidt, J. H. Membrane protein folding on the example of outer membrane protein A of Escherichia coli. *Cellular and Molecular Life Sciences (CMLS)* **60,** 1547–1558 (2003).
388. Mahalakshmi, R. & Marassi, F. M. Orientation of the Escherichia coli Outer Membrane Protein OmpX in Phospholipid Bilayer Membranes Determined by Solid-State NMR. *Biochemistry* **47,** 6531–6538 (2008).
389. Stanley, A. M., Treubrodt, A. M., Chuawong, P., Hendrickson, T. L. & Fleming, K. G. Lipid Chain Selectivity by Outer Membrane Phospholipase A. *J. Mol. Biol.* **366,** 461–468 (2007).
390. Conlan, S. & Bayley, H. Folding of a Monomeric Porin, OmpG, in Detergent Solution †. *Biochemistry* **42,** 9453–9465 (2003).
391. Surrey, T. & Jähnig, F. Kinetics of folding and membrane insertion of a beta-barrel membrane protein. *J. Biol. Chem.* **270,** 28199–28203 (1995).
392. Debnath, D. K. & Otzen, D. E. Cell-free synthesis and folding of transmembrane OmpA reveals higher order structures and premature truncations. *Biophysical Chemistry* **152,** 80–88 (2010).
393. Goemans, C., Denoncin, K. & Collet, J.-F. Folding mechanisms of

periplasmic proteins. *BBA - Molecular Cell Research* 1–12 (2013). doi:10.1016/j.bbamcr.2013.10.014
394. Kelly, S. M., Jess, T. J. & Price, N. C. How to study proteins by circular dichroism. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics* **1751,** 119–139 (2005).
395. Bartlett, A. I. Kinetic analysis of protein folding energy landscapes. PhD thesis, University of Leeds (2010).
396. Kleinschmidt, J. H. & Tamm, L. K. Structural Transitions in Short-Chain Lipid Assemblies Studied by 31P-NMR Spectroscopy. *Biophysj* **83,** 994–1003 (2002).
397. Pautsch, A. & Schulz, G. E. Structure of the outer membrane protein A transmembrane domain. *Nat. Struct. Biol.* **5,** 1013–1017 (1998).
398. Vandeputte-Rutten, L., Kramer, R. A., Kroon, J., Dekker, N., Egmond, M. R. & Gros, P. Crystal structure of the outer membrane protease OmpT from Escherichia coli suggests a novel catalytic site. *EMBO J.* **20,** 5033–5039 (2001).
399. van den Berg, B., Black, P. N., Clemons, W. M. & Rapoport, T. A. Crystal structure of the long-chain fatty acid transporter FadL. *Science* **304,** 1506–1509 (2004).
400. Cowan, S. W., Schirmer, T., Rummel, G., Steiert, M., Ghosh, R., Pauptit, R. A., Jansonius, J. N. & Rosenbusch, J. P. Crystal structures explain functional properties of two E. coli porins. *Nature* **358,** 727–733 (1992).
401. Sugimura, K. & Nishihara, T. Purification, characterization, and primary structure of Escherichia coli protease VII with specificity for paired basic residues: identity of protease VII and OmpT. *J. Bacteriol.* (1988).
402. Duche, D., Issouf, M. & Lloubes, R. Immunity Protein Protects Colicin E2 from OmpT Protease. *Journal of Biochemistry* **145,** 95–101 (2008).
403. Hritonenko, V. & Stathopoulos, C. Omptin proteins: an expanding family of outer membrane proteases in Gram-negative Enterobacteriaceae(Review). *Molecular Membrane Biology* **24,** 395–406 (2009).
404. Kramer, R. A., Vandeputte-Rutten, L., de Roon, G. J., Gros, P., Dekker, N. & Egmond, M. R. Identification of essential acidic residues of outer membrane protease OmpT supports a novel active site. *FEBS Letters* **505,** 426–430 (2001).
405. Nunn, W. D. & Simons, R. W. Transport of long-chain fatty acids by Escherichia coli: mapping and characterization of mutants in the fadL gene. *Proceedings of the National Academy of Sciences* **75,** 3377–3381 (1978).
406. Black, P. N. The fadL gene product of Escherichia coli is an outer membrane protein required for uptake of long-chain fatty acids and involved in sensitivity to bacteriophage T2. *J. Bacteriol.* **170,** 2850–2854 (1988).
407. Black, P. N. Characterization of FadL-specific fatty acid binding in Escherichia coli. *Biochim. Biophys. Acta* **1046,** 97–105 (1990).
408. van den Berg, B. The FadL family: unusual transporters for unusual substrates. *Current Opinion in Structural Biology* **15,** 401–407 (2005).
409. Black, P. N. & DiRusso, C. C. Transmembrane Movement of Exogenous Long-Chain Fatty Acids: Proteins, Enzymes, and Vectorial Esterification. *Microbiol. Mol. Biol. Rev.* **67,** 454–472 (2003).
410. Arora, A., Abildgaard, F., Bushweller, J. H. & Tamm, L. K. Structure of outer membrane protein A transmembrane domain by NMR spectroscopy. *Nat. Struct. Biol.* **8,** 334–338 (2001).
411. Yamashita, E., Zhalnina, M. V., Zakharov, S. D., Sharma, O. & Cramer, W. A. Crystal structures of the OmpF porin: function in a colicin translocon. *EMBO J.* **27,** 2171–2180 (2008).
412. Seshadri, K., Garemyr, R., Wallin, E., Heijne, G. V. & Elofsson, A. Architecture of β-barrel membrane proteins: Analysis of trimeric porins.

Protein Science **7,** 2026–2032 (1998).
413. Masi, M. & Pagès, J.-M. Structure, Function and Regulation of Outer Membrane Proteins Involved in Drug Transport in Enterobactericeae: the OmpF/C - TolC Case. *Open Microbiol J* **7,** 22–33 (2013).
414. Housden, N. G., Hopper, J. T. S., Lukoyanova, N., Rodriguez-Larrea, D., Wojdyla, J. A., Klein, A., Kaminska, R., Bayley, H., Saibil, H. R., Robinson, C. V. & Kleanthous, C. Intrinsically disordered protein threads through the bacterial outer-membrane porin OmpF. *Science* **340,** 1570–1574 (2013).
415. Mitchell, D. C. Progress in understanding the role of lipids in membrane protein folding. *BBA - Biomembranes* **1818,** 951–956 (2012).
416. Pocanschi, C. L., Patel, G. J., Marsh, D. & Kleinschmidt, J. H. Curvature Elasticity and Refolding of OmpA in Large Unilamellar Vesicles. *Biophysj* **91,** L75–L77 (2006).
417. Volkmer, B. & Heinemann, M. Condition-Dependent Cell Volume and Concentration of Escherichia coli to Facilitate Data Conversion for Systems Biology Modeling. *PLoS ONE* **6,** e23126–6 (2011).
418. Tafer, H., Hiller, S., Hilty, C., Fernández, C. & Wüthrich, K. Nonrandom Structure in the Urea-Unfolded Escherichia coliOuter Membrane Protein X (OmpX) †. *Biochemistry* **43,** 860–869 (2004).
419. Kräutler, V., Hiller, S. & Hünenberger, P. H. Residual structure in a peptide fragment of the outer membrane protein X under denaturing conditions: a molecular dynamics study. *Eur Biophys J* **39,** 1421–1432 (2010).
420. Pattnaik, B. R., Ghosh, S. & Rajeswari, M. R. Selective excitation of tryptophans in OmpF: A fluorescence emission study. *IUBMB Life* **42,** 173–181 (1997).
421. O'Keeffe, A. H., East, J. M. & Lee, A. G. Selectivity in Lipid Binding to the Bacterial Outer Membrane Protein OmpF. *Biophysj* **79,** 2066–2074 (2000).
422. Rollauer, S. E., Sooreshjani, M. A., Noinaj, N. & Buchanan, S. K. Outer membrane protein biogenesis in Gram-negative bacteria. *Philos. Trans. R. Soc. Lond., B, Biol. Sci.* **370,** 20150023 (2015).
423. Brack, A. & Orgel, L. E. Beta structures of alternating polypeptides and their possible prebiotic significance. *Nature* **256,** 383–387 (1975).
424. Danoff, E. J. & Fleming, K. G. Aqueous, Unfolded OmpA Forms Amyloid-Like Fibrils upon Self-Association. *PLoS ONE* **10,** e0132301–10 (2015).
425. Knowles, T. P. J., Vendruscolo, M. & Dobson, C. M. The amyloid state and its association with protein misfolding diseases. *Nat. Rev. Mol. Cell Biol* **15,** 384–396 (2014).
426. Murakami, S., Nakashima, R., Yamashita, E. & Yamaguchi, A. Crystal structure of bacterial multidrug efflux transporter AcrB. *Nature* **419,** 587–593 (2002).
427. Danoff, E. J. & Fleming, K. G. The soluble, periplasmic domain of OmpA folds as an independent unit and displays chaperone activity by reducing the self-association propensity of the unfolded OmpA transmembrane β-barrel. *Biophysical Chemistry* **159,** 194–204 (2011).
428. Wülfing, C. & Plückthun, A. Protein folding in the periplasm of Escherichia coli. *Mol. Microbiol.* **12,** 685–692 (1994).
429. Hernández, H. & Robinson, C. V. Determining the stoichiometry and interactions of macromolecular assemblies from mass spectrometry. *Nat Protoc* **2,** 715–726 (2007).
430. Ruotolo, B. T. Evidence for Macromolecular Protein Rings in the Absence of Bulk Water. *Science* **310,** 1658–1661 (2005).
431. Hopper, J. T. S., Yu, Y. T.-C., Li, D., Raymond, A., Bostock, M., Liko, I., Mikhailov, V., Laganowsky, A., Benesch, J. L. P., Caffrey, M., Nietlispach, D. & Robinson, C. V. Detergent-free mass spectrometry of membrane protein complexes. *Nature Methods* **10,** 1206–1208 (2013).

432. Calabrese, A. N., Watkinson, T. G., Henderson, P. J. F., Radford, S. E. & Ashcroft, A. E. Amphipols Outperform Dodecylmaltoside Micelles in Stabilizing Membrane Protein Structure in the Gas Phase. *Anal. Chem.* **87,** 1118–1126 (2015).
433. Venter, H., Ashcroft, A. E., Keen, J. N., Henderson, P. J. F. & Herbert, R. B. Molecular dissection of membrane-transport proteins: mass spectrometry and sequence determination of the galactose-H+ symport protein, GalP, of Escherichia coli and quantitative assay of the incorporation of [ring-2-13C]histidine and (15)NH(3). *Biochem. J.* **363,** 243–252 (2002).
434. Jones, L. N., Baldwin, S. A., Henderson, P. J. F. & Ashcroft, A. E. Defining topological features of membrane proteins by nanoelectrospray ionisation mass spectrometry. *Rapid Commun. Mass Spectrom.* **24,** 276–284 (2010).
435. Weber, K. & Osborn, M. The Reliability of Molecular Weight Determinations by Dodecyl Sulfate-Polyacrylamide Gel Electrophoresis. *J. Biol. Chem.* **244,** 4406–4412 (1969).
436. Shelimov, K. B. & Jarrold, M. F. Conformations, unfolding, and refolding of apomyoglobin in vacuum: An activation barrier for gas-phase protein folding. *J. Am. Chem. Soc.* **119,** 2987–2994 (1997).
437. Kanu, A. B., Dwivedi, P., Tam, M., Matz, L. & Hill, H. H., Jr. Ion mobility-mass spectrometry. *J. Mass Spectrom.* **43,** 1–22 (2008).
438. Benesch, J. L. & Ruotolo, B. T. Mass spectrometry: come of age for structural and dynamical biology. *Current Opinion in Structural Biology* **21,** 641–649 (2011).
439. Ruotolo, B. T. & Robinson, C. V. Aspects of native proteins are retained in vacuum. *Curr Opin Chem Biol* **10,** 402–408 (2006).
440. Pacholarz, K. J., Porrini, M., Garlish, R. A., Burnley, R. J., Taylor, R. J., Henry, A. J. & Barran, P. E. Dynamics of Intact Immunoglobulin G Explored by Drift-Tube Ion-Mobility Mass Spectrometry and Molecular Modeling. *Angew. Chem. Int. Ed.* **53,** 7765–7769 (2014).
441. Zhou, M., Politis, A., Davies, R. B., Liko, I., Wu, K.-J., Stewart, A. G., Stock, D. & Robinson, C. V. Ion mobility-mass spectrometry of a rotary ATPase reveals ATP-induced reduction in conformational flexibility. *Nature Chemistry* **6,** 208–215 (2014).
442. Quan, S., Wang, L., Petrotchenko, E. V., Makepeace, K. A., Horowitz, S., Yang, J., Zhang, Y., Borchers, C. H. & Bardwell, J. C. Super Spy variants implicate flexibility in chaperone action. *eLife* **3,** e01584–22 (2014).
443. Stirling, P. C., Bakhoum, S. F., Feigl, A. B. & Leroux, M. R. Convergent evolution of clamp-like binding sites in diverse chaperones. *Nat Struct Mol Biol* **13,** 865–870 (2006).
444. Estrada Mallarino, L., Fan, E., Odermatt, M., Müller, M., Lin, M., Liang, J., Heinzelmann, M., Fritsche, F., Apell, H.-J. & Welte, W. TtOmp85, a β-Barrel Assembly Protein, Functions by Barrel Augmentation. *Biochemistry* **54,** 844–852 (2015).
445. Fleming, K. G. A combined kinetic push and thermodynamic pull as driving forces for outer membrane protein sorting and folding in bacteria. *Philos. Trans. R. Soc. Lond., B, Biol. Sci.* **370,** (2015).
446. Hartl, F. U., Bracher, A. & Hayer-Hartl, M. Molecular chaperones in protein folding and proteostasis. *Nature* **475,** 324–332 (2011).
447. Kim, Y. E., Hipp, M. S., Bracher, A., Hayer-Hartl, M. & Ulrich Hartl, F. Molecular Chaperone Functions in Protein Folding and Proteostasis. *Annu. Rev. Biochem.* **82,** 323–355 (2013).
448. Fleming, K. G. Energetics of Membrane Protein Folding. *Annu. Rev. Biophys.* **43,** 233–255 (2014).
449. Clantin, B., Delattre, A.-S., Rucktooa, P., Saint, N., Méli, A. C., Locht, C., Jacob-Dubuisson, F. & Villeret, V. Structure of the membrane protein FhaC:

a member of the Omp85-TpsB transporter superfamily. *Science* **317,** 957–961 (2007).
450. Chimento, D. P., Mohanty, A. K., Kadner, R. J. & Wiener, M. C. Substrate-induced transmembrane signaling in the cobalamin transporter BtuB. *Nat. Struct. Biol.* **10,** 394–401 (2003).
451. Wu, E. L., Fleming, P. J., Yeom, M. S., Widmalm, G., Klauda, J. B., Fleming, K. G. & Im, W. E. coli outer membrane and interactions with OmpLA. *Biophys. J.* **106,** 2493–2502 (2014).
452. Sinnige, T., Weingarth, M., Renault, M., Baker, L., Tommassen, J. & Baldus, M. Solid-State NMR Studies of Full-Length BamA in Lipid Bilayers Suggest Limited Overall POTRA Mobility. *J. Mol. Biol.* **426,** 2009–2021 (2014).
453. Paula, S., Volkov, A. G., Van Hoek, A. N., Haines, T. H. & Deamer, D. W. Permeation of protons, potassium ions, and small polar molecules through phospholipid bilayers as a function of membrane thickness. *Biophysj* **70,** 339–348 (1996).
454. Bennett, W. F. D., Sapay, N. & Tieleman, D. P. Atomistic Simulations of Pore Formation and Closure in Lipid Bilayers. *Biophysj* **106,** 210–219 (2014).
455. Nielsen, L. K., Bjørnholm, T. & Mouritsen, O. G. Fluctuations caught in the act. *Nature* **404,** 352–352 (2000).
456. Tokumasu, F., Jin, A. J. & Dvorak, J. A. Lipid membrane phase behaviour elucidated in real time by controlled environment atomic force microscopy. *J Electron Microsc (Tokyo)* **51,** 1–9 (2002).
457. Heinz, E. & Lithgow, T. A comprehensive analysis of the Omp85/TpsB protein superfamily structural diversity, taxonomic occurrence, and evolution. *Front Microbiol* **5,** 370 (2014).
458. Henke, K., Welte, W. & Hauser, K. Direct Monitoring of β-Sheet Formation in the Outer Membrane Protein TtoA Assisted by TtOmp85. *Biochemistry* **55,** 4333–4343 (2016).
459. Jerabek-Willemsen, M., Wienken, C. J., Braun, D., Baaske, P. & Duhr, S. Molecular Interaction Studies Using Microscale Thermophoresis. *ASSAY and Drug Development Technologies* **9,** 342–353 (2011).
460. Milev, S. Isothermal titration calorimetry: Principles and experimental design. *General Electric* (2013).
461. Duff, M. R., Grubbs, J. & Howell, E. E. Isothermal titration calorimetry for measuring macromolecule-ligand affinity. *J Vis Exp* (2011). doi:10.3791/2796
462. Ferguson, A. D., Welte, W., Hofmann, E., Lindner, B., Holst, O., Coulton, J. W. & Diederichs, K. A conserved structural motif for lipopolysaccharide recognition by procaryotic and eucaryotic proteins. *Structure/Folding and Design* **8,** 585–592 (2000).
463. Lewis, B. A. & Engelman, D. M. Lipid bilayer thickness varies linearly with acyl chain length in fluid phosphatidylcholine vesicles. *J. Mol. Biol.* **166,** 211–217 (1983).
464. Lomize, A. L., Pogozheva, I. D., Lomize, M. A. & Mosberg, H. I. Positioning of proteins in membranes: A computational approach. *Protein Sci.* **15,** 1318–1333 (2006).
465. Killian, J. A. Hydrophobic mismatch between proteins and lipids in membranes. *Biochim. Biophys. Acta* **1376,** 401–415 (1998).
466. Morgado, L., Zeth, K., Burmann, B. M., Maier, T. & Hiller, S. Characterization of the insertase BamA in three different membrane mimetics by solution NMR spectroscopy. *J Biomol NMR* **61,** 333–345 (2015).
467. Sinnige, T., Houben, K., Pritisanac, I., Renault, M., Boelens, R. & Baldus,

M. Insight into the conformational stability of membrane-embedded BamA using a combined solution and solid-state NMR approach. *J Biomol NMR* **61,** 321–332 (2015).
468. Tellez, R. & Misra, R. Substitutions in the BamA β-barrel domain overcome the conditional lethal phenotype of a ΔbamB ΔbamE strain of Escherichia coli. *J. Bacteriol.* **194,** 317–324 (2012).
469. Rigel, N. W., Ricci, D. P. & Silhavy, T. J. Conformation-specific labeling of BamA and suppressor analysis suggest a cyclic mechanism for β-barrel assembly in Escherichia coli. *Proc. Natl. Acad. Sci. U.S.A.* **110,** 5151–5156 (2013).
470. Raivio, T. L. Everything old is new again: an update on current research on the Cpx envelope stress response. *Biochim. Biophys. Acta* **1843,** 1529–1541 (2014).
471. Ulrich, T. & Rapaport, D. Biogenesis of beta-barrel proteins in evolutionary context. *Int. J. Med. Microbiol.* **305,** 259–264 (2015).
472. Maurya, S. R., Chaturvedi, D. & Mahalakshmi, R. Modulating lipid dynamics and membrane fluidity to drive rapid folding of a transmembrane barrel. *Sci. Rep.* **3,** 1–6 (2013).
473. Anfinsen, C. B. Principles that govern the folding of protein chains. *Science* **181,** 223–230 (1973).
474. Ellis, R. J. & Hartl, F. U. Principles of protein folding in the cellular environment. *Current Opinion in Structural Biology* **9,** 102–110 (1999).
475. Duke, T. A. J., Le Novère, N. & Bray, D. Conformational spread in a ring of proteins: a stochastic approach to allostery. *J. Mol. Biol.* **308,** 541–553 (2001).
476. Goodsell, D. S. & Olson, A. J. Structural symmetry and protein function. *Annu Rev Biophys Biomol Struct* **29,** 105–153 (2000).
477. Baker, T. A. & Sauer, R. T. ClpXP, an ATP-powered unfolding and protein-degradation machine. *BBA - Molecular Cell Research* **1823,** 15–28 (2012).
478. Lehr, U., Schutz, M., Oberhettinger, P., Ruiz-Perez, F., Donald, J. W., Palmer, T., Linke, D., Henderson, I. R. & Autenrieth, I. B. C-terminal amino acid residues of the trimeric autotransporter adhesin YadA of Yersinia enterocolitica are decisive for its recognition and assembly by BamA. *Mol. Microbiol.* **78,** 932–946 (2010).
479. Lee, J., Xue, M., Wzorek, J. S., Wu, T., Grabowicz, M., Gronenberg, L. S., Sutterlin, H. A., Davis, R. M., Ruiz, N., Silhavy, T. J. & Kahne, D. E. Characterization of a stalled complex on the β-barrel assembly machine. *Proc. Natl. Acad. Sci. U.S.A.* **113,** 8717–8722 (2016).
480. Murzin, A. G., Lesk, A. M. & Chothia, C. Principles determining the structure of β-sheet barrels in proteins II. The observed structures. *J. Mol. Biol.* **236,** 1382–1400 (1994).
481. Schulz, G. E. The structure of bacterial outer membrane proteins. *Biochim. Biophys. Acta* **1565,** 308–317 (2002).
482. Eichner, T. & Radford, S. E. A Diversity of Assembly Mechanisms of a Generic Amyloid Fold. *Molecular Cell* **43,** 8–18 (2011).
483. Delattre, A.-S., Clantin, B., Saint, N., Locht, C., Villeret, V. & Jacob-Dubuisson, F. Functional importance of a conserved sequence motif in FhaC, a prototypic member of the TpsB/Omp85 superfamily. *FEBS Journal* **277,** 4755–4765 (2010).
484. Leonard-Rivera, M. & Misra, R. Conserved Residues of the Putative L6 Loop of Escherichia coli BamA Play a Critical Role in the Assembly of -Barrel Outer Membrane Proteins, Including That of BamA Itself. *J. Bacteriol.* **194,** 4662–4668 (2012).
485. Remaut, H., Rose, R. J., Hannan, T. J., Hultgren, S. J., Radford, S. E., Ashcroft, A. E. & Waksman, G. Donor-strand exchange in chaperone-

assisted pilus assembly proceeds through a concerted beta strand displacement mechanism. *Molecular Cell* **22,** 831–842 (2006).
486. Stansfeld, P. J., Goose, J. E., Caffrey, M., Carpenter, E. P., Parker, J. L., Newstead, S. & Sansom, M. S. P. MemProtMD: Automated Insertion of Membrane Protein Structures into Explicit Lipid Membranes. *Structure* **23,** 1350–1361 (2015).
487. Clark, P. L. & Elcock, A. H. Molecular chaperones: providing a safe place to weather a midlife protein-folding crisis. *Nat Struct Mol Biol* **23,** 621–623 (2016).
488. Serra-Batiste, M., Ninot-Pedrosa, M., Bayoumi, M., Gairí, M., Maglia, G. & Carulla, N. Aβ42 assembles into specific β-barrel pore-forming oligomers in membrane-mimicking environments. *Proc. Natl. Acad. Sci. U.S.A.* **113,** 10866–10871 (2016).
489. Horne, J. E. & Radford, S. E. A growing toolbox of techniques for studying β-barrel outer membrane protein folding and biogenesis. *Biochem. Soc. Trans.* **44,** 802–809 (2016).
490. Leitner, A., Walzthoeni, T., Kahraman, A., Herzog, F., Rinner, O., Beck, M. & Aebersold, R. Probing native protein structures by chemical cross-linking, mass spectrometry, and bioinformatics. *Mol. Cell Proteomics* **9,** 1634–1649 (2010).
491. Rappsilber, J. The beginning of a beautiful friendship: Cross-linking/mass spectrometry and modelling of proteins and multi-protein complexes. *Journal of Structural Biology* **173,** 530–540 (2011).
492. Baud, C., Guérin, J., Petit, E., Lesne, E., Dupré, E., Locht, C. & Jacob-Dubuisson, F. Translocation path of a substrate protein through its Omp85 transporter. *Nat Commun* **5,** 5271 (2014).
493. Calabrese, A. N., Ault, J. R., Radford, S. E. & Ashcroft, A. E. Using hydroxyl radical footprinting to explore the free energy landscape of protein folding. *Methods* **89,** 38–44 (2015).
494. Geibel, S., Procko, E., Hultgren, S. J., Baker, D. & Waksman, G. Structural and energetic basis of folded-protein transport by the FimD usher. *Nature* **496,** 243–246 (2014).
495. Gault, J., Donlan, J. A. C., Liko, I., Hopper, J. T. S., Gupta, K., Housden, N. G., Struwe, W. B., Marty, M. T., Mize, T., Bechara, C., Zhu, Y., Wu, B., Kleanthous, C., Belov, M., Damoc, E., Makarov, A. & Robinson, C. V. High-resolution mass spectrometry of small molecules bound to membrane proteins. *Nature Methods* **13,** 333–336 (2016).
496. Hayat, S., Sander, C., Marks, D. S. & Elofsson, A. All-atom 3D structure prediction of transmembrane β-barrel proteins from sequences. *Proceedings of the National Academy of Sciences* **112,** 5413–5418 (2015).
497. Mori, N., Ishii, Y., Tateda, K., Kimura, S., Kouyama, Y., Inoko, H., Mitsunaga, S., Yamaguchi, K. & Yoshihara, E. A peptide based on homologous sequences of the β-barrel assembly machinery component BamD potentiates antibiotic susceptibility of Pseudomonas aeruginosa. *J. Antimicrob. Chemother.* **67,** 2173–2181 (2012).
498. Ruiz, N., Wu, T., Kahne, D. & Silhavy, T. J. Probing the Barrier Function of the Outer Membrane with Chemical Conditionality. *ACS Chem. Biol.* **1,** 385–395 (2006).
499. Cain, R., Narramore, S., McPhillie, M., Simmons, K. & Fishwick, C. W. G. Applications of structure-based design to antibacterial drug discovery. *Bioorg. Chem.* **55,** 69–76 (2014).