UNIVERSITY OF LEEDS

Mechanisms of Outer Membrane Protein Folding: Effects of the Lipid Environment and Periplasmic Chaperones

Lindsay Madeline McMorran

Submitted in accordance with the requirements for

the degree of Doctor of Philosophy

The University of Leeds

Astbury Centre for Structural Molecular Biology

September 2013

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.

Details of jointly authored publications and the contributions of other authors to these manuscripts:

Chapter 3 contains work from the following publication:

Leney A. C., McMorran L. M., Radford S. E., Ashcroft A. E. Amphipathic Polymers Enable the Study of Functional Membrane Proteins in the Gas Phase. *Anal. Chem.* (2012) **22**: 9841–9847. © 2012 American Chemical Society.

In this publication, I provided the plasmids and expression/purification protocols for the production of HT PagP and OmpT. I was also involved in the design of all biochemical and biophysical assays, data interpretation and manuscript preparation. A. C. Leney carried out experimental work, data interpretation and manuscript preparation. S. E. Radford and A. E. Ashcroft provided help with scientific discussions, data interpretation and manuscript preparation.

Chapter 5 contains work from the following publication:

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.* (2013) **425:** 3178–3191

In this publication, I designed and performed experimental work, data interpretation, literature research and manuscript preparation. A. I. Bartlett also designed and performed experimental work, data interpretation, literature research and manuscript preparation. G. H. M. Huysmans provided scientific discussions and helped with manuscript preparation. S. E. Radford and D. J. Brockwell provided help with scientific discussions, data interpretation and manuscript preparation.

ACKNOWLEDGEMENTS

I could probably write a second thesis on the amazing people I've worked with and how they've helped me, but in case anyone actually reads this I shall endeavour not to. Firstly, I would like to express my gratitude to my supervisors Sheena Radford and David Brockwell for their dedication, enthusiasm and guidance throughout my Ph.D. David taught me all of the molecular biology techniques used in this thesis but, while I am very grateful, I maintain MolBiol is actually voodoo. Thank you to Sheena for expanding my vocabulary with useful phrases such as "up the wazoo" and "oodles" as well as giving plenty of pep talks when things were not going according to plan.

My "OMP friend", Alice Bartlett, taught me lots about OMPs and how to avoid killing myself in the lab. Thank you so much for all your time, advice and friendship over the last four years, you're a star! Thank you also for introducing me to Gerard Huysmans, whose inexhaustible knowledge on all things membrane-related (coupled with a healthy dose of sarcasm) has proved invaluable throughout this project. Thanks are also due to Remco Rodenburg for forcing me to learn to explain myself better and to the newest addition to the team, Bob Schiffrin, for so many useful chats and humouring my wild conjecture. I'm also grateful to Alison Ashcroft and "Team OMP – Mass Spectacular division" for their collaboration. Thanks to Bethny Morrissey, who initiated the OMP mass spec work, as well as being the host of awesome parties and an all-round amazeface, party-arms friend; Aneika Leney, who cleverly steered us on to amphipols and forced me to go climbing; and Tom Watkinson for attempting to spread early morning enthusiasm.

Thanks to my proof reading minions for finding my typos and grammatical failings: Jan Saunders, Rhys Thomas, Alice Bartlett, Claire Sarell, Clare Pashley and Bob Schiffrin.

Thank you to all the other members of the Radford and Brockwell labs, in particular: Gareth Morgan for experimental baked goods, Clare Pashley for keeping us all entertained (intentionally or otherwise), Alessandro Sicorello for providing me with TEV protease, Theo Karamanos for helping to make the OmpP homology model, Sophie Goodchild for teaching me about FRET and lots of useful membrane-related things, Wei-Feng Xue for sharing his personal chocolate supply and knowledge on error

analysis, Paul Devine and Charlie Scarff for dusting off their sock puppets and explaining various aspects of mass spectrometry literature to me, Chris Wilson for helping me to decipher DJB's writing and Claire Sarell for referring to me as a "marginally less offensive but more attractive version of Frankie Boyle". Uh… thanks? A special mention goes to Jan Saunders, even if she did keep stealing my pens and once set my lab coat on fire, thanks for keeping me laughing both in and out of the lab! Keith Ainley and Nasir Khan have provided excellent technical support and a truckload of LB over the last four years, thank you both for your hard work in keeping the lab running smoothly. Thanks also to Nasir for keeping me well fed with amazing home-made curries and lots of biscuits!

I'm grateful to my "BAM buddy" Dennis Gessmann for our many OMP-related discussions, sending paper links to things I might have missed and for making liposome preps more entertaining by introducing me to the phrase "making LUV" too.

I'd also like to give a quick mention to Reviewer 1 – I'm glad you found my lysozyme experiments "particularly amusing" and thanks for qualifying me for the title of "most entertaining reviewer's comment" in the Radford/Brockwell labs.

I would like to thank all my family and friends, especially my parents for their unconditional love and support in everything I do. When I was younger, I once called my brother stupid and he replied that I could call him that when I was more qualified than him – Leigh, you're stupid but thank you for accidentally providing the inspiration to keep learning. Thanks to Donnie McAusland for his support and trying to find ways to cheer me up when I needed it. Last, but definitely not least, thank you to my thesis-writing cheerleader and favourite ginge, Pete Westwood, for all his encouragement and attempts to make thesis-writing less depressing (including but not limited to bad movies, meerkats and baby murlocs).

During the last four years I met a lot of awesome people but also had to say goodbye to a few other others. I realise I'm well on the way to being a crazy cat lady, but thanks to Emily "Ba" the cat for being an awesome furball and friend to me for 12 years. Finally, I would like to dedicate this thesis to Ryan, my beautiful nephew, and my grandparents who gave me so much support and encouragement while always reminding me not to underestimate the importance of common sense. If only I had some!

ABSTRACT

In contrast with the wealth of information on the folding of soluble, cytosolic proteins, little is known about the folding of integral membrane proteins. The outer membrane proteins (OMPs) of Gram-negative bacteria have a β-barrel structure and are essential for cell survival. The mechanisms of OMP transport across the periplasm and how these proteins subsequently fold and insert into the outer membrane remain to be elucidated.

The work presented herein examines the folding and membrane insertion of four different OMP constructs. Two homologous bacterial OMPs, OmpT and OmpP, were cloned, over-expressed and purified before biochemical and biophysical methods were employed to examine their folding properties. This work demonstrates that small differences in primary sequence can have large effects on folding efficiency and stability.

In spite of both OmpT and OmpP being able to fold under a variety of conditions, it was not possible to establish conditions under which folding was completely reversible. Examination of the origins of irreversible OMP folding was carried out using OmpT, as well as both hexa-histidine tagged and untagged constructs of the outer membrane acyltransferase enzyme, PagP. This study revealed evidence that lipid adhesion of the protein in the unfolded state may be important in preventing aggregation and promoting reversibility.

Finally, conditions were established to promote the folding of untagged PagP in low urea concentrations to allow the study of OMPs in the presence of the periplasmic chaperones, SurA and Skp. SurA was shown to have little effect on the folding of PagP into liposomes with zwitterionic or negatively charged membrane surfaces, while Skp was shown to exhibit holdase activity and to modulate PagP folding rate, dependent on the lipid composition. The results present the first detailed insights into the mechanism by which Skp and SurA act to facilitate PagP folding *in vitro*.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

Amino acids are abbreviated according to their standard three-letter or single-letter codes. Other abbreviations are as follows:

Abz	*ortho*-aminobenzoic acid
AFM	Atomic force microscopy
ANS	8-Anilinonaphthalene-1-sulfonic acid
Apol	Amphipol or amphipathic polymer
ATP	Adenosine triphosphate
BAM	β-barrel assembly machine
BMP	bis(monoacylglycero)phosphate
bO	Bacterioopsin
bR	Bacteriorhodopsin
BSA	Bovine serum albumin
CD	Circular dichroism
CHAPS	3-[(3-cholamidopropyl)dimethylammonio]-1-propanesulfonate
CL	Cardiolipin
di$C_{x:y}$PZ	A phospholipid with two identical acyl chains: x indicates the number of carbon atoms in the acyl chain y indicates the number of double bonds per acyl chain Z represents the head-group moiety; choline (C), ethanolamine (E), glycerol (G), inositol (I) or serine (S)
Dansyl	5-(dimethylamino)naphthalene-1-sulfonyl
DDM	*n*-dodecyl-β-D-maltopyranoside
dH_2O	Deionised water (greater than 15 MΩ resistance)
DNA	Deoxyribonucleic acid
dNTP	deoxyribonucleotide triphosphate
dsDNA	Double stranded deoxyribonucleic acid
E. coli	*Escherichia coli*
EDTA	Ethylenediaminetetraacetic acid
En-HD	Engrailed homeodomain
ESI	Electrospray ionisation

EtOH	Ethanol
fnIII	Fibronectin type III
FRET	Förster resonance energy transfer
GPI	Glycophosphatidylinositol
GTP	Guanosine triphosphate
GuHCl	Guanidine hydrochloride
HT	His-tagged
hVDAC	Human voltage-dependent anion-selective channel
Ig	Immunoglobulin
IM	Inner membrane
IMS	Ion mobility spectrometry
IPTG	Isopropyl-β-D-thiogalactoside
LB	Lysogeny broth
LDAO	*N*-lauryl-*N*,*N*-dimethylammonium-*N*-oxide
LPR	Lipid to protein ratio
LPS	Lipopolysaccharide
LUV	Large unilamellar vesicle
MALDI	Matrix-assisted laser desorption ionisation
MD	Molecular dynamics
MS	Mass spectrometry
MWCO	Molecular weight cut-off
NaCl	Sodium chloride
NBF	Nucleotide binding fold
N. meningitides	*Neisseria meningitides*
NMR	Nuclear magnetic resonance
OM	Outer membrane
OMP	Outer membrane protein
PA	Phosphatidic acid
PAGE	Polyacrylamide gel electrophoresis
PBS	Phosphate-buffered saline
PC	Phosphocholine
PCR	Polymerase chain reaction
PDB	Protein Data Bank
PDZ	PSD-95, Discs-large, ZO-1

PE	Phosphoethanolamine
PG	Phosphoglycerol
pI	Isoelectric point
PI	Phosphoinositol
PMSF	Phenylmethanesulfonylfluoride
POTRA	Polypeptide-transport-associated
PPIase	Peptidyl-prolyl isomerise
PS	Phosphoserine
PVDF	Polyvinylidene fluoride
RMSD	Root mean squared deviation
SDS	Sodium dodecyl sulphate
SDS-PAGE	Sodium dodecyl sulphate polyacrylamide gel electrophoresis
SEM	Standard error of the mean
SM	Sphingomyelin
SOB	Super optimal broth
SOC	Super optimal broth with catabolite repression
SRP	Signal recognition particle
SUV	Small unilamellar vesicle
TAE	Tris-acetate-EDTA
TBS	Tris-buffered saline
TCA	Trichloroacetic acid
TEMED	Tetramethylethylenediamine
TEVp	Tobacco etch virus protease
TPR	Tetratricopeptide repeat
Tris	Tris (hydroxymethyl)-aminomethane
UV	Ultra violet
X-Gal	5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside

1 INTRODUCTION

1.1 PRINCIPLES OF PROTEIN FOLDING

1.1.1 HOW WATER-SOLUBLE PROTEINS FOLD

The biological activity of many proteins is reliant on their ability to adopt a specific, three-dimensional structure. Failure to achieve this functional, or native, state can have serious consequences, as evidenced by the prevalence of diseases caused by protein misfolding and aggregation[1; 2; 3]. The information required for a polypeptide chain to assume its native structure was shown to be contained within its primary sequence of amino acids by Anfinsen's experiments on ribonuclease A[4; 5]. Ribonuclease A was demonstrated to regain its native conformation when dialysis was carried out following complete denaturation of this small, globular protein with urea and reduction of the disulphide bonds using dithiothreitol[4]. This seemingly simple conclusion has led to decades of scientific research on what remains one of the key questions in the study of protein folding: how does the amino acid sequence of a protein confer its native structure[6]?

If folding were to occur by random sampling of all possible conformations of the polypeptide chain, finding the native state would take an astronomically long time, however, proteins are able to fold on biologically relevant timescales[7]. Levinthal suggested that this apparent paradox could be resolved if proteins fold *via* defined pathways[8]. Following this conclusion, several mechanisms were proposed to describe the pathways traversed by a folding protein *en route* to the native state. Analysis of the refolding kinetics of ribonuclease A revealed two distinct phases[9] and led to the suggestion of a "framework" mechanism of folding wherein the secondary structure is formed prior to docking together to form the tertiary structure[6; 10]. Further investigation of the refolding of ribonuclease A revealed that one of the phases observed was not, in fact, due to the presence of an observable folding intermediate, but to proline *cis-trans* isomerisation[11]. This realisation, alongside the characterisation of the folding of chymotrypsin inhibitor 2 which revealed a simple two-state folding system[12], saw the framework mechanism become disfavoured due to its implication that folding intermediates should be present[6; 10]. To explain folding in the absence of detectable

intermediates, the nucleation condensation mechanism was proposed. This mechanism involves the formation of a small nucleus of structure stabilised by weak, long-range contacts and the subsequent rapid collapse around this folding nucleus to yield the native state[6; 10]. A combination of simulation and experimental data on members of the homeodomain-like super-family of proteins revealed that these two seemingly contrasting folding mechanisms could be considered as two extremes of a single folding mechanism, with the relative stability of the secondary and tertiary structure determining if these elements are formed in series or in parallel[13].

More recently it has been realised that proteins fold *via* a collection of parallel pathways which make up a funnel shaped energy landscape[14]. At the top of the funnel, the unfolded state represents a large ensemble of high-entropy conformations of the polypeptide chain, which may be biased by weak, residual interactions[15]. Indeed, an unfolded variant of the bacterial immunity protein Im7 has been studied recently under non-denaturing conditions, revealing conformational restriction in the regions of the protein which ultimately form the native helices and demonstrating the importance of such interactions in the initiation of folding[16]. By undergoing many reorganisations, the polypeptide chain forms stabilising interactions between side chains, the protein backbone and the solvent as the protein approaches the native state[14]. The landscape view is an attractive one as it does not place restrictions on whether secondary structure must form before, or at the same time, as the tertiary structure[14]. Additionally, the funnel-shaped landscape predicts the robustness of the folding process to destabilising mutations: so long as the final fold remains the most stable state relative to the unfolded ensemble, a mutation may block some of the pathways to the native state but alternative folding pathways can be utilised[14].

For proteins in which only the native and unfolded states are populated, known as a two-state folding mechanism, the energy landscape to be traversed by the folding polypeptide is relatively smooth (Figure 1.1A)[17]. This simple folding scenario is rarely observed even for small proteins, however, and the folding landscape is often rugged with local energy minima representing the population of one or more folding intermediates *en route* to the native state (Figure 1.1B)[17; 18].

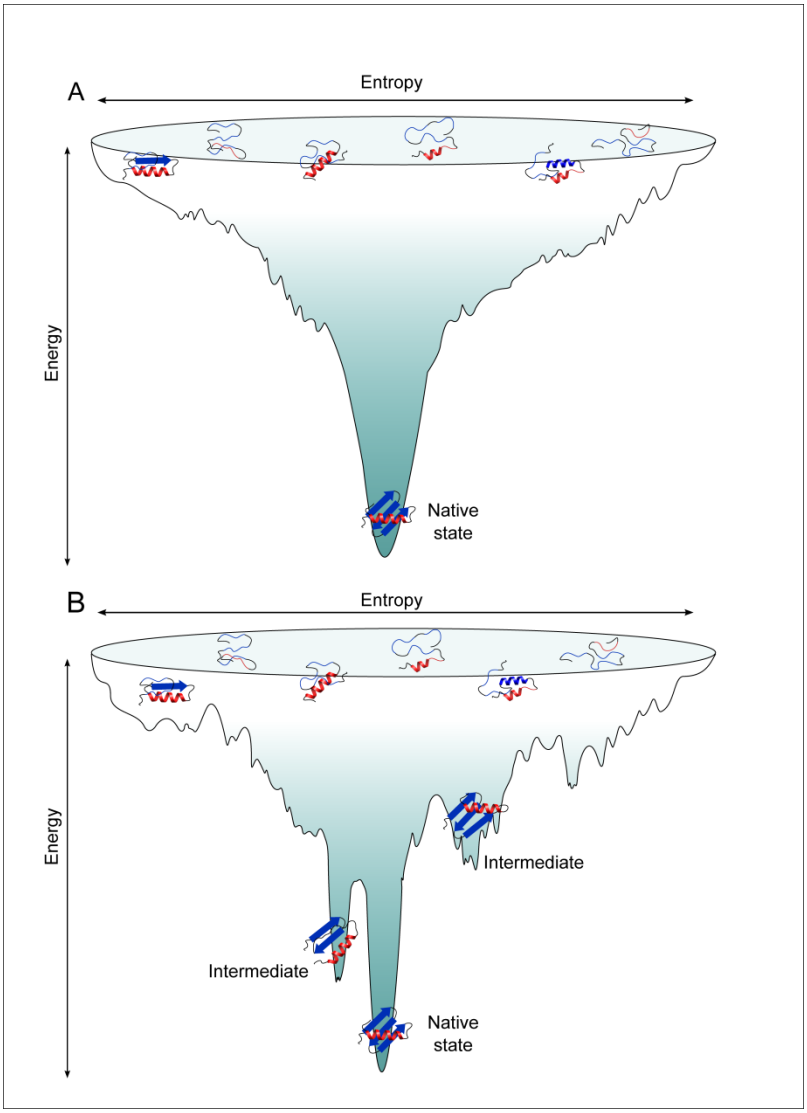

Figure 1.1 A schematic representation of funnel shaped folding landscapes. Examples of (A) a smooth folding landscape expected for a two-state folding mechanism where only the native and unfolded states are populated and (B) a rugged landscape wherein the polypeptide chain populates one or more intermediate structures which represent local energy minima. This figure was taken from Bartlett and Radford, 2009[17].

Rugged folding landscapes are a consequence of the need to form the thousands of weak, stabilising interactions which are present in the native state[17]. During the structural reorganisations required for the protein to traverse the folding pathway, it is not always possible to minimise the energy of all of these interactions simultaneously, leading to "frustration" in the landscape[19; 20]. Theory predicts that a rough landscape will lead to slower folding by limiting the rate at which the protein can find the native state[19]. Recently the effects of solvent viscosity on the folding kinetics of homologous domains of α-spectrin were studied, revealing the first explicit experimental evidence

that differences in landscape ruggedness can indeed influence folding kinetics of structurally homologous proteins[21]. It has been suggested previously that frustration, and hence ruggedness in energy landscapes, arises due to conflicting pressures on proteins to evolve sequences which can reliably fold to a stable native structure that is also capable of carrying out a defined biological role[19; 20]. While this may be true in some cases, the differences in landscape ruggedness of structurally and functionally similar spectrin domains suggest that increased frustration in some domains may act to deliberately slow unfolding in this protein which has an unusually long half-life *in vivo*[21]. Furthermore, folding of the Trp-cage mini protein was shown to be slowed when the folding intermediate was destabilised by the presence of chemical denaturant or a helix-breaking mutation[22] suggesting that in this case local energetic minima in the folding landscape act to separate the conformational search into multiple, smaller problems (reviewed in Brockwell and Radford, 2007[18]). Whether folding intermediates, in general, act to disrupt or promote the folding process is still under debate. The utility of intermediate species in providing insights into the structural regulation of biological function, however, as well as their role in initiation of protein aggregation, makes them valuable targets of study[23].

The landscape theory of protein folding also predicts the existence of "downhill folding", that is, folding in the absence of significant energy barriers[14]. In such a folding scenario, non-exponential folding kinetics would be observed due to the lack of a bottleneck in the pathway and the overall folding time would be expected to approach the folding "speed limit" predicted by theory[14; 24]. Downhill folding would, in principle, allow the direct study of all the intermediate conformations adopted by the polypeptide chain *en route* to the native state, providing a wealth of information about these transiently populated species[17]. Even with the most powerful single molecule techniques available, however, unequivocal identification of barrier-less folding is still difficult experimentally[25]. Analysis of the folding rates of known "fast-folding" proteins suggested that even these extreme examples could still be engineered to fold up to ten times faster[24], in contrast with the expected results for a truly downhill folding system[14]. Use of simulation methods to probe the folding of such fast-folding proteins suggested that although, in some cases, no barrier could be identified, many do experience energy barriers to folding[26]. While it is clear that many fundamental questions remain to be answered about the folding of globular proteins, recent advances

in experimental and computational techniques have provided significant progress towards understanding this complex and important biological process.

1.1.2 The Membrane Protein Folding Problem

In contrast with the wealth of information available on the folding of water-soluble proteins, progress in the field of membrane protein folding has lagged significantly[27]. In water-soluble proteins, folding is driven by the need to bury hydrophobic side chains, in order to prevent aggregation, and facilitate formation of the most stable structure[28]. Conversely, the folding of integral membrane proteins is complicated by the need to insert the polypeptide chain into a lipid membrane, as well as to fold to the native state[27]. Following insertion, most of the surface of the protein is in contact with the hydrophobic membrane interior, with hydrophilic residues restricted to regions which contact the polar head-groups of the membrane lipids or being exposed to the aqueous environment on either side of the membrane[27]. Additionally, the membrane environment *in vivo* is highly dynamic and heterogeneous with regions of varying lipid composition[29]. Recreating this environment *in vitro* has proved more difficult than the simple aqueous environment needed to fold water-soluble proteins, and this has greatly limited studies on the folding of integral membrane proteins[27; 29]. Recently, however, more research is being focused on understanding the folding, stability and function of integral membrane proteins as these proteins represent 60 % of current drug targets and have many roles in the cell, including signalling, transport and biogenesis[30; 31]. In spite of recent advances in this field, understanding of membrane protein folding and membrane protein structure databases still lag behind those of water-soluble proteins[29]. This illustrates the need for research into the underlying principles which govern the folding of membrane embedded proteins.

1.2 Classes of Membrane Proteins

The proteins present in biological membranes can be categorised by two distinct families: the lipid-anchored proteins, which have a covalently attached fatty-acid moiety to associate them with the membrane, and the integral membrane proteins, which

contain membrane-spanning regions. Bacterial lipid-anchored proteins, or lipoproteins, are anchored to the membrane *via* an N-terminal N-acyl-diacylglycerylcysteine[32] moiety. The type and location of these linkers vary in eukaryotic organisms, with examples including prenylation, found in the Rab family of monomeric G-proteins[33], and glycophosphatidylinositol (GPI) linkers[34].

In contrast with the myriad of lipid linkages available to lipid-associated proteins, the integral membrane proteins are constrained by the need to compensate for the energetic cost of burying a peptide bond in the lipid bilayer[35], estimated to be 1.2 kcal/mol[36]. As a consequence of this, it was predicted that membrane spanning regions would form regular secondary structural elements in order to maximise the hydrogen bonding potential of the peptide backbone[35]. Formation of secondary structure reduces the energetic cost of incorporation of peptide bonds into a bilayer by 0.4 kcal/mol per peptide bond for α-helical structure and 0.5 kcal/mol per peptide bond for β-sheet structure[36]. The first α-helical membrane protein structure was solved in 1975 by Henderson and Unwin using electron microscopy to generate a three-dimensional image of the purple membrane of *Halobacterium salinarum*[37]. The resulting 7 Å resolution image revealed the structure of bacteriorhodopsin (bR) to be a transmembrane seven helical bundle[37]. The structural resolution has since been increased using X-ray crystallography and the most detailed structural information on bR available in the Protein Data Bank (PDB) is currently at a resolution of 1.43 Å[38]. Since the structure of bR was elucidated, a wide variety of helical transmembrane structures have been solved and deposited in the Membrane Protein Data Bank[39] showing a diverse range of sizes and functions across all organisms. Some examples are depicted in Figure 1.2.

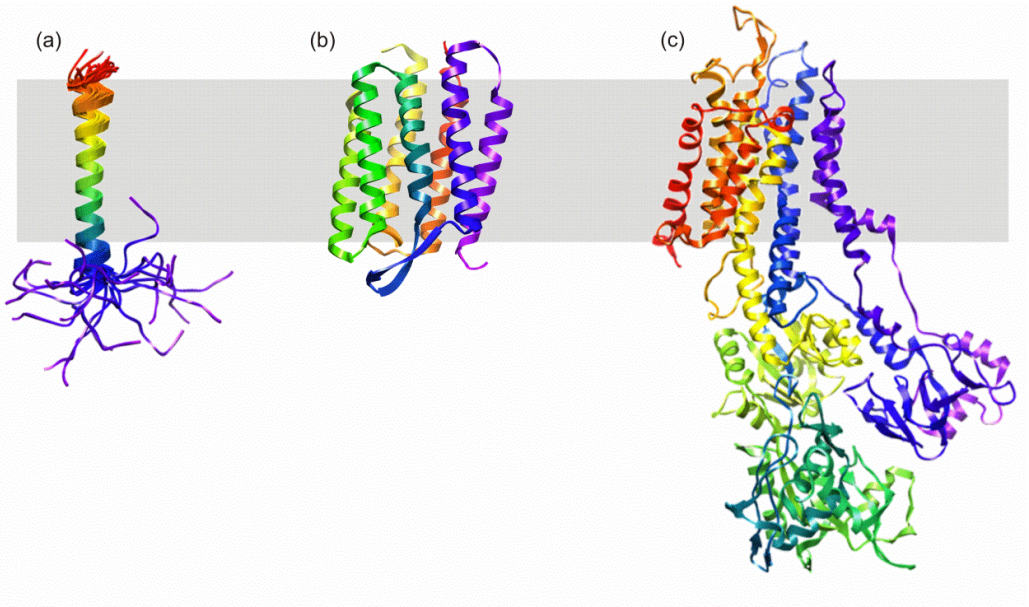

Figure 1.2 Example structures of α-helical integral membrane proteins. Structures of (a) the transmembrane segment of a glycophorin A monomer from human erythrocyte membranes solved by NMR spectroscopy (1AFO[40]); (b) bacteriorhodopsin, a seven-helical bundle from the purple membrane of *Halobacterium salinarum* solved by X-ray crystallography (1C3W[41]) and (c) calcium ATPase 1 from the sarcoplasmic reticulum membrane of *Oryctolagus cuniculus*, a ten-helical bundle with large cytoplasmic domain solved by X-ray crystallography (1IWO[42]). Proteins are coloured rainbow: violet (N-terminus) to red (C-terminus). The approximate position of the membrane is indicated with grey shading with the cytoplasmic side of the membrane being represented by the bottom edge of the shaded area. All images were generated from the Protein Data Bank[43] files from the accession numbers given in brackets using UCSF Chimera molecular visualisation application[44].

In contrast with the ubiquitous distribution of α-helical transmembrane proteins, membrane proteins containing β-sheet secondary structure are found only in the outer membrane (OM) of Gram-negative bacteria, mitochondria and chloroplasts[45; 46]. In order to satisfy all of the possible hydrogen bonds in the peptide backbone, each β-strand associates laterally with its neighbours resulting in an overall cylindrical topology, known as a β-barrel[45]. With predominantly non-polar side chains exposed to the hydrophobic membrane interior and each of the backbone hydrogen bonds satisfied, the resulting structure has a very high thermodynamic stability[36; 47]. The size of the β-barrel is highly variable, with known structures containing between 8 and 24 β-strands, and the protein may contain periplasmic or extracellular domains[45; 46]. The regions between β-strands often alternate between tight turns on the periplasmic side of the membrane and longer, more flexible loops on the outer surface, which are exposed to

the external environment[45]. One almost entirely conserved structural feature is an even number of β-strands, with only one known exception to this rule discovered thus far: the human mitochondrial voltage-dependent anion-selective channel (hVDAC)[46; 48]. Some examples of β-barrel membrane protein structures are depicted in Figure 1.3.

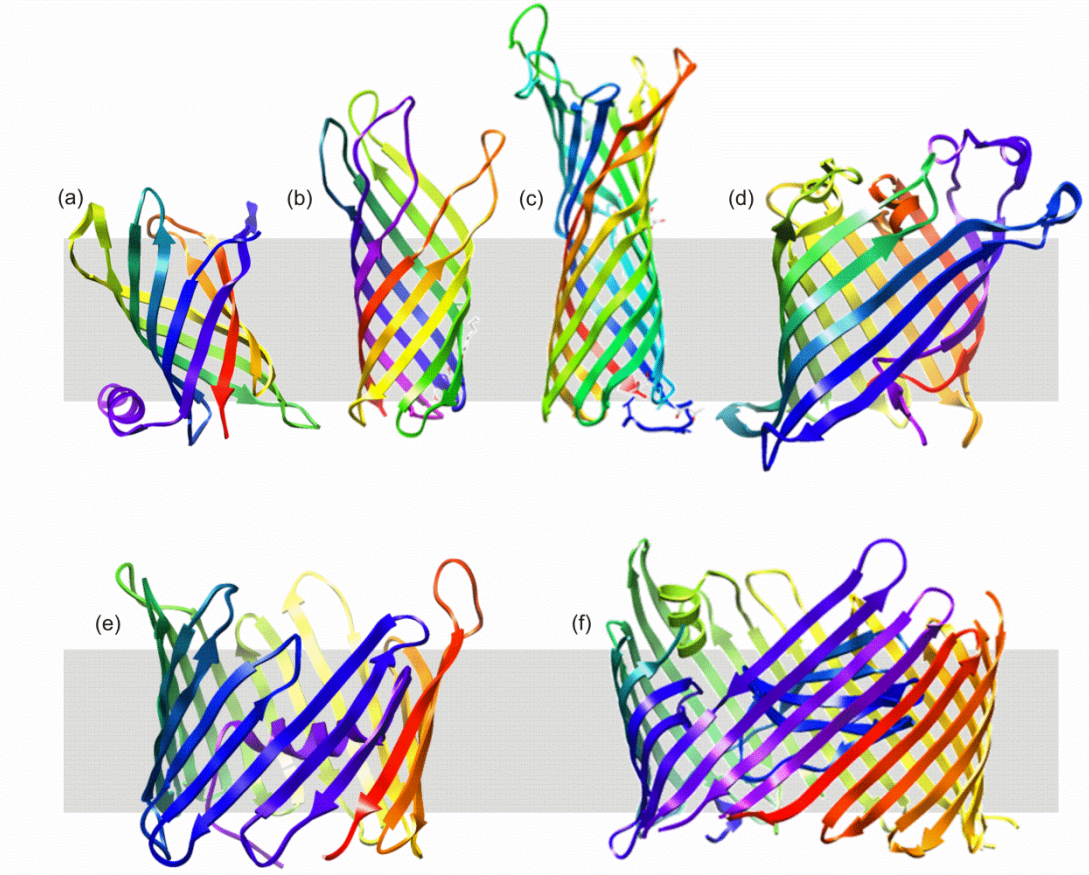

Figure 1.3 Example structures of β-barrel integral membrane proteins. Structures of (a) PagP, an 8-stranded palmitoyl transferase enzyme from *Escherichia coli* (1THQ[49]); (b) the 8-stranded transmembrane domain of OmpA, an ion channel from *Escherichia coli* (1BXW[50]); (c) the 10-stranded OM protease, OmpT, from *Escherichia coli* (1I78[51]); (d) the 12-stranded, colicin-secreting phospholipase A, OmpLa, from *Escherichia coli* (1QD5[52]); (e) the 19-stranded voltage-dependent anion-selective channel (hVDAC) from human mitochondria (2JK4[53]) and (f) the 24-stranded translocation domain of PapC from *Escherichia coli* (3FIP[54]). All structures shown were solved by X-ray crystallography. Proteins are coloured rainbow: violet (N-terminus) to red (C-terminus). The approximate position of the membrane is indicated with grey shading with the extracellular face (or cytoplasmic face for hVDAC) represented by the top edge of the shaded area. All images were generated from the Protein Data Bank[43] files from the accession numbers given in brackets using UCSF Chimera molecular visualisation application[44].

As more and more research is focused on the determination of membrane protein structure, a total of almost 400 unique structures have now been reported with both α-helical and β-barrel proteins represented[55]. Structure determination has been most successful using X-ray diffraction of crystallised proteins, although 100 of the reported structures were solved by nuclear magnetic resonance (NMR) spectroscopy, demonstrating the utility of both techniques in the membrane protein tool box[56]. While integral membrane proteins still account for only 1 % of the protein structures in the PBD almost 30 years after the first membrane protein structure was reported, significant progress is now being made in this area with more structures being determined every year[57].

1.3 BIOLOGICAL MEMBRANES

Lipids in cells have three main functions: energy storage, signal transduction and forming the matrix of biological membranes – the approximately 30 Å thick layer which encloses organelles within the cell, and the cell itself[58]. The fulfilment of these functions depends on the ability of the cell to synthesise a variety of lipid structures (for examples, see Figure 1.4), demonstrated by the fact that approximately 5 % of the eukaryotic genome is dedicated solely to lipid synthesis[58]. Cellular membranes are composed of polar lipids, which self associate into bilayers to shield the hydrophobic regions from the aqueous environment in a process entropically driven by water molecules[58]. The first evidence for the bilayer structure of membranes came in 1925 when Gorter and Grendel solubilised the lipid components of erythrocytes and measured the area covered when the lipids were deposited in a monolayer[59]. This basic model was replaced with the fluid mosaic model of membrane structure in which globular proteins can freely diffuse through the two-dimensional membrane in which they are embedded[60]. These simple models view the membrane as being predominantly homogeneous, and fail to predict the relatively slow diffusion constants of both lipids and proteins within the bilayer[61]. This observation can be explained using the picket and fence model, which suggests that transmembrane proteins anchored to the membrane skeleton act to compartmentalise the membrane[62; 63]. As a result, membrane components undergo "hop diffusion" where they are temporarily confined within a membrane compartment and undergo infrequent hops to adjacent compartments[62].

The structure and composition of bilayers vary greatly even within an organism, as shown in Figure 1.5, allowing variation in the properties of different membranes. For example, the plasma membrane of a eukaryotic cell is enriched with lipids such as sphingomyelin and cholesterol (for structures, see Figure 1.4) in order to decrease fluidity and resist mechanical stress, while late endosomes contain bis(monoacylglycero)phosphate (BMP) to aid in membrane fusion processes[58]. The mitochondrial membrane is much more reflective of bacterial membranes, being the only eukaryotic membrane to contain cardiolipin[58]. Modulation of lipid composition to adapt to different functional requirements implies the evolutionary advantage of an extensive and complex lipid repertoire[58].

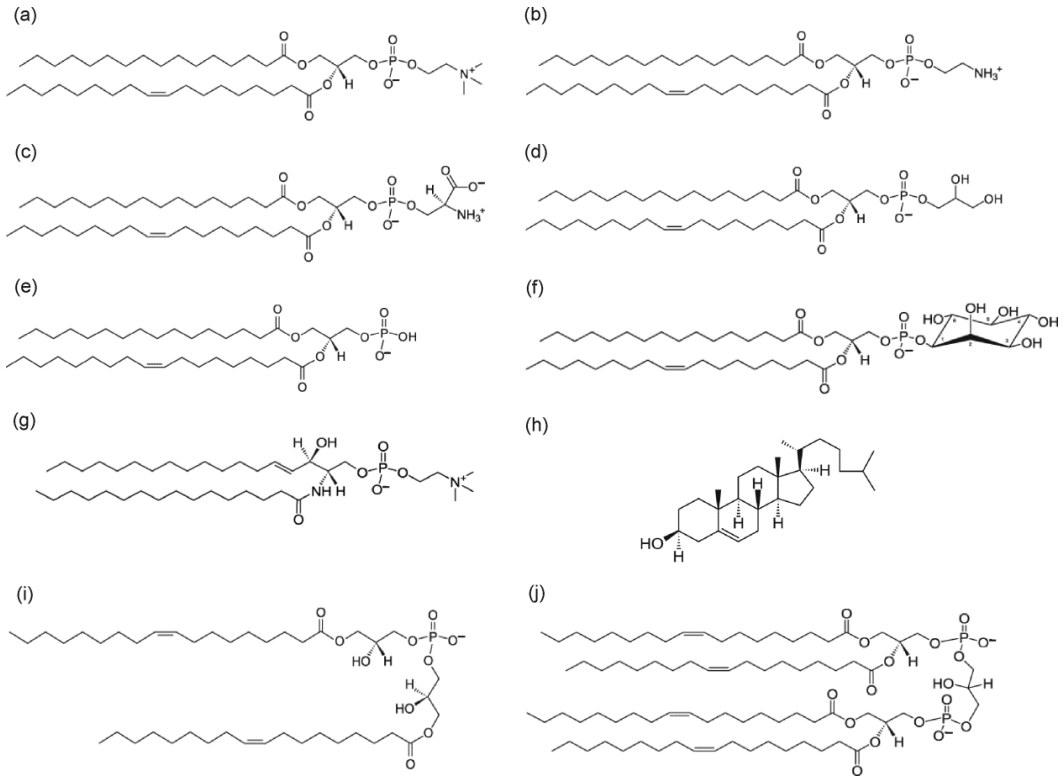

Figure 1.4 Example structures of structural lipids found in biological membranes. (a) phosphocholine (PC); (b) phosphoethanolamine (PE); (c) phosphoserine (PS); (d) phosphogylcerol (PG); (e) phosphatidic acid (PA); (f) phosphoinositol (PI); (g) sphingomyelin (SM); (h) cholesterol; (i) bis(monoacylglycero)phosphate (BMP); and (j) cardiolipin. Example chain lengths and saturation are depicted.

In addition to varying lipid compositions, membranes can have either a symmetrical or asymmetrical distribution of lipids between the two leaflets of the bilayer. Examples of

symmetrical membranes include the eukaryotic endoplasmic reticulum membrane and the inner membrane (IM) of Gram-negative bacteria, while asymmetrical membranes include the eukaryotic plasma membrane, endosomal membranes and the OM of Gram negative bacteria[32; 53; 64]. Within a leaflet, favourable interactions between some of the lipid components can generate domains of specific lipid compositions, known as lipid rafts, which are thought to be involved in localising membrane proteins[58]. These variable properties of biological membranes demonstrate the highly dynamic and heterogeneous nature of the lipid environment in which integral membrane proteins must insert, fold and function.

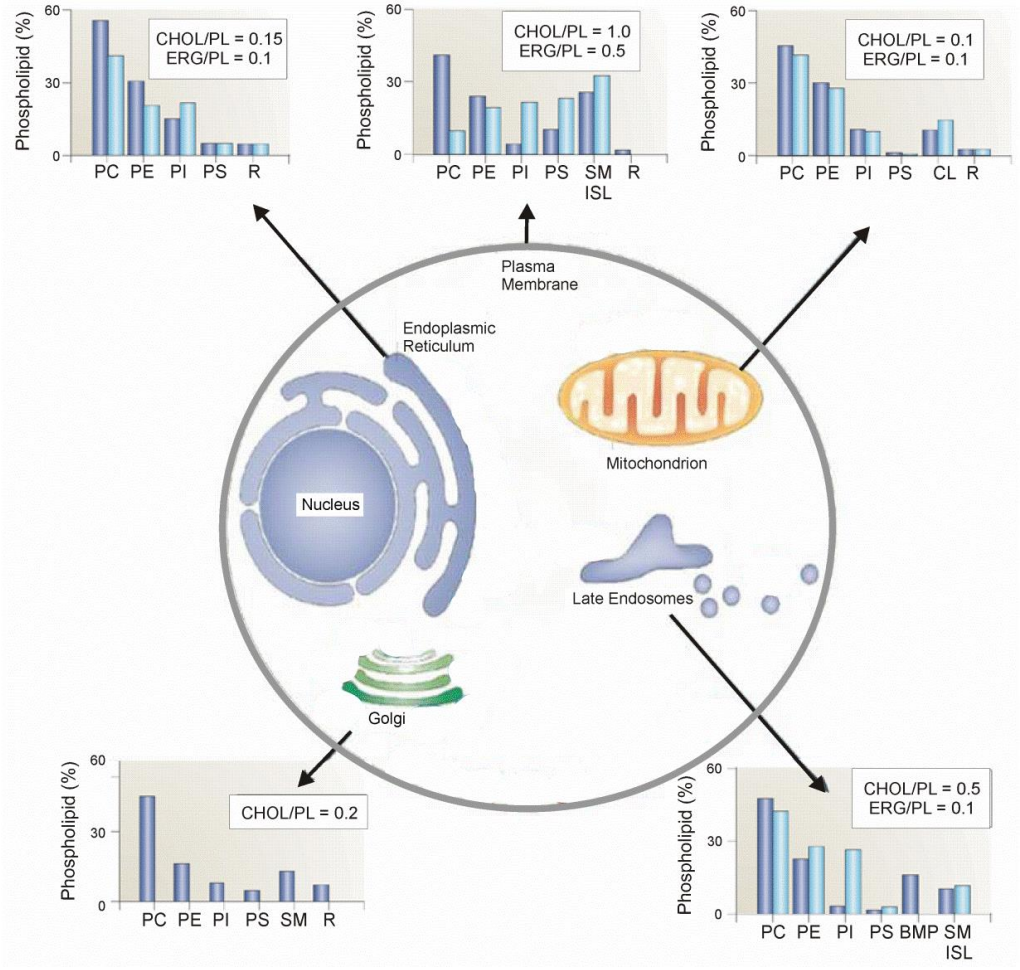

Figure 1.5 Lipid composition of different eukaryotic membranes. The lipid composition data (in graphs) are expressed as a percentage of the total phospholipid (PL) in mammals (dark blue) and yeast (light blue). As a measure of sterol content, the molar ratio of cholesterol (CHOL; in mammals) and ergosterol (ERG; in yeast) to phospholipid is also included. Lipid abbreviations are as follows: phosphocholine (PC), phosphoethanolamine (PE), phosphoinositol (PI), phosphoserine (PS), sphingomyelin (SM), yeast inositol sphingolipid (ISL), cardiolipin (CL), bis(monoacylglycero)phosphate (BMP) and all remaining lipids (R). Example lipid structures are depicted in Figure 1.4. This figure was adapted from van Meer *et al*, 2008[58].

1.4 The Cell Envelope of Gram-Negative Bacteria

The hallmark of Gram-negative bacteria, such as *Escherichia coli* (*E. coli*), is the cell envelope, which is composed of two lipid membranes enclosing an aqueous compartment called the periplasm (Figure 1.6)[64; 65]. The IM is a symmetric phospholipid bilayer composed of approximately 70 % PE, 25 % phosphoglycerol (PG) and 5 % or less cardiolipin and forms the barrier between the cytoplasm and the periplasm[64; 65; 66]. There are two types of protein associated with the IM; lipoproteins which undergo lipid modifications of the N-terminal cysteine to anchor them to the periplasmic face of the IM and the α-helical integral membrane proteins[64]. IM proteins are responsible for many cellular processes, including lipid synthesis and small molecule transport[64].

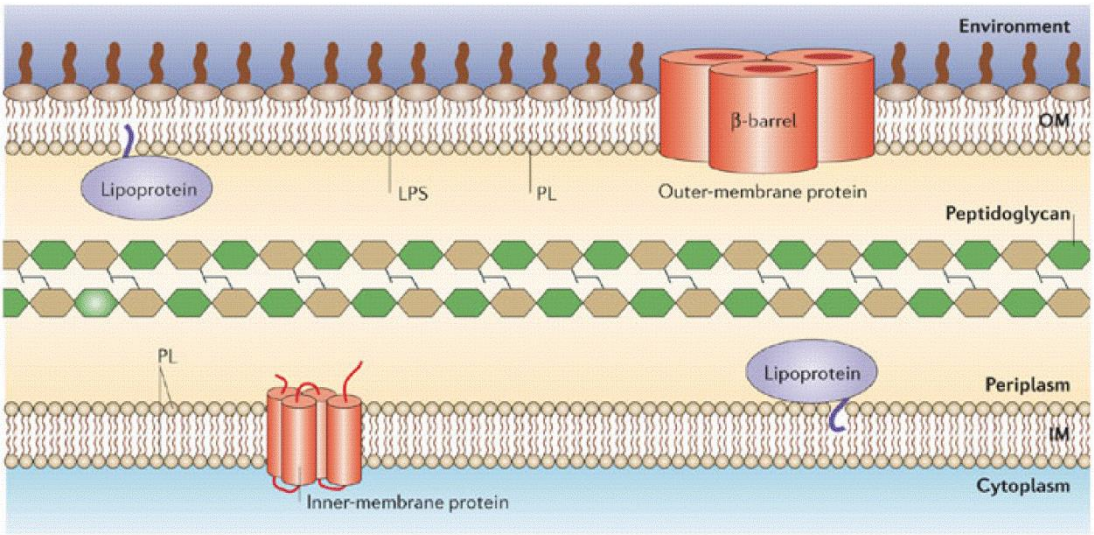

Figure 1.6 The cell envelope of Gram-negative bacteria. The cytoplasm of *E. coli* is surrounded by the inner membrane (IM), the periplasm and the outer membrane (OM). The IM is a symmetric phospholipid (PL) bilayer containing α-helical membrane proteins. The OM is an asymmetric bilayer of PL and lipopolysaccharide (LPS) and contains β-barrel integral membrane proteins. The periplasm is the aqueous compartment between the two membranes where the peptidoglycan cell wall is found. Both membranes have associated lipoproteins on their periplasmic faces. This figure was taken from Ruiz *et al* (2006)[64].

The periplasm is the compartment between the IM and the OM, which comprises around 10 % of the total cell volume and contains soluble proteins, as well as a peptidoglycan cell wall[64]. The peptidoglycan layer plays an important role in maintaining the shape of the cell and preventing lysis, while the periplasmic proteins are involved in maintaining the integrity of the cell envelope[64]. All of the processes which take place in the periplasm are independent of nucleotide hydrolysis, since no adenosine triphosphate (ATP) is present in this compartment[64; 67]. When energy is required, the cell relies on complex coupling mechanisms which often use the proton-motive force of the IM as the primary energy source[64].

In contrast with the IM, the OM is an asymmetric bilayer. The inner leaflet of the OM is comprised of phospholipids and is similar in composition to the IM, however, the PE content is enriched compared with the IM[64; 66]. The outer leaflet of the OM consists of lipopolysaccharide (LPS), a glycolipid typically consisting of lipid A, a core oligosaccharide and an O-antigen (Figure 1.7)[65; 68]. The structure of the O-antigen is highly variable even within species, with approximately 170 variants being recorded in *E. coli*[69]. The high number of fatty acid chains on LPS compared with phospholipids, and the fact that these chains are saturated, facilitates tight packing of the LPS molecules in the OM[68]. This confers low fluidity to the membrane and it is this property that is responsible for the low permeability of the OM[68]. As a consequence, the OM prevents the rapid diffusion of small, hydrophobic molecules and thus Gram-negative bacteria tend to be less susceptible to hydrophobic antibiotics than Gram-positive bacteria[68]. As with the IM, integral membrane proteins and lipoproteins are associated with the OM but in contrast, the integral outer membrane proteins (OMPs) usually have a β-barrel structure[45; 64]. The barrel is formed by membrane spanning β-strands which are held together by a lateral hydrogen bond network[45]. The structure of OMPs is discussed in more detail in Section 1.2 and several examples of integral membrane protein structures of both classes can be found in Figure 1.2 and Figure 1.3.

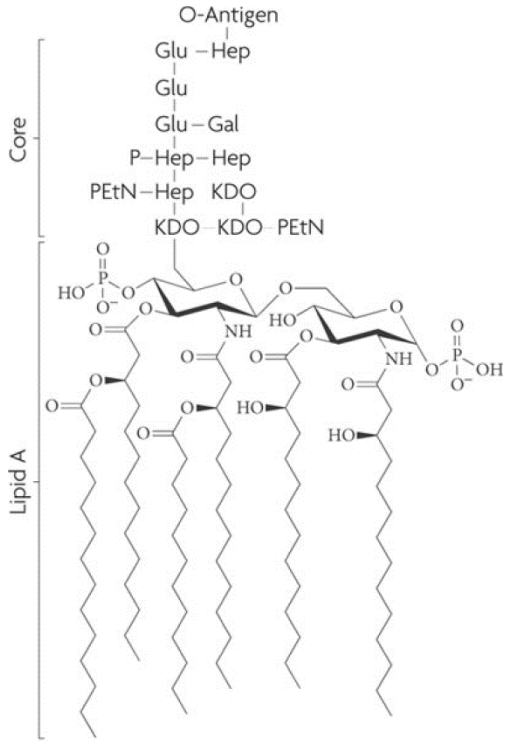

Figure 1.7 Structure of lipopolysaccharide in *E. coli*. The structure of the core region from K-12 strains is shown. Abbreviations: EtN, ethanolamine; Gal, D-galactose; Glu, D-glucose; Hep, L-glycero-D-manno-heptose; KDO, 3-deoxy-D-manno-oct-2-ulosonic acid; P, phosphate. Taken from Ruiz *et al*, (2009)[68].

1.5 MEMBRANE PROTEIN BIOGENESIS *IN VIVO*

1.5.1 SYNTHESIS AND TRANSLOCATION

Approximately 30 % of bacterial proteins function either as secreted proteins or in the cell envelope and as such they must be translocated across the IM following their synthesis in the cytosol[70]. The major translocation route in *E. coli* is the general secretion, or Sec, pathway[71]. The Sec translocase is comprised of a proteinaceous, heterotrimeric membrane protein complex, SecYEG, which forms a passive channel through the membrane, and a peripherally bound ATPase or ribosome[71]. The functional oligomeric state of the SecYEG complex has been debated due to the failure to detect oligomers using chemical cross-linking[71]. It has since been shown, however, that the cross-linking process itself actually interferes with the oligomerisation process[71]. Recent data demonstrate that SecYEG dimerisation in a back-to-back manner is required for activation of the translocation partner, the SecA ATPase, providing strong

evidence that the SecYEG dimer is the functional state, at least for SecA-dependent translocation[71].

There are two main targeting pathways to the Sec translocon, the first of which is used by IM proteins and occurs co-translationally[70]. As the nascent chain emerges from the ribosome, the hydrophobic transmembrane regions are recognised by the signal recognition particle (SRP)[70; 72]. The SRP interacts with its receptor, which is embedded in the cytosolic membrane and through the interaction of the SRP receptor with the SecYEG complex, the ribosome-associated nascent chain is targeted to the translocation pore[72]. The nascent chain diffuses through the channel in a manner independent of SecA (and hence nucleotide hydrolysis); however, it is coupled to chain elongation by the ribosome[72]. Transmembrane segments insert directly into the membrane *via* opening of a lateral gate in the SecYEG pore[72; 73; 74], while cytoplasmic domains emerge from the ribosome directly into the cytosol without entering the channel[72].

The second targeting pathway to the Sec translocon is utilised by secretory proteins and many OMPs after completion of translation (Figure 1.8)[71]. These proteins have an N-terminal signal sequence which causes them to be trafficked to the secretion-dedicated chaperone, SecB[70]. SecB is a homotetrameric protein which binds to unfolded proteins prior to their translocation to prevent aggregation in the cytosol[70]. Binding of SecB to the dimeric form of SecA can then occur and allows the transfer of the unfolded polypeptide to SecA, which is peripherally bound to the translocation channel[70]. Hydrolysis of ATP by the two nucleotide binding fold (NBF) domains in SecA provides the energy required to transfer the polypeptide into the SecYEG channel and translocation then occurs by a "pushing" mechanism, although the precise details of this mechanism remain unresolved[72]. A membrane-bound signal peptidase enzyme on the periplasmic face of the IM cleaves the signal sequence from translocating polypeptides as they emerge into the periplasm[70; 75].

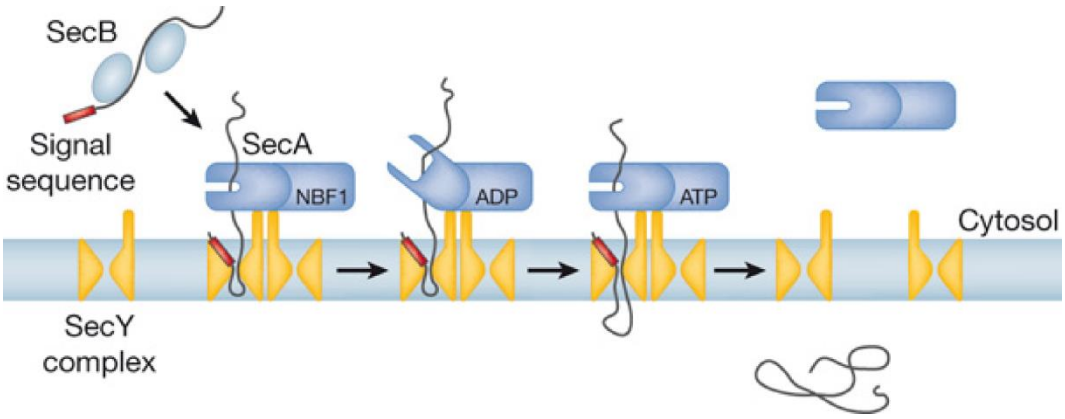

Figure 1.8 A model of post-translational translocation in bacteria. The protein to be translocated is maintained in an unfolded form by the SecB chaperone and is then targeted to the SecA ATPase. SecA is bound to the SecYEG complex and uses ATP hydrolysis to provide the pushing force for translocation through the SecYEG pore and into the periplasm. NBF1 is an abbreviation of nucleotide binding fold 1. This figure was taken from Rapoport (2007)[72].

1.5.2 Traversing the Periplasm

Following translocation into the periplasm, unfolded OMPs must be protected from aggregation, successfully traverse the periplasm including the peptidoglycan layer, and then correctly fold and insert into the OM[76]. These observations suggest that transport across the periplasm and membrane insertion are facilitated processes and, indeed, a number of periplasmic and OM proteins have been implicated in the OMP assembly pathway[76]. These proteins can be roughly grouped into three categories: proteases, chaperones which stabilise unfolded and non-native conformations of their client proteins, and folding catalysts, which catalyse rate-limiting steps in folding (Figure 1.9)[76].

Given the range of essential functions carried out by OMPs[46], it is unsurprising that the presence of unfolded OMPs in the periplasm is a strong inducer of the envelope stress response[76]. Exposed C-terminal residues of misfolded OMPs are recognised by the PDZ domain of the inner membrane-associated protease, DegS, and cause activation of the DegS protease domain[77]. PDZ is an acronym created from the first letter of three proteins which were first identified to contain such domains (post synaptic density protein (PSD-95), disc-large tumour suppressor and zonula occludens-1)[78]. PDZ domains share a common structural motif, are commonly found in signalling proteins in most organisms and bind to C-terminal regions of client proteins[79]. Activated DegS

triggers cleavage of the transmembrane regulator, RseA, which in turn causes transcriptional upregulation of genes of the σ^E-regulon (Figure 1.10)[76; 77]. Using the assumption that depletion of periplasmic chaperones will cause an increase is misfolded OMPs, identification of many of the known periplasmic folding factors arose by the genetic study of bacterial strains showing high σ^E activity[67]. To date, many of the studies of the roles of periplasmic chaperones have utilised *in vivo* depletion of the protein of interest. In order to fully understand the interactions of the OMP folding factors with their clients, however, more biochemical and biophysical studies are being conducted.

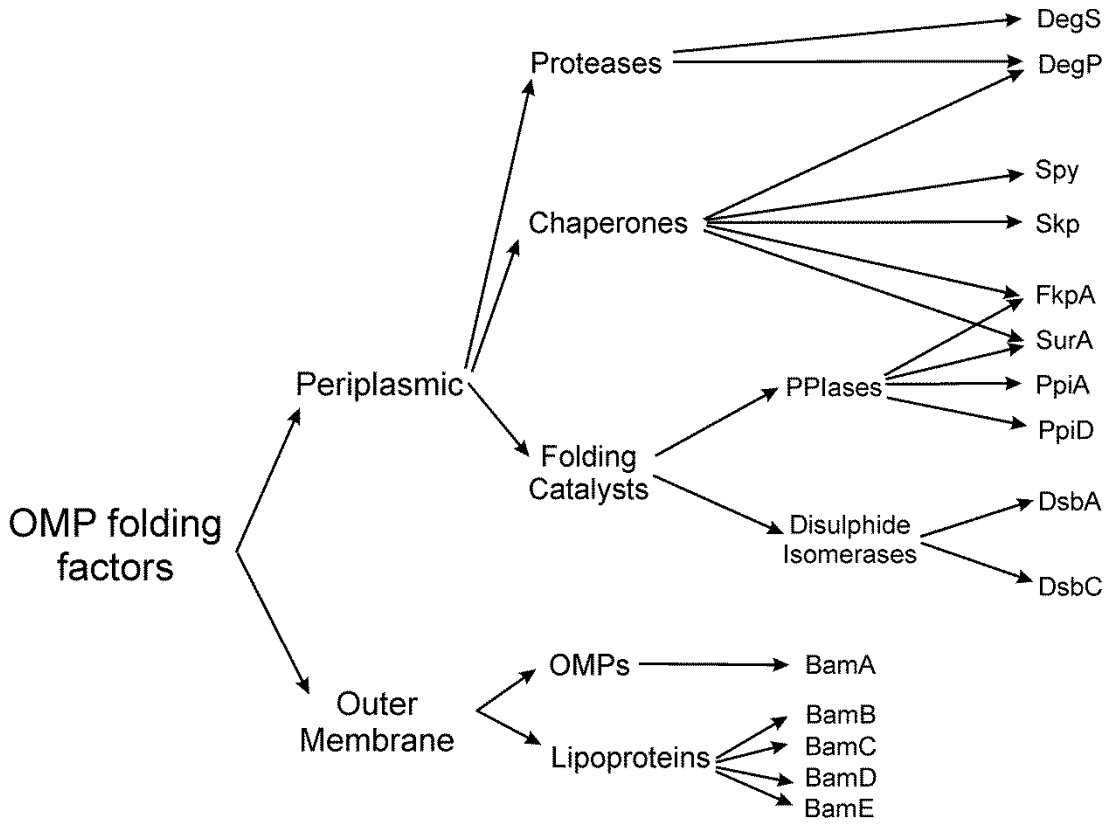

Figure 1.9 Overview of known folding factors in *E. coli* OMP biogenesis. Folding factors involved in OMP folding can be divided into OM associated proteins and periplasmic proteins. Periplasmic chaperones have been grouped into proteases, folding chaperones and folding catalysts, but some of the proteins have dual functions as indicated. Adapted from Mogensen and Otzen (2005)[76].

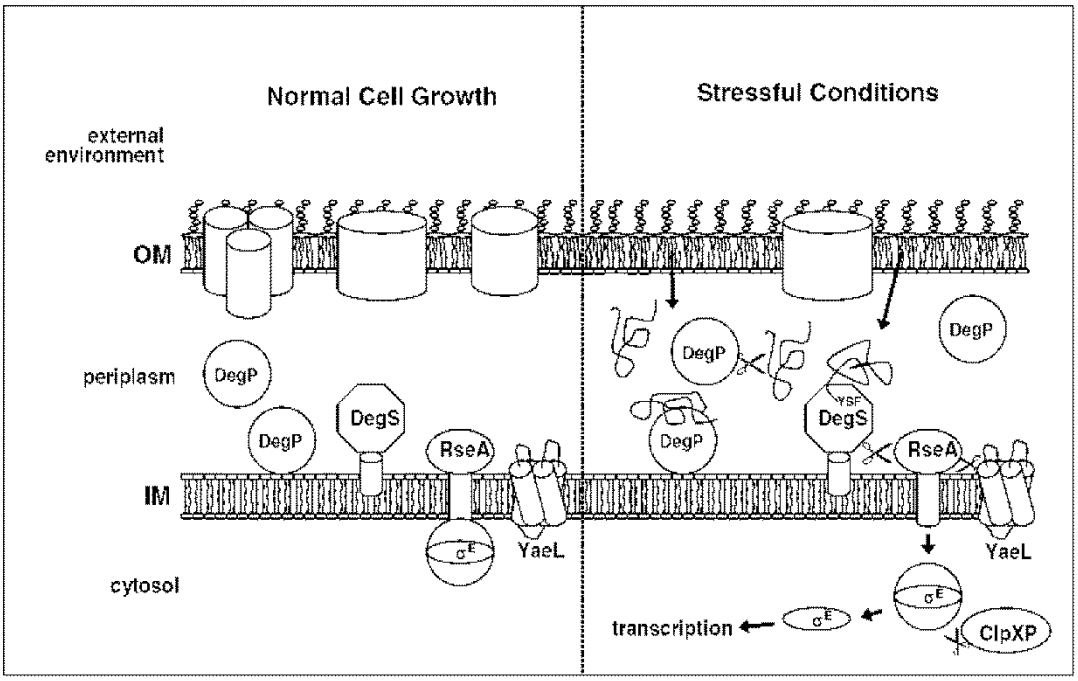

Figure 1.10 Activation of the σ^E stress response by unfolded OMPs. Under normal growth conditions (left), the σ^E transcription factor is sequestered in a complex with the cytoplasmic domain of the inactive protease, RseA. The membrane-associated periplasmic protease DegS is inactive. Under stress conditions (right), the exposed C-terminal residues of denatured OMPs activate DegS, which cleaves RseA. The IM protease YaeL (now known as RseP) then cleaves RseA again to release its cytoplasmic domain. Degradation of this RseA fragment by cytosolic proteases such as ClpXP releases σ^E for binding to RNA polymerase and initiates transcription of periplasmic chaperones and proteases. Abbreviations: IM, inner membrane; OM, outer membrane. This figure was taken from Mogensen and Otzen (2005)[76].

1.5.2.1 SurA

SurA was first identified in 1990 when it was shown to be required for the survival of *E. coli* in the stationary phase[80]. Initial characterisation described SurA as a parvulin-like peptidyl-prolyl isomerase (PPIase) involved in the proper assembly of major OMPs[67; 81; 82]. Trypsin digestion of OMPs in *surA* mutants showed that not all OMPs showed increased trypsin sensitivity in the absence of SurA, leading to the conclusion that SurA is not an essential folding factor[81]. However, the amounts of FadL, LamB, OmpA, OmpC, OmpF, OmpX and LptD were all found to be lower in *surA* mutants[83] and this is often a consequence of SurA depletion causing reduced mRNA levels as a result of σ^E induction[83; 84]. Similar results were obtained using proteomic analysis of *surA* mutants,

which also showed decreased levels of FadL, LamB, OmpA, OmpF, OmpX and LptD and an upregulation of proteins under the control of the σ^E regulon[85].

Crystallisation of SurA revealed a four-domain protein with two PPIase domains (P1 and P2) sandwiched between the N- and C-terminal domains (Figure 1.11)[86]. Of the two PPIase domains, P1 is packed against the core structure of the N- and C-terminal domains and does not show significant activity, while the more active P2 domain extends away from the core structure[82; 86; 87]. Deletion of both PPIase domains did not cause a significant loss of SurA function *in vivo* and the isolated PPIase domains failed to complement activity in *surA* mutants[87]. This led to the conclusion that the main function of SurA is as a chaperone[87]. Interestingly, mutations which would be expected to cause a loss of PPIase function in the P1 domain destabilised SurA *in vitro*, but increased chaperone activity *in vivo*[88]. This result suggests a regulatory function of the P1 domain and provides an explanation of the lack of PPIase activity displayed by this domain[88].

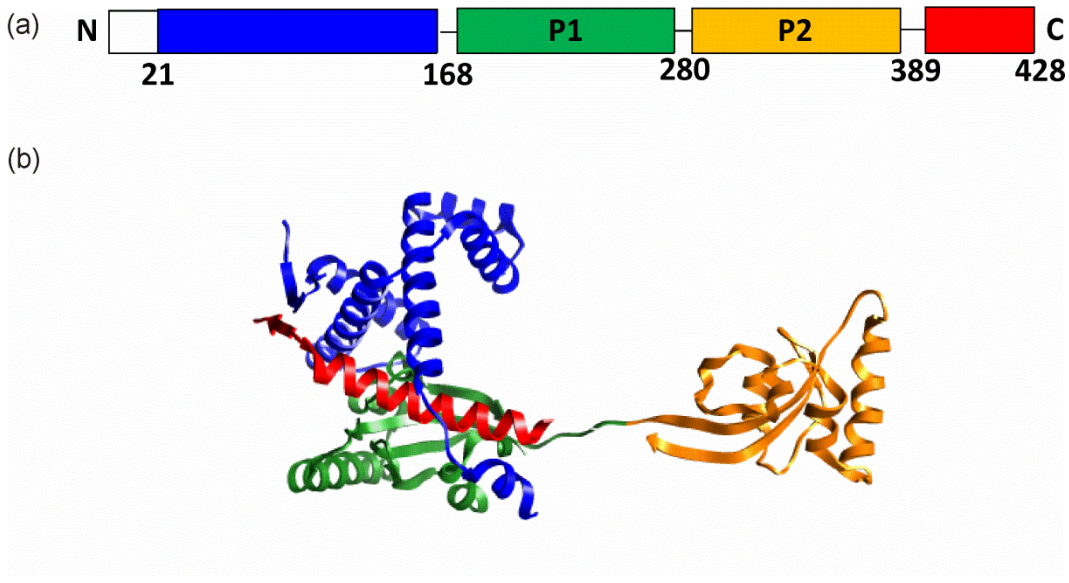

Figure 1.11 Crystallographic structure of SurA from *E. coli*. (a) Schematic of the domains of SurA: Signal sequence (white), N-terminal domain (blue), PPIase domain P1 (green), PPIase domain P2 (yellow) and C-terminal domain (red). Amino acid numbers shown below the schematic correspond to the N-terminal residue of each domain. (b) Ribbon diagram of SurA using the same colour coding as in (a) from PDB file 1M5Y[86] created in UCSF Chimera molecular visualisation application[44].

Binding studies using peptide sequences have been carried out on SurA *in vitro*, revealing a preference for aromatic-rich sequences with an Ar-X-Ar motif[89; 90]. Sequences containing this motif bind strongly to SurA with dissociation constants in the micromolar range[89], suggesting that SurA binds preferentially to OMPs due to the above average frequency of this motif in OMP sequences, particularly in the C-terminal region[83; 91]. Interestingly, this C-terminal motif is also responsible for activating DegS and triggering the σ^E stress response[76; 83]. In contrast with the peptide binding data, very few studies have reported binding of full length OMPs to SurA[92; 93]. The available data, however, suggest that unfolded OMPs are bound by SurA in preference to unfolded soluble proteins, while folded proteins do not bind[93]. While the polypeptide binding site of SurA has not been identified conclusively, an extended crevice, located in the core region, was observed in the crystal structure of SurA and is thought to be the site through which SurA:client interaction occurs[86].

The work of Sklar and co-workers demonstrated that depletion of SurA caused a loss of OM density that was not seen when other periplasmic folding factors, Skp and DegP, were depleted[94]. Additionally, SurA has been shown to be localised at the OM[89] and is the only periplasmic folding factor to have been successfully cross-linked to BamA of the β-barrel assembly machine (BAM) complex *in vivo*[94]. It is not yet known, however, if SurA binds directly to BamA, or if this interaction is through SurA-bound substrates[91]. Together, these observations have led to the hypothesis that SurA is the main chaperone for OMP transport *in vivo*[94; 95].

1.5.2.2 Skp

Although Skp was originally misidentified as a DNA-binding protein, an OMP and an LPS binding protein, it is now known that Skp is a periplasmic protein, as evidenced by the N-terminal signal sequence which targets Skp for translocation across the inner membrane[76]. Depletion of Skp *in vivo* led to a moderate reduction of OmpC, OmpF, OmpA and LamB in the OM fraction, however, these cells were still viable demonstrating that Skp is non-essential[96]. Further genetic studies revealed that the *skp surA* double mutant is bacteriostatic and that the *degP surA* double mutant results in a synthetic lethal phenotype[95]. Synthetic lethality describes the situation where a

combination of mutations leads to cell death, but single mutations of the same genes produce a viable phenotype[97]. This led to the prevailing hypothesis that the periplasmic chaperones Skp and DegP act on a separate, redundant pathway to SurA[95]. Proteomic analysis of the OMPs in a *skp* mutant suggested that none of the OMPs identified were significantly affected by this mutant; however, depletion of SurA in the *skp* strain reduced the levels of almost all OMPs consistent with the parallel pathways hypothesis[98]. While this hypothesis explains many of the observations above, the accumulation of protein aggregates in the periplasm of *skp degP* double mutants suggests that Skp may have an important role in maintaining the solubility of OMPs prior to folding[99]. Additionally, *skp fkpA* double mutants showed defects in the assembly of the essential OMP LptD, which could not be compensated by over-expression of SurA[100], suggesting that the chaperone pathways in the periplasm may not be completely redundant.

The chaperone activity of Skp was confirmed by the fact Skp can bind readily to unfolded OMPs[96; 101; 102], forming tight complexes with dissociation constants in the nanomolar range[101]. Furthermore, binding to OmpA was shown to occur only within the transmembrane domain of OmpA, and not the C-terminal periplasmic domain[96]. Analysis of the Skp:OmpA complex by NMR spectroscopy confirmed that the transmembrane domain of OmpA is Skp bound, while the periplasmic domain is free in solution and can fold independently[102]. Interaction of unfolded OMPs with Skp is thought to occur early after translocation, as evidenced by the cross-linking of Skp to PhoE at the periplasmic side of the IM in spheroplasts[103]. Furthermore, Skp has been shown to be required for the release of newly-translocated OmpA from the periplasmic side of the IM in spheroplasts[99], suggesting that the N-terminal residues of the client protein may interact with Skp first as they emerge from the translocon. This hypothesis has been confirmed using Förster resonance energy transfer (FRET) experiments which showed that the N-terminal residues of OmpC enter the Skp cavity first[104]. Molecular dynamics (MD) simulations also suggested that N-terminal OMP fragments enter the Skp cavity while C-terminal OMP fragments do not[104].

In solution, Skp is a trimeric protein with a structure that has been described as "jellyfish-like" (Figure 1.12a, b)[105; 106]. A β-barrel domain forms the body of the jellyfish, with long α-helical "tentacles" protruding outwards from the β-barrel[105; 106]. Skp is a basic protein, with a theoretical isoelectric point (pI) around 9.5, a property

which has been attributed to its tendency to mislocalise in cell fractionation experiments[76]. The charge distribution of Skp is notable, with many basic residues clustered around the tips of the α-helices resulting in a large macrodipole moment (Figure 1.12c, d)[105; 106], suggesting a role of Skp in substrate delivery to the OM. The α-helical domains define a central cavity enriched with hydrophobic residues (Figure 1.12e, f), which has been hypothesised to be involved in sequestering unfolded substrates to prevent their misfolding and aggregation[105; 106]. Indeed, this has now been shown to be correct by both NMR spectroscopy and MD simulation data[92; 102; 104]. In spite of low sequence identity, structural comparisons have been made between Skp and prefoldin, a cytosolic chaperone from *Methanobacterium thermoautotrophicum*, which protects partially folded proteins from aggregation and passes its clients to other chaperones to promote folding[105; 106]. Additionally, a putative LPS-binding site was identified on the outer edge of the Skp α-helices, similar to an LPS-binding motif found on the OMP FhuA[106]. This finding has led to the suggestion that Skp may act as a stress chaperone, acting to rescue OMPs which fall off the SurA chaperone pathway[107].

The ability of Skp to prevent aggregation of soluble proteins has also been tested against lysozyme[106] and single chain antibodies[108], revealing that Skp is able to chaperone soluble clients as well as OMPs. This may point to Skp as a dual-function chaperone in the periplasm, which interacts with all unfolded proteins following their translocation *via* the Sec machinery into the periplasm. In contrast with SurA, no common binding motif has been identified for the interaction of Skp with its substrates[76]. Both electrostatic and hydrophobic interactions have been implicated as being important in binding[101; 104; 108], with OMP interactions being shown to rely heavily on electrostatic interactions[104], and soluble client interactions suggested to be driven more by hydrophobic interactions[108].

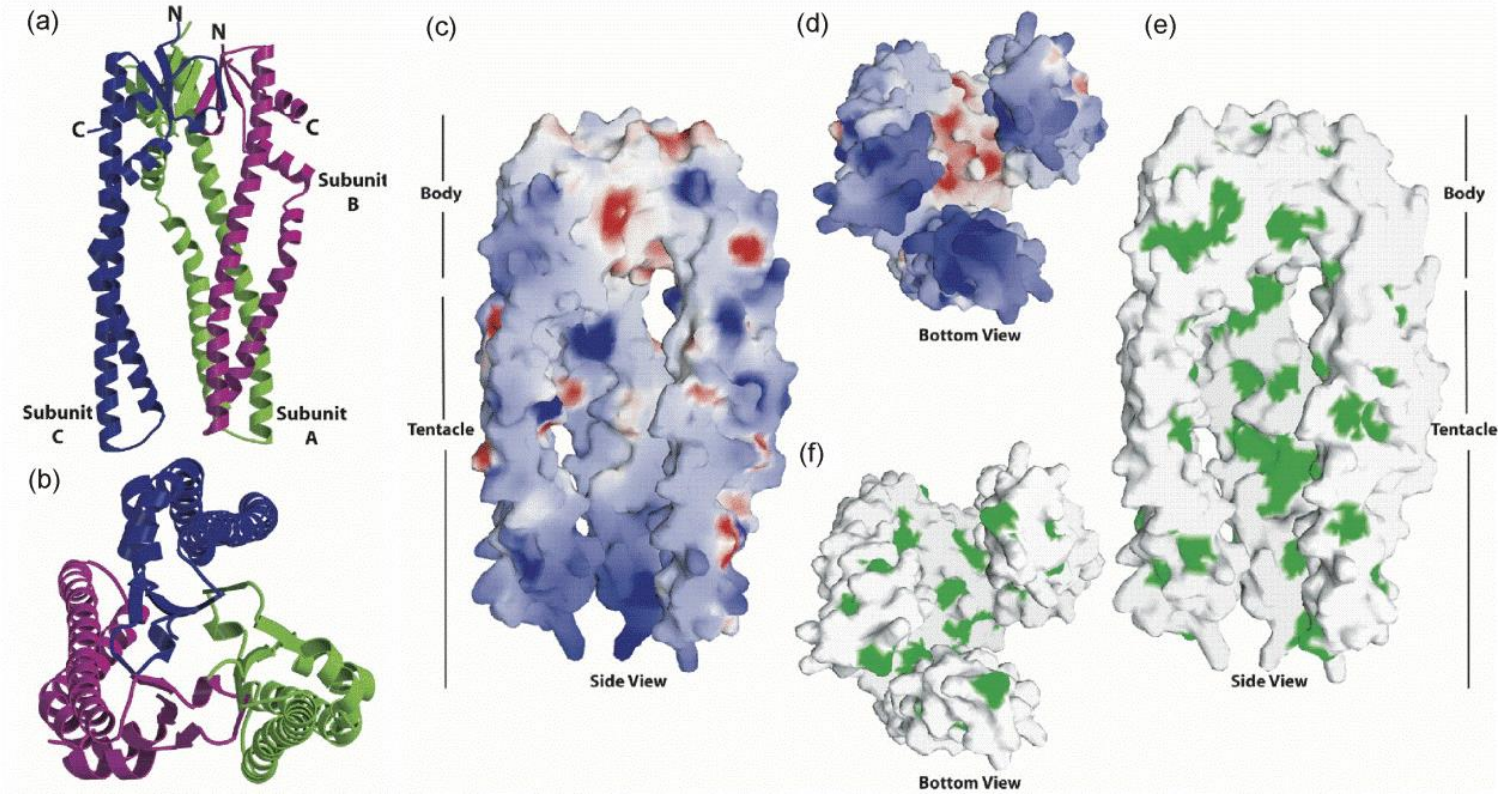

Figure 1.12 Crystallographic structure and surface properties of Skp from *E. coli*. Ribbon diagram representing the (a) side and (b) bottom views of the Skp trimer with the subunits A, B and C coloured in green, magenta and blue, respectively. Surface representations of the Skp trimer from the (c) side and (d) bottom with electrostatic potential mapped in red (negative) and blue (positive). Surface representations of the Skp trimer from the (e) side and (f) bottom with hydrophobic residues mapped in green. This figure has been adapted from Walton and Sousa (2004)[106].

1.5.2.3 Other Periplasmic Folding Factors

While SurA and Skp remain the most studied of the periplasmic folding factors, many others have been identified and shown to have varying effects on OMP biogenesis. Two disulphide bond isomerases, DsbA and DsbC, have been shown to function alongside the IM proteins DsbB and DsbD, respectively[109]. DsbA and DsbB are thought to catalyse initial disulphide bond formation, while DsbC and DsbD act as a proof-reading mechanism to rearrange any disulphide bonds which have been paired incorrectly[109]. Disulphide bonds are a rare structural feature in OMPs, particularly in the transmembrane domains, leading to the conclusion that DsbA and DsbC most likely have little involvement in OMP assembly[76].

The PPIase activity of SurA has been shown to be non-essential *in vivo*, as deletion of both domains did not affect OMP biogenesis[87]. One explanation for this is the presence of other PPIase proteins in the periplasm which can interact with OMPs. Three other PPIase enzymes are currently known: PpiA, PpiD and FkpA[65]. While PpiA has been shown to have the highest activity of these proteins *in vitro*, its deletion did not have any detectable effect on the assembly of OMPs *in vivo* and, as yet, no evidence has directly linked PpiA to OMP biogenesis[65; 76]. In contrast, deletion of the inner membrane-anchored PpiD was reported to cause a reduction in the levels of OmpC, OmpF, OmpA and LamB in the OM of the mutant cells and a double *ppiD surA* mutant was reported to be lethal, implicating the involvement of PPiD in OMP folding[65; 76]. Later studies did not replicate these findings, however, and the role of PpiD, if any, in OMP assembly remains to be elucidated[65].

FkpA, like SurA, is an example of a dual PPIase-chaperone folding factor and *fkpA* null mutants have been shown to be viable, but to display increased OM permeability and an up-regulation of periplasmic proteases[67]. Structural studies revealed that FkpA has an N-terminal chaperone domain, through which dimerisation occurs forming a v-shaped cleft, and a C-terminal PPIase domain (Figure 1.13)[110]. It has been hypothesised that substrate binding occurs in this cleft[110], however, no data are currently available to support this hypothesis. Until recently, chaperone activity of FkpA had been reported only for soluble proteins and the involvement of FkpA in OMP biogenesis was not well supported[65; 110]. Creation of a *skp fkpA* double mutant by Schwalm and co-workers

showed that assembly of LptD is compromised in this strain and has provided the first conclusive evidence that FkpA does indeed chaperone OMPs in the periplasm[100].

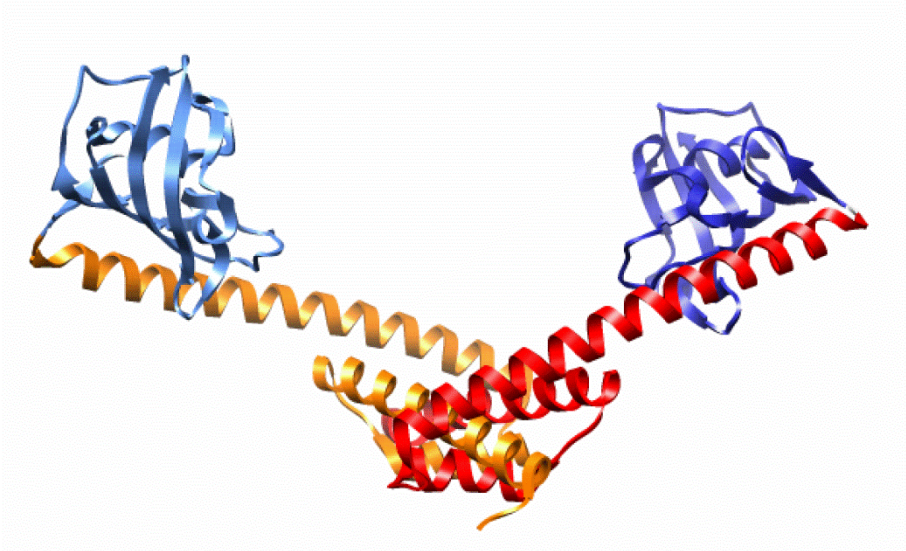

Figure 1.13 Crystallographic structure of FkpA from *E. coli*. Ribbon diagram of the FkpA dimer showing the N-terminal chaperone domains (red and orange) through which dimerisation occurs and the C-terminal PPIase domains (blue and light blue). This figure was created using PDB file 1Q6H[110] in UCSF Chimera molecular visualisation application[44].

Another example of a dual-function chaperone in the periplasm is the protease-chaperone DegP which, alongside DegS, belongs to the HtrA family of proteases[76]. Temperature changes were initially thought to cause the switch between the two activities of DegP, with the chaperone activity being dominant at 28 °C and the protease activity becoming dominant at 42 °C[76]. Structure-function analysis of DegP concluded that the resting state of DegP is a hexamer, in which the interactions between subunits block the protease sites and the central cavity is large enough to accommodate unfolded substrates[111]. Similarly to DegS, DegP recognises the C-terminal residues of misfolded OMPs *via* its PDZ domains and this was thought to cause structural reorganisation to larger proteolytically active oligomers comprised of 12 or 24 DegP monomers[83; 112; 113]. Cyro-electron microscopy of DegP in the presence of OmpC revealed a tetrahedral arrangement of DegP trimers with the inner cavity filled by a cylindrical area of electron density into which the structure of folded OmpC could be modelled (Figure 1.14)[112].

This led to the conclusion that the fate of DegP-bound substrates lies in their ability to adopt their native structure within the cavity, as only unfolded substrates can be degraded[112]. Recent data using cage-deficient mutants of DegP, which can only associate into trimers, revealed that these mutants are able to bind and degrade substrates without the need to form higher order oligomers formation, suggesting that cage formation in response to substrate binding may be linked solely to the chaperone function of DegP[114].

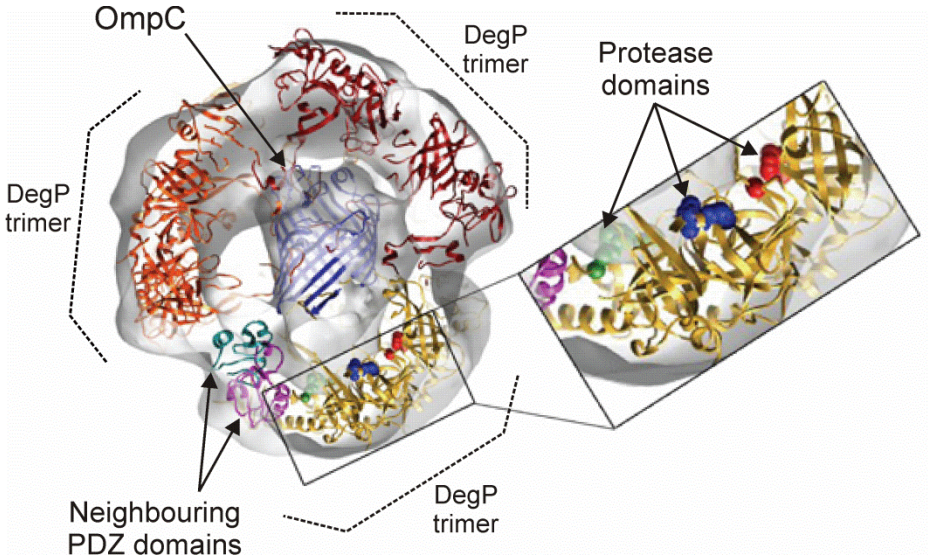

Figure 1.14 Cryo-electron microscopy structure of dodecameric DegP encapsulating OmpC. A central section of the electron microscopy map of DegP is depicted showing three DegP trimers in maroon, orange and gold. Adjacent PDZ domains interacting between neighbouring trimers are shown in cyan and magenta. Three protease sites are shown in green, red and blue and are magnified on the right. A folded OmpC monomer (blue) has been modelled into the electron density in the central cavity. This figure has been adapted from Krojer *et al* (2008)[112].

The most recently discovered periplasmic chaperone, Spy, was identified by its induced over-expression in bacterial strains expressing unstable periplasmic proteins[115]. *In vitro* characterisation of Spy revealed it to be an effective chaperone, suppressing aggregation and protecting substrates including the IM protein DsbB from inactivation by tannins[115]. Spy has a novel α-helical cradle structure, which is unlike that of any known chaperone[115]. It is not currently known if Spy is involved in OMP biogenesis, but the

discovery of a previously unknown chaperone only two years ago suggests that there is much still to be learned about the complex chaperone network in the *E. coli* periplasm. Additionally, studies on the periplasmic chaperones of *Neisseria meningitides* (*N. meningitides*) show than SurA deletion causes no detectable OMP assembly defects, while deletion of Skp caused lower levels of some porins, but did not affect all OMPs[116]. Complementation of the *skp* null mutant was not possible using *E. coli* Skp, suggesting that periplasmic chaperones may act in a species-specific manner and that generic principles of periplasmic chaperoning may be challenging to understand[116].

1.5.3 Insertion Into the Outer Membrane

Following chaperone-assisted transport across the periplasm, unfolded OMPs must insert and fold into the OM, a process which is assisted by the BAM complex in *E. coli*[117]. The BAM complex (Figure 1.15) is comprised of the OM protein BamA (previously YaeT) and four accessory lipoproteins, BamB (YfgL), BamC (NlpB), BamD (YfiO) and BamE (SmpA)[107; 117]. BamA is an essential protein, the depletion of which causes an accumulation of aggregated OMPs in the periplasm, leading to cell death[118]. BamA belongs to the conserved Omp85 super-family and was first identified in *N. meningitides*[119]. Homologues are found also in mitochondria (Sam50) and chloroplasts (Toc75)[120]. All of the homologues identified thus far have a C-terminal transmembrane β-barrel domain and a water-soluble N-terminal region[117]. Before its recent elucidation, the β-barrel domain of BamA was hypothesised to be structurally similar to that of FhaC, a secretory protein in *Bordetella pertussis*[117]. FhaC forms a 16-stranded β-barrel with a conserved loop which can fold over the pore to close it[121]. The conserved VRGY sequence at the tip of the loop has been shown to be functionally important in FhaC and current evidence suggests that this loop is also functionally important in BamA[117; 122; 123]. While the BamA β-barrel is also 16-stranded, the root mean squared deviation (RMSD) of the BamA and FhaC β-barrels is greater than 10 Å due to the differences in shear number and β-barrel architecture[122].

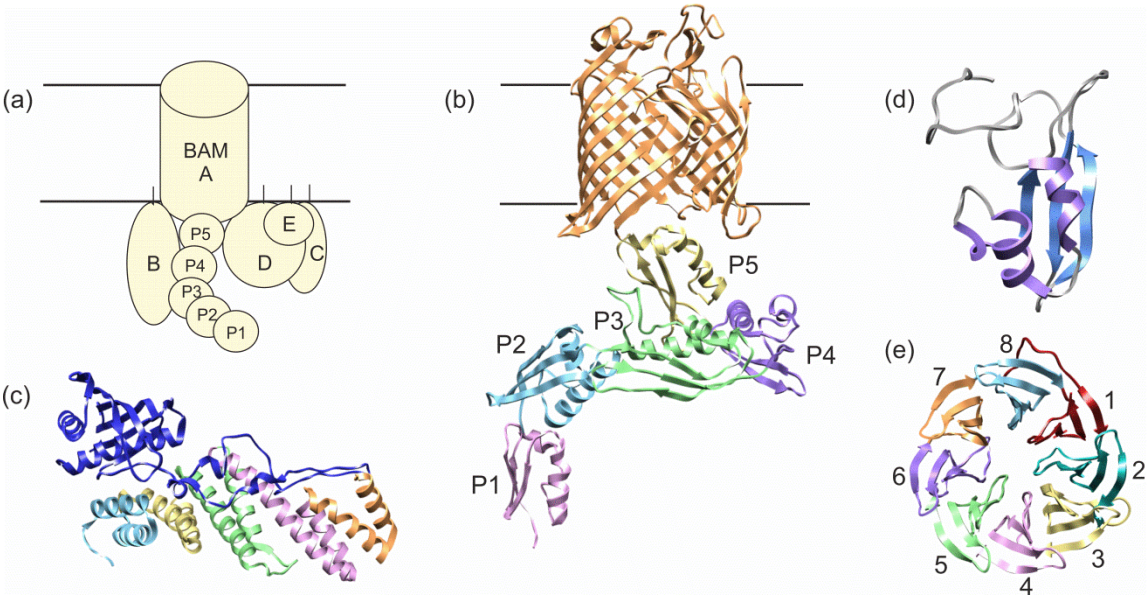

Figure 1.15 Structure of the BAM complex. (a) Schematic of the BAM complex with BAM proteins labelled A–E and POTRA domains labelled P1–5 taken from McMorran *et al* (2013)[124] (b) Crystal structure of *N. gonorrhoeae* BamA (4K3B [122]). The β-barrel domain is shown in orange. POTRA domains are labelled as in (a) and are shown in pink, blue, green, purple and yellow. (c) Crystal structure of *E. coli* BamB (3P1L [125]). The blades of the β-propeller structure are coloured individually and labelled 1–8. (d) Crystal structure of the N-terminal domain of *E. coli* BamC bound to BamD (2YHC [126]). BamC is shown in blue and the five TPR motifs of BamD coloured in light blue, yellow, green, pink and orange. (e) Lowest energy structure of *E. coli* BamE solved by NMR spectroscopy (2KXX [127]). Secondary structural elements of BamE are coloured. All images were generated from the PDB files from the accession numbers given in brackets using UCSF Chimera molecular visualisation application [44].

The soluble region of BamA is made up of a varying number of polypeptide transport-associated (POTRA) domains. Bacterial BamA homologues tend to have five POTRA domains, however, in chloroplasts Toc75 has three POTRA domains and in mitochondria Sam50 has only one POTRA domain[117]. Crystallisation of the four N-terminal POTRA domains (POTRAs 1–4) of *E. coli* revealed that in spite of low sequence identity between domains, the overall fold, a 3-stranded β-sheet with 2 α-helices, is conserved[120]. Previous binding studies of BamA revealed a direct interaction with unfolded OMPs and peptides mimicking the conserved aromatic-rich C-terminal sequence of OMPs[128], while evidence of β-augmentation between POTRA domains of neighbouring proteins in the crystal structure suggested a possible mechanism for substrate binding[120]. Deletion of individual POTRA domains from

BamA demonstrated that one of the functions of this soluble region is to scaffold the lipoproteins which make up the BAM complex alongside BamA[120]. Interestingly, while POTRAs 3–5 are essential for cell viability in *E. coli*, only POTRA 5 is essential in *N. meningitides*[117], suggesting differences in the roles of these conserved domains between bacterial species.

Of the four accessory lipoproteins in the BAM complex, only BamD has been shown to be essential for cell viability in *E. coli*, implying an important role in OMP biogenesis[117]. The crystal structure of BamD showed it to be a predominantly α-helical protein with five tetratricopeptide repeat (TPR) motifs, which shares structural similarity to other proteins which recognise C-terminal targeting sequences[129]. BamD interacts with the BamA POTRA 5 domain[117] and recent co-crystallisation with the non-essential lipoprotein BamC revealed that a binding pocket in the N-terminal region of BamD is the site of interaction of these two lipoproteins[126]. A truncated variant of BamD was used to demonstrate binding of this lipoprotein to synthetic peptides based on the OMP C-terminal targeting sequence, and this interaction occurred at the same binding pocket as the interaction with BamC[117; 126]. Together, these data have led to the hypothesis that BamC may act as a regulator of BamD substrate binding[117]. It was recently discovered, however, that regions of the BamC structure are surface exposed in intact *E. coli* cells, an incredibly unusual feature of an *E. coli* lipoprotein, and the functional implications of this observation, as yet, are unclear[130].

The most recently discovered of the accessory lipoproteins of the BAM complex is BamE. This lipoprotein associates with the BAM complex *via* interaction with BamD and has been shown to cause only minor OMP assembly defects when depleted[117; 131]. NMR spectroscopy revealed that BamE binds preferentially to PG lipids[131], which have been shown to enhance the insertion of OMPs into liposomes *in vitro*[117]. These data suggest that the role of BamE may be to promote OMP insertion into the OM. While BamC and BamE are associated with BamA *via* the interaction of BamD with the soluble POTRA domains, BamB is associated directly with BamA *via* POTRAs 2–5[117; 120]. Although not essential for cell viability, deletion of BamB results in reduced OMP assembly and a phenotype similar to SurA depletion mutants[132]. The crystal structure of BamB revealed it to an 8-bladed β-propeller with conserved residues on one face corresponding to the BamA interaction site[125; 133; 134; 135]. Crystal contacts are observed

on the outermost β-strands of each blade, suggesting possible substrate interaction sites[133]. In BamB deletion mutants, the OMPs most affected were those with larger β-barrels and, thus, it has been suggested that BamB has a role in substrate delivery to BamA, perhaps by increasing the substrate binding capacity of the BAM complex[117]. Indeed, BamB and BamD have been recently shown to bind to unfolded OmpA and BamA in the absence of other BAM components[136]. Both BamB and BamD accelerated BamA assembly into liposomes, however, OmpA folding was not similarly assisted by either lipoprotein, suggesting that one role of the BAM lipoproteins is to facilitate assembly of BamA[136].

Recently, a groundbreaking study was undertaken in which the BAM complex was over-expressed and purified as two sub-complexes, then reconstituted *in vitro*[137]. The complex was found to have a 1:1:1:1 ratio of BamA:B:C:D, but the stoichiometry of BamE present could not be determined due to its small size[137]. The activity of the BAM complex was monitored using the folding of the OM β-barrel peptidase enzyme, OmpT (Figure 1.3c), whose activity, and hence folding, can be monitored using a fluorogenic peptide[138]. OmpT was denatured in urea and then incubated with SurA before being added to a solution of liposomes containing the reconstituted BAM complex, LPS and the fluorogenic peptide[137]. Although OmpT could fold into liposomes in the absence of the BAM complex, the rate of assembly was increased in its presence. Moreover, solutions of liposomes containing Bam subcomplexes, BamAB or BamACDE, also showed reduced OmpT activity. The lower rates of OmpT folding seen in experiments using BAM subcomplexes, or in the absence of SurA, is consistent with observations of reduced viability of bacterial strains which lack either *bamB* or *surA*, while cells lacking both genes show severely reduced OMP assembly. These results have led to the conclusion that BamB and SurA have similar, but not redundant, roles *in vivo*[117; 137].

Based on all of the data currently available, the leading hypothesis suggests that newly translocated OMPs are chaperoned across the periplasm by SurA and delivered to the BAM complex (Figure 1.16)[107]. OMPs which are diverted from this pathway for any reason are picked up by the parallel Skp/DegP pathway where they may be degraded or refolded[107]. The mechanism of action of the BAM complex in OMP folding and membrane insertion is currently unknown, but several models have been proposed[117]. Possible models suggest that substrates are first translocated across the OM before

folding into the OM from outside the cell; that the BamA β-barrel acts as a scaffold for substrate β-sheet formation following insertion from the periplasmic side of the membrane; or that BamA forms multimers which create a pore in the OM into which the substrate can insert[117]. The elucidation of the BamA crystal structure, however, shows a large cavity in the β-barrel that is closed on the extracellular side and within which a substrate could be accommodated[122]. Additionally, MD simulation data point to the possible existence of a lateral gate in the BamA β-barrel, allowing substrates to move from the barrel interior into the membrane[122]. While the emergence of recent structural data has fuelled much progress in the study of OMP biogenesis, the exact roles of each of the BAM complex components and their combined mechanism of action remain to be elucidated.

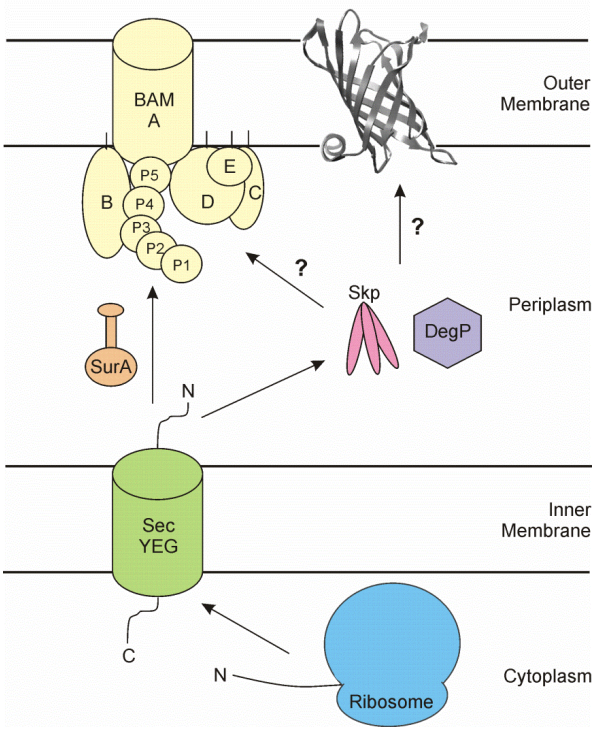

Figure 1.16 Schematic of the current model of biogenesis and chaperoning of OMPs in *E. coli*. OMPs are synthesised on the ribosome before post-translational translocation across the IM by the SecYEG translocon. Unfolded OMPs are then chaperoned across the periplasm to the β-barrel assembly (BAM) complex, which aids folding and insertion into the OM. BAM complex proteins are labelled A-E, and the periplasmic polypeptide transport-associated (POTRA) domains of BamA are labelled P1-5. Horizontal black lines indicate the approximate position of the inner and outer membranes. Taken from McMorran *et al* (2013)[124].

1.6 FOLDING STUDIES ON OUTER MEMBRANE PROTEINS *IN VITRO*

The study of OMP folding *in vitro* has been pioneered by work on the ion channel OmpA (Figure 1.3b) from *E. coli*, with initial experiments conducted by Schweizer *et al* as early as 1978[139] demonstrating that OmpA can fold in the presence of LPS and Triton X-100 detergent using a combination of sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE) and protease digestion. This study was aided by the observation that OMP samples prepared for SDS-PAGE without boiling (cold SDS-PAGE) migrate such that the folded and unfolded species have different apparent molecular weights as a consequence of the high stability of these proteins to denaturation with SDS when correctly folded[140; 141]. It was not until over a decade later that Surrey and Jähnig reported folding of OmpA into lipid bilayers[142], opening the door to studies of membrane protein folding that had lagged significantly behind studies of soluble protein folding. OmpA, solubilised in 8 M urea, was shown by these authors to insert spontaneously into *di*$C_{14:0}$PC small unilamellar vesicles (SUVs) upon rapid urea dilution[142]. Detailed kinetic studies were carried out subsequently by Surrey and Jähnig, as well as Kleinschmidt and Tamm, on the refolding of urea unfolded OmpA into liposomes using cold SDS-PAGE, far-ultra-violet (UV) circular dichroism (CD) and tryptophan fluorescence emission[143; 144]. Three kinetically distinct phases, as well as a rapid phase which manifested through a fluorescence increase upon initiation of the folding reaction, were observed[143; 144]. A scheme for the refolding pathway was proposed, beginning with collapse of the protein in the aqueous solution, followed by adsorption to the lipid membrane and folding to the native state *via* a speculative partially inserted intermediate (Figure 1.17)[143; 144]. The observed rates were concluded to be slower than would be likely *in vivo*, even at high temperatures, so the role of periplasmic chaperones in folding was implied[144].

Development of a method of time-resolved distance determination by fluorescence quenching, in which single tryptophan mutants were refolded into bilayers composed of brominated lipids, allowed a more detailed analysis of the folding mechanism of OmpA[145]. The results showed that the eight β-strands which make up the OmpA β-barrel insert into the membrane at the same time and so the previously suggested kinetic scheme was further developed to describe the intermediates as progressively

penetrating deeper into the membrane as the β-structure forms and packs together[145]. This provided the first evidence of the concerted nature of OMP folding and membrane insertion[145; 146].

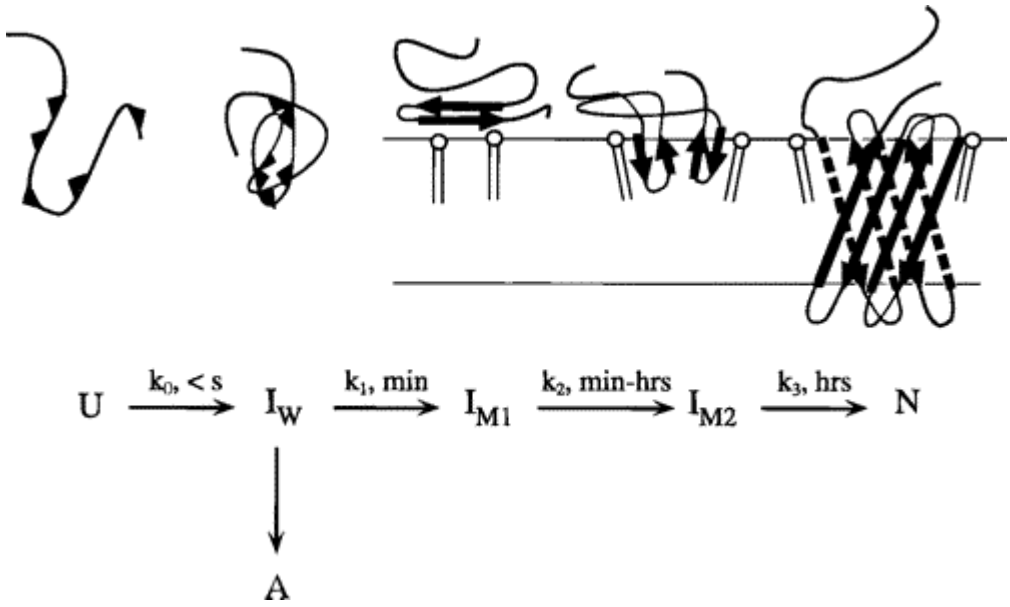

Figure 1.17 Proposed mechanism of OmpA folding and insertion into lipid bilayers *in vitro*. U is the unfolded state, I_w the collapsed state in water, A represents off-pathway aggregates, I_{M1} is a membrane-bound state, I_{M2} is a partially folded and inserted state, and N is the native state. Black triangles indicate hydrophobic regions. The timescale on which each step in the pathway occurs is indicated. Taken from Kleinschmidt and Tamm (1996)[144].

Analysis of the effect of membrane thickness on the kinetics of OmpA insertion and folding revealed that folding is slowed in large unilamellar vesicles (LUVs) in which the length of the lipid hydrocarbon chains is increased from ten to fourteen carbon atoms, at which point spontaneous insertion is no longer observed[147]. When SUVs of the same lipids were used, however, insertion was possible into thicker membranes due to the higher curvature of the vesicle[147]. This observation was supported by a study of amphipathic helix insertion into single liposomes of different diameters by Hatzakis *et al*, which suggested that high curvature would introduce packing defects that could aid insertion[148]. Kleinschmidt and Tamm suggested that since SUVs are more curved than biological membranes, OMPs would only insert into thinner membranes *in vivo* and this

could be a possible mechanism by which OMPs are prevented from inserting into the thick bacterial IM following translocation into the periplasm[147]. The same study also investigated the kinetics of secondary structure formation (using far-UV CD spectroscopy) and tertiary structure formation (using tryptophan fluorescence), revealing second order rate constants that were indistinguishable[147]. The simultaneous formation of secondary and tertiary structure was proposed to be necessary to form a hydrophobic surface capable of inserting into the bilayer[147].

More recently, measurement of the stability of OmpA has been achieved in several membrane mimetic environments[149; 150; 151]. Finding experimental conditions under which membrane protein folding is completely reversible and reaches equilibrium on an appropriate timescale is challenging and continues to impede studies on membrane protein stability[152]. Folding of OmpA into SUVs composed of 92.5 % $C_{16:0}C_{18:1}$PC and 7.5 % $C_{16:0}C_{18:1}$PG was reported to be reversible and analysis of the folding reaction by both tryptophan fluorescence emission and cold SDS-PAGE revealed good agreement between the two techniques, implying a two-state folding mechanism[149]. Introduction of guest lipids into this reversible system allowed the effects of bilayer thickness and the extent of lipid acyl chain saturation on OmpA stability to be measured, revealing that forces within the bilayer, such as curvature stress, affect thermodynamic stability[149]. Kinetic analysis of OmpA refolding into LUVs of varying diameter confirmed the role of curvature elastic stress in modulating folding rate, suggesting that irreversibility of folding of OmpA in LUVs may be a kinetic trapping effect[153].

The established conditions under which OmpA folds reversibility in 92.5 % $C_{16:0}C_{18:1}$PC and 7.5 % $C_{16:0}C_{18:1}$PG SUVs were utilised to conduct double mutant cycle analyses[154; 155]. The first application of this method was used to probe the interactions of residues in the suggested pore region of OmpA, revealing that salt bridge formation and rearrangement provide stabilising interactions, but also allow opening and closing of the OmpA pore[154]. Furthermore, replacement of the Glu52-Arg138 salt bridge with a disulphide bond not only blocked the pore but also caused growth defects in cells subjected to osmotic shock, suggesting that these interactions are functionally relevant *in vivo*[154]. A second application of this technique was used to probe how the interactions between aromatic residues in the intrafacial region of the β-barrel contribute to stability[155]. In both α-helical and β-sheet membrane proteins, aromatic residues are

enriched in the intrafacial region forming "aromatic girdles"[156]. This has been hypothesised to be stabilising, based on the partitioning energies of small model peptides into a model bilayer from aqueous solution[157]. Indeed, double mutant cycle analysis of OmpA provided the first evidence that in a fully assembled OMP clustering of aromatic residues provides a driving force for folding[155].

Analysis of OmpA folding in octyl maltoside (Figure 1.18a) micelles showed that, although OmpA folding remained irreversible when urea was used as a denaturant, reversibility could be obtained using guanidine hydrochloride (GuHCl) as an alternative denaturant[150]. A full kinetic characterisation was undertaken and revealed that OmpA folds *via* a three-state folding mechanism in octyl maltoside, and the calculated thermodynamic stability was much higher than that measured in SUVs ($\Delta G°_{octylmaltoside} \approx$ -65.2 kJ/mol, compared with $\Delta G°_{PC:PG\text{-}SUVs} \approx$ -14.2 kJ/mol), however, no explanation has been put forward for this seemingly counterintuitive result[149; 150]. This result was confirmed in a subsequent study of OmpA folding into detergent micelles of *N*-lauryl-*N*,*N*-dimethylammonium-*N*-oxide (LDAO, Figure 1.18b), revealing that the thermodynamic stability of OmpA in LDAO micelles is comparable to that in octyl maltoside micelles ($\Delta G°_{LDAO} \approx$ -60 kJ/mol) and that high kinetic stability is a barrier to reversibility studies[151]. The thermodynamic stability of OmpA in the Apol A8-35 (Figure 1.18c) was also measured and found to be much lower than that in detergent ($\Delta G°_{A8\text{-}35} \approx$ -8 kJ/mol), and the time taken to reach equilibrium was reduced to 25 days, compared with 52 days in LDAO[151]. Due to the low stability of liposomes in aqueous solution, incubation on these timescales to achieve equilibrium is not possible[158]. These studies clearly demonstrate the utility of alternative bilayer mimics to study membrane protein stability and function.

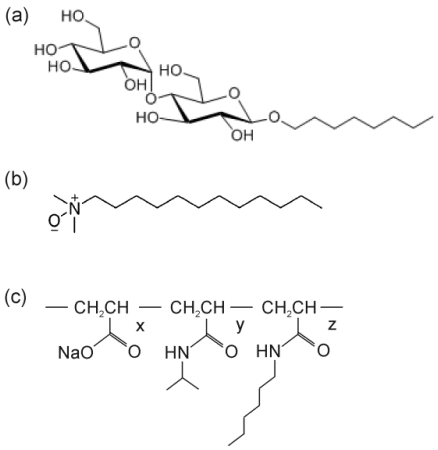

Figure 1.18 Structure of selected amphipatic molecules. Detergents (a) octylmaltoside and (b) LDAO, and (c) amphipol A8-35. Typically, x = 29–34 %, y = 25–28 %, and z = 39–44 %.

OmpA has also been used as the model folding system in the first *in vitro* study of the effect of a periplasmic chaperone on OMP folding into bilayers. Skp (see Section 1.5.2.2) was shown to solubilise unfolded OmpA in low urea concentrations (≈ 0.4 M) and to accelerate OmpA folding into negatively charged SUVs composed of 50 % *di*$C_{18:1}$PC, 30 % *di*$C_{18:1}$PE, and 20 % *di*$C_{18:1}$PG, but inhibit folding into zwitterionic *di*$C_{18:1}$PC SUVs over a 4 h time-course[159]. Additionally, pre-binding of the Skp:OmpA complex to LPS prior to addition to lipid was shown to promote folding into either lipid composition and LPS was suggested to have a role in OMP assembly[159]. This investigation was furthered recently by the inclusion of BamA in the OmpA refolding assay. Full-length BamA incorporated into liposomes increased the folding rate constant of OmpA into *di*$C_{12:0}$PC LUVs in a manner that was found to be dependent on the periplasmic POTRA domains[160]. While the transmembrane domain of BamA alone gave a small, two-fold increase in the folding rate, this was comparable to the rate increase seen when FomA, an OMP not implicated in OMP assembly, was incorporated into the liposomes. This suggests that the rate increase observed with the BamA β-barrel alone is due to non-specific perturbations of the local lipid environment by the incorporated OMP[160]. This study also showed that incorporation of 20 % *di*$C_{12:0}$PE into the *di*$C_{12:0}$PC LUVs has an inhibitory effect on OmpA folding which was overcome by the incorporation of BamA into the liposomes, suggesting how OMPs may overcome this inhibition *in vivo* as the OM is composed of approximately 80 % PE lipids[160]. Finally, the effect of BamA on the Skp-mediated folded of OmpA was investigated,

revealing that in the presence of BamA the previously observed retardation of OmpA folding into zwitterionic bilayers in the presence of Skp is ablated, possibly through the interaction of negatively charged patches on the surface of the POTRA domains with positively charged Skp[160; 161]. The available data on OmpA folding has provided valuable insights into the forces that govern the folding of OMPs. Lack of such extensive data on other OMPs, however, makes general conclusions hard to draw.

A recent study by Burgess *et al* adopted a screening approach using cold SDS-PAGE analysis to determine the folding yield of nine different OMPs (OmpX, OmpW, OmpA, PagP, OmpT, OmpLa, FadL, Omp85 (BamA) and OmpF) under different conditions[162]. The effects of pH, temperature, vesicle size and lipid composition were investigated, revealing that, in general, folding yield was increased by using higher pH, lipids with shorter acyl chains and smaller vesicle diameter. A single condition to maximise the folding yield of all nine OMPs could not be identified, however[162]. In previous studies, higher temperatures were reported to increase the folding yield of both OmpA[144] and FomA[163], however, increasing the reaction temperature had varying effects on the folded yield of the nine different OMPs[162]. For example, higher temperatures decreased the folding efficiency of OmpX, increased the folding efficiency of PagP, and had little effect on the folding efficiency of OmpW[162]. This variation was proposed to arise from the different aggregation propensities of these OMPs at higher temperatures[162]. Sedimentation velocity analytical ultra-centrifugation was employed to determine the aggregation propensities of the nine OMPs in the urea-unfolded state revealing varied results which did not correlate well with folding efficiency[164]. Some trends were observed, however, showing that the presence of salt and lower pH increase self-association, while the presence of $\geq$ 4 M urea kept all the OMPs in a monomeric state[164]. A thorough screen was then conducted to establish conditions for the reversible folding of OmpLa, which was attained using 100 mM citrate, pH 3.8 in *di*$C_{12:0}$PC LUVs at 37 °C[165]. These conditions also allowed reversible folding of PagP and OmpW, but not the other OMPs studied. Other studies providing useful information on the folding of OMPs into liposomes *in vitro* have not been discussed in detail here, however, a summary is provided in Table 1.1.

Protein	Notes	Techniques Used	Selected References
OmpA	Folding studies of OmpA are described in detail in Section 1.6	See Section 1.6	Kleinschmidt (2006)[146]
PagP	Folding studies of a C-terminally his-tagged construct of PagP (HT PagP) are described in detail in Section 1.8.2	See Section 1.8.2	Bishop *et al* (2000)[166] Ahn *et al* (2004)[49] Huysmans *et al* (2010)[167]
hVDAC	Human voltage-dependent anion-selective channel (hVDAC) can be folded to the native state in LUVs composed of lipids of varying acyl chain length (*di*$C_{10:0}$PC to *di*$C_{18:1}$PC). Folding yield was estimated to be 94 % in *di*$C_{12:0}$PC LUVs. Secondary structure content of VDAC in *di*$C_{12:0}$PC LUVs was not affected by changing pH from 7.0 to 3.0.	SDS-PAGE, far-UV CD, Trp fluorescence, sucrose density gradient centrifugation and proteolysis	Shanmugavadivu *et al* (2007)[168]
FomA	FomA can be folded to the native state in both *di*$C_{10:0}$PC and *di*$C_{18:1}$PC (LUVs and SUVs). Kinetic analysis of folding into *di*$C_{10:0}$PC and *di*$C_{18:1}$PC SUVs suggested that FomA folds *via* parallel pathways into both lipids. The folding halftime is dependent on acyl chain length and reaction temperature.	SDS-PAGE, Trp fluorescence, far-UV CD	Pocanschi *et al* (2006)[163]
OmpG	The porin OmpG reconstituted in native *E. coli* lipids is gated by conformational changes in extracellular loops in a pH-dependent manner (closed at pH 5.0). Unfolding under force reveals each β-hairpin unfolds individually. Refolding from this mechanically unfolded state also proceeds by sequential folding of individual β-hairpins.	Atomic force miscroscopy (AFM)	Sapra *et al* (2009)[169] Damaghi *et al* (2010)[170] Mari *et al* (2010)[171] Damaghi *et al* (2011)[172]
OmpF	Refolding of urea-solubilised OmpF into *di*$C_{14:0}$PC SUVs occurred at only 15 % yield. Refolding kinetics were biphasic but much slower than OmpA.	SDS-PAGE, Trp fluorescence, far-UV CD	Surrey *et al* (1996)[173]

Table 1.1 Summary of *in vitro* folding studies of OMPs into lipid bilayers.

1.7 Folding Studies on Protein Families: Using a Common Topology to Find Generic Principles

The OMP folding screens undertaken by the Fleming group[162; 164; 165; 174] are the first examples of comparative folding studies on a family of membrane proteins, however, the study of homologous proteins has been widely utilised in the field of protein folding. This approach has yielded valuable insights into the role of protein topology, amino acid sequence and secondary structure propensity on folding mechanism[175]. Following the detailed characterisation of the folding mechanism of the engrailed homeodomain (En-HD), four other members of the homeodomain-like super-family, c-Myb, hRAP1, Pit1 and hTRF1 were studied[13; 176]. This revealed that the mechanism varied from a pure "framework" mechanism (En-HD, Figure 1.19a) to a pure nucleation-condensation type mechanism (hTRF1), with the other proteins studied exhibiting mixed folding mechanisms[13; 176]. A similar scenario was seen in the peripheral subunit-binding domain superfamily and, as with the homeodomain-like superfamily proteins, the observed folding mechanism is determined by the inherent stability of individual secondary structural elements[175; 177].

Like the homeodomain-like proteins, spectrin domains also have an α-helical bundle topology and display varied folding mechanisms[177]. The spectrin domains R15, R16 and R17 (Figure 1.19b) have approximately 30 % sequence identity and similar thermodynamic stabilities, however, R15 folds around a thousand times faster than R16 and R17[21]. This observation could not be explained by simulation data or by characterising the transition state for folding, as all three domains showed transition states with similar structure and stability[175]. Instead, this difference was caused by increased folding landscape ruggedness, or internal friction, in the R16 and R17 domains[21]. By grafting the core of the R15 domain into R16 and R17, which was possible due to the nearly identical structure of these proteins, the internal friction of the latter domains was reduced and folding rate increased dramatically[21].

Differences in the kinetic profiles of related proteins do not always indicate a different folding pathway. For example, the bacterial immunity protein Im9 folds through a simple two-state pathway while its homologue, Im7 (Figure 1.19c), shows three-state folding *via* a populated intermediate[178; 179]. In spite of this difference, the

characterisation of the rate-limiting transition states of both proteins revealed that the transition states are remarkably similar overall[179]. Both transition states are compact and have three of the native helices formed. In the case of the Im7, however, the transition state was more conformationally restricted[179]. The hypothesis that the two proteins fold *via* the same pathway was confirmed by using protein engineering on Im9 to mimic the non-native interactions which stabilise the Im7 folding intermediate[180]. This resulted in a shift to a three-state folding mechanism of Im9, as the engineered stabilising interactions lowered the energy of the intermediate species on the folding pathway sufficiently for it to become significantly populated. Furthermore, the three-state folding variant of Im9 folded approximately twenty times faster than wild-type Im9, indicating that intermediate formation can increase folding rate even if non-native contacts are made[180].

Proteins with an all-β, immunoglobulin-like (Ig-like) fold have a characteristic "Greek-key" topology, with varying numbers and arrangements of β-strands, but there is low sequence identity between families[175]. Characterisation of the transition states of members of the fibronectin type III (fnIII) family revealed a highly conserved "obligate nucleus" between four hydrophobic residues from neighbouring β-strands[177]. Formation of this nucleus is required to initiate correct folding of the complex folded topology of these proteins, and only one Ig-like domain, CAfn2, was identified which lacked this domain[177]. Analysis of the transition state of this protein revealed that the nucleus was in fact still present, but had "slipped" down to an adjacent set of hydrophobic residues[181]. The evolutionarily unrelated Ig-like domain titin I27 (Figure 1.19b) was also found to have an obligate nucleus of structurally similar residues to those in the fnIII domains; however, I27 was also found to have an alternative folding pathway through a second obligate nucleus revealing the malleability of the folding pathways in Ig-like proteins[182].

Overall, the "fold approach" to studying protein folding mechanisms has been a useful tool for the understanding of the folding of soluble proteins. By selecting common topologies and evaluating the similarities and differences in the folding kinetics and structural mechanisms, many valuable insights have been gained[175; 177]. The application of this approach to the folding of membrane proteins will be critical in understanding the underlying principles in this challenging field.

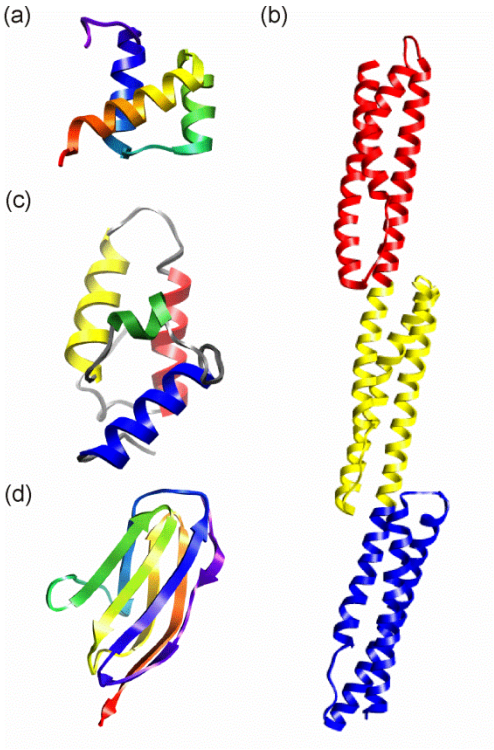

Figure 1.19 Structures of selected members of protein families which have been the subject of comparative folding studies. (a) Engrailed homeodomain (En-HD, 2JWT[183]) coloured violet (N-terminus) to red (C-terminus); (b) domains R15 (red), R16 (yellow) and R17 (blue) of α-spectrin (1U4Q[184]); (c) Im7 with helix I coloured red, helix II coloured yellow, helix III coloured green and helix IV coloured blue (1AYI[185]) and (d) titin I27 (1TIT[186]) coloured violet (N-terminus) to red (C-terminus). All images were generated from the Protein Data Bank[43] files from the accession numbers given in brackets using UCSF Chimera molecular visualisation application[44].

1.8 Application of Protein Folding Methods to the Study of Membrane Protein Folding

In recent years, many techniques have been developed to study the folding of water-soluble proteins yielding a near-atomistic view of the folding landscapes of these proteins[17]. In contrast, the experimental toolbox for the study of membrane protein folding thus far has been more limited, with fewer techniques successfully employed, often due to the interference of the membrane mimetic with spectroscopic signatures and the small differences in spectroscopic signals between the folded and unfolded states[187]. Much of the information available on membrane protein folding is the result

of simple techniques such as cold SDS-PAGE, tryptophan fluorescence emission and far-UV CD on a limited subset of proteins, as indicated by the studies summarised in Section 1.6 and Table 1.1. In spite of the challenges faced, new methodologies are continually being developed to aid the study of membrane protein folding.

The use of NMR spectroscopy to study residual structure in the unfolded ensemble of water-soluble proteins has revealed significant deviations from a random coil structure, for example in hen egg-white lysozyme[188] and in Im7 both under denaturing[189] and non-denaturing conditions[16]. It has been suggested that this residual structure is important in initiating folding by facilitating collapse of the polypeptide chain, and regions of residual structure in the natively unfolded variant of Im7 have been shown to correspond to the helices in the folded state[16]. This approach has been applied to the urea-denatured state of OmpX, revealing that the protein is globally unfolded but has two regions of non-random structure: one hydrophobic cluster and one helical region[190]. Analysis of peptides corresponding to these regions showed independent binding of the clusters to detergent micelles suggesting a role of residual structure in the initiation of folding and membrane-insertion[191]. The application of NMR spectroscopy to folded membrane proteins is complicated by the need to find a suitable membrane mimetic environment. Detergent micelles are widely used; however, problems with long term stability and maintaining proteins in a functional state are commonly encountered[192]. As a result, development of alternative non-micellar systems to stabilise the folded state of membrane proteins has become an active research area[192; 193; 194].

Protein engineering methods have been the most successful in probing the folding mechanisms of membrane proteins, having been used on the IM proteins bR[195; 196] (see Section 1.8.1) and DsbB[197] as well as the OMP, PagP[167] (see Section 1.8.2). Φ-value analysis is a powerful protein engineering technique, which can be used to map the formation of contacts in the transition states and intermediates formed during protein folding (Figure 1.20)[198; 199; 200]. Specific side chain interactions are deleted *via* strategic mutations, and the effect of these on the kinetic and thermodynamic parameters of the variant protein is measured and compared with wild-type[200]. The mutation causes a change in the free energy of activation (kinetic) and the equilibrium (thermodynamic) free energy of the folding reaction. The ratio of these changes is the Φ-value, which is usually between 0 and 1, and gives a measure of the change of stability of the transition state (or intermediate), compared with the change in stability of the native state[200]. Both

stabilities are measured relative to the denatured state, whose energy is assumed to be unaffected by the mutation[200]. Φ-values close to zero indicate that the transition state is unfolded in the region of the amino acid substitution, while higher Φ-values indicate that the region is folded in the transition state[198; 199]. Partial Φ-values are also observed and can be interpreted in several ways, for example, that native contacts have partially formed in the transition state or that multiple folding routes exist[200]. Φ-values can provide a wealth of information on the structure of the transition state and hence the folding mechanism of the protein[200]. Furthermore, the use of experimental Φ-values as constraints in protein folding simulations allows atomistic models of these ensembles to be created[201]. The main stumbling block for the application of this analysis is the need to find experimental conditions that promote reversible folding for the protein of interest, which has proved difficult for many membrane proteins[48; 150; 165]. Another potential pitfall of this technique is the relevance of measuring the equilibrium between folded and membrane inserted protein with unfolded protein in aqueous solution since it is not known if this equilibrium is likely to exist *in vivo*. It should be noted, however, that water-soluble proteins which do not unfold during their lifetime in the cell do exist and have been studied by kinetic and thermodynamic methods[202; 203; 204].

A promising approach for the study of transiently populated folding intermediates involves the use of a laser pulse to generate hydroxyl radicals (HO•) from a dilute solution of hydrogen peroxide (0.04 %)[205]. Exposure to HO• causes labelling of solvent accessible side chains at reaction rates close to the diffusion limit[205]. Laser pulses can be reduced to a duration of approximately 1 μs, allowing characterisation of rapid conformational changes[205]. Following HO• labelling, tryptic digestion and subsequent mass spectrometry analysis provides information on the protein conformation at the time the laser pulse was initiated[205]. This approach has been used successfully to study the dynamics of apo-myoglobin revealing structural flexibility of helix F[206] and the conformation of the intermediate in acid-induced myoglobin unfolding revealing a molten globule state[207]. It has also been used to study the temperature induced folding of barstar, which indentified a cluster of residues involved in folding nucleation by their decreasing oxidation levels in the initial stages of folding[208]. This technique has also been used to probe the conformation of bR[209; 210] as described in Section 1.8.1.

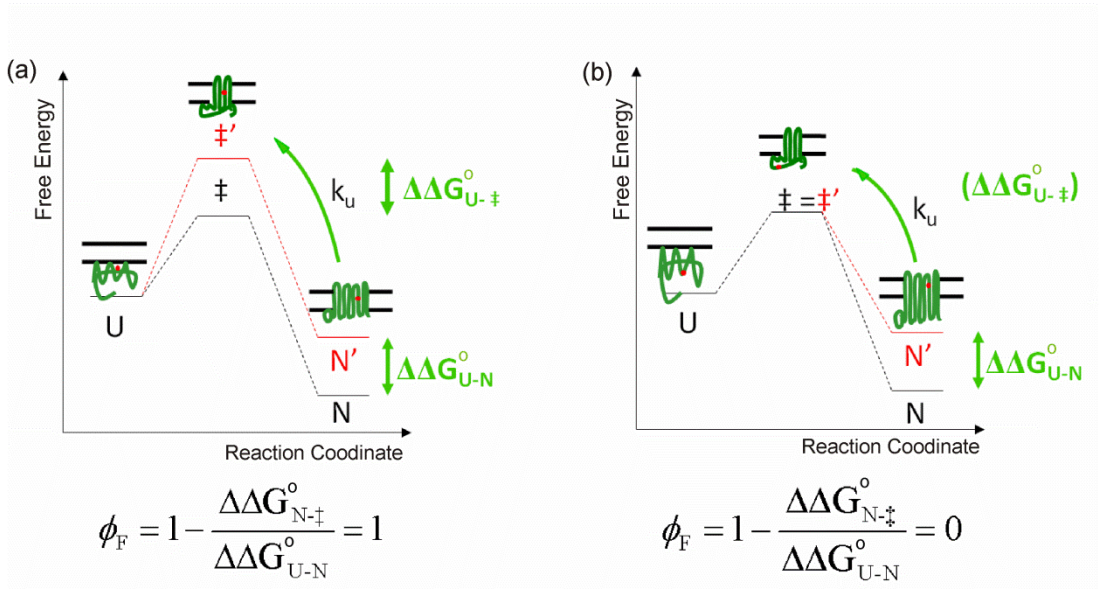

Figure 1.20 Schematic of the principles of Φ-value analysis. (a) A mutation (shown as a red dot) is made in a region of the protein which is native-like in the transition state (‡) leading to equal destabilisation of ‡ and the native state (N) resulting in a Φ-value of 1 or (b) a mutation is made in a region of the protein which is unfolded in ‡, leading to destabilisation of N only and a Φ-value of 0. It is assumed that the mutation does not affect the energy of the unfolded ensemble (U). Abbreviations: $\Delta\Delta G^{o}_{U-N}$ refers to the difference in the free energy change of folding upon mutation, $\Delta\Delta G^{o}_{U-‡}$ refers to the difference in the free energy change between U and ‡ upon mutation, $\Delta\Delta G^{o}_{N-‡}$ refers to the difference in the free energy change between N and ‡ upon mutation and k_u is the unfolding rate constant. This image was adapted with permission from Dr Gerard H. M. Huysmans (Institut Pasteur).

Another recently developed technique to probe protein folding is pulse proteolysis, which exploits the increased susceptibility of unfolded proteins to proteolytic digestion compared with that of folded proteins[211]. This method uses SDS-PAGE to quantify the digestion, circumventing the need to have large differences in spectroscopic signals between the native and unfolded states[211]. Additionally, small quantities of protein are required and the analysis can be carried out on crude protein samples compared with the large quantities of pure protein required for biophysical analysis[211]. This method was first tested on RNase H and the resulting free energy of unfolding was comparable to that calculated using far-UV CD as a folding probe[211]. Additionally, pulse proteolysis can be used to probe ligand binding. The stability of maltose binding protein was calculated in the presence and absence of maltose, and the change in free energy of unfolding used to calculate the binding free energy and dissociation constant of the complex. This yielded values in excellent agreement with those calculated by

fluorescence titration methods[211]. Recently, pulse proteolysis has been used to probe the stability of bR[212] as described in Section 1.8.1.

1.8.1 BACTERIORHODOPSIN

bR is a seven-helical bundle protein, which functions as a proton pump in the purple membrane of *Halobacterium salinarium* (Figure 1.2b)[213]. Oesterhelt and Stoeckenius successfully isolated bR in 1971, identifying the retinal co-factor[214] and went on to show that retinal binding is reversible[215]. In the early 1980s, Khorana and co-workers showed for the first time that bR denatured using SDS could be refolded to the native state *in vitro*, the first time this had been achieved with a membrane protein[216; 217]. The ability of soluble proteins to refold from a denatured state was demonstrated twenty years previously by Anfinsen's experiments on RNase A[4], illustrating how knowledge of membrane protein folding lags behind that of soluble proteins. Work by Popot and colleagues found that chymotrypsin-cleaved bR formed two fragments – one composed of helices A and B and the other of helices C–G, and these could be reconstituted into separate liposomes[218]. Following liposome fusion, the two fragments re-associated, bound to retinal and regenerated the cleaved bR structure.[218] This work led to the proposal that folding of bR follows a two step model where the helices form and then associate[29]. This model assumes that each helix can form stable secondary structure in isolation. The study of the stability of isolated bR helices in phospholipid vesicles, however, has shown that not all of the helices are independently stable, suggesting that simultaneous formation of secondary and tertiary contacts is required for the protein to fold[29; 219].

Kinetic analysis of SDS-denatured bR (bacterioopsin, bO) revealed that although bO retains approximated 42 % helical structure in the denatured state, its refolding kinetics are multiphasic[220]. Refolding in the absence of retinal occurs *via* three exponential phases to yield a non-native apo-protein containing helical content equivalent to seven α-helices[220]. When retinal is added, a further two exponential phases are seen which are attributable to the non-covalent then covalent binding of this co-factor[221]. Formation of the helices to yield the apo-protein is found to be the rate-limiting step, and this process can be slowed by increasing the proportion of *di*$C_{14:0}$PC in the mixed micelles into

which folding takes place to increase their rigidity, or by increasing the pH[222]. This effect was also observed also in lipid bilayers as increasing the amount of PE lipids present in PC vesicles slowed bR folding and this was proposed to be due to the increased lateral pressure in the chain region of the membrane either slowing insertion or slowing rearrangement of the inserted helices[223].

A complete thermodynamic and kinetic analysis was undertaken in mixed micelles of 3-[(3-cholamidopropyl)dimethylammonio]-1-propanesulfonate (CHAPS) and *di*$C_{14:0}$PC for denaturation from the folded state of bR to a partially unfolded retinal-bound state using SDS[224]. The natural logarithm of the folding and unfolding rate constants in different concentrations of SDS were found to give a chevron plot with linear folding and unfolding branches over the limited range of SDS concentrations used in the experiment. The results are consistent with the two-state assumption (Figure 1.21)[224]. It should be noted, however, that a long extrapolation is required to obtain the folding and unfolding rate constants in the absence of SDS[224]. In spite of the experimental limitations, this study showed a reversible unfolding transition of bR, with an associated large free energy change of approximately 83 kJ/mol[224].

Using the reversible folding conditions found in this study, a Φ-value analysis of bR was carried out by conducting an alanine scan on helix B[195] and helix G[196] (Figure 1.22). In helix B, most of the calculated Φ-values were close to one, indicating that this region of protein is highly structured in the transition state, with many native contacts having formed[195]. Two residues with lower Φ-values were observed, Y43 and T46, where helix B packs on to other helices in the native state[195]. The fractional Φ-values observed in these positions suggest partial formation of tertiary structure at these residues in the transition state as the helices begin to pack together[195]. In contrast, the Φ-values in helix G were mostly close to 0 indicating that this helix is mostly unstructured in the transition state and suggests folding is polarised with the N-terminal helices folding first[196]. Double mutant cycle analysis revealed that the low Φ-value at T46 is due to a missing hydrogen bond with neighbouring helix C and missing hydrogen bonds were also found between helices F and G, most likely due to the lack of structure in this region[196]. While this study paved the way for further application of this technique to integral membrane proteins, the results concentrate mostly on two neighbouring helices and do not give a full picture of the transition state for folding of bR.

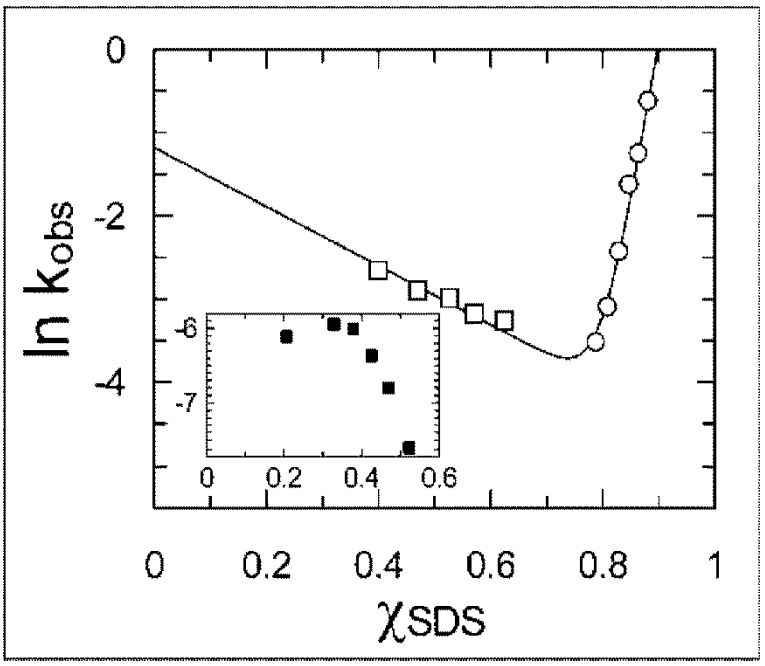

Figure 1.21 The linear chevron plot obtained for the folding and unfolding of bacteriorhodopsin. The natural logarithm of the observed folding (open squares) and unfolding (open circles) rate constants (ln k_{obs}) is plotted against the mole fraction of SDS (χSDS) and fitted to a two-state equation. Non linear refolding from bO_{390} (inset, filled squares) is observed. This figure is taken from Curnow *et al*, 2007[224].

The four helix bundle DsbB has also been subjected recently to a Φ-value analysis in mixed *n*-dodecyl-β-D-maltopyranoside (DDM)/SDS micelles based on kinetic measurements[197; 225]. Kinetics were thought to represent a three-state process with a reaction intermediate, and assume the reaction reaches equilibrium in the timescale used[225]. Twelve mutations were made throughout the structure of DsbB and used to investigate the structure of the rate-limiting transition state, which leads to the formation of the folding intermediate[197]. The Φ-values obtained, in general, are close to 0 indicating a lack of structure in this transition state, however, three higher Φ-values clustered on one end of the helices suggest polarisation in the transition state similar to that seen for bR[197; 226].

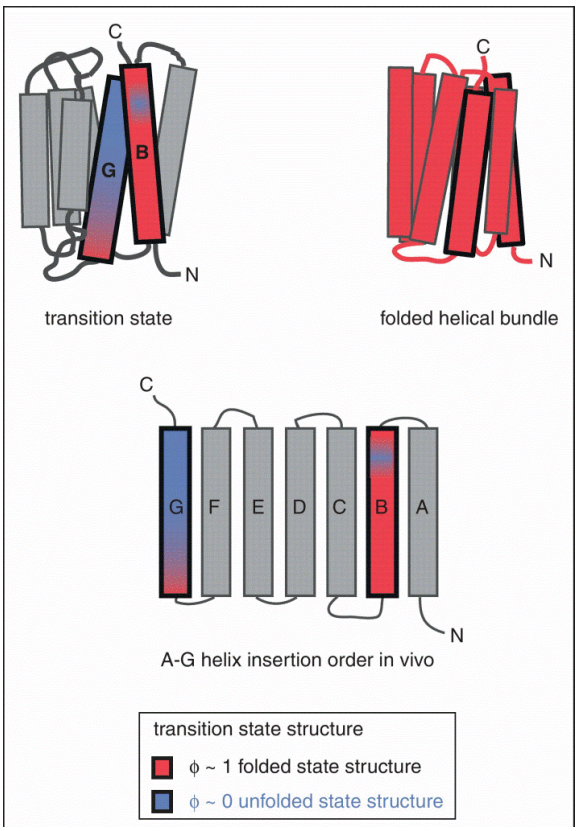

Figure 1.22 Schematic of the transition state structure of bacteriorhodopsin from Φ-value analysis. Alanine scanning revealed that in the transition state (top left) helix B is mostly helical as in the native state (top right) while helix G is mostly unstructured. Helix B inserts into the membrane earlier than helix G *in vivo* (bottom). Regions with Φ-values close to 1 are shown in red, regions with Φ-values close to 0 are shown in blue and undetermined Φ-values are grey. Helices B and G are shown outlined in bold. The cytoplasmic side is at the top and extracellular side at the bottom of the diagram. This figure was taken from Booth (2012)[226].

The technique of pulse proteolysis (described in Section 1.8) has also been used to study the folding of bR[212]. The results were compared with those obtained when using the absorbance at 560 nm (A_{560}), which reports on the retinal co-factor, as a probe. These studies demonstrated good agreement between the two techniques[212]. Refolding using pulse proteolysis revealed a large hysteresis suggesting that bR does not refold on this timescale as had been previously suggested by Curnow and Booth[224; 227]. In fact, the refolding rate constant was approximately forty-fold lower than that reported at the same mole fraction of SDS, a discrepancy arising from experimental differences. In the pulse proteolysis method the mole fraction of *di*$C_{14:0}$PC/CHAPS was kept constant as SDS was varied, however, in the Booth method the mole fraction of *di*$C_{14:0}$PC/CHAPS was allowed to vary as SDS varied[224; 227]. Even at very high concentrations of

di$C_{14:0}$PC/CHAPS, the pulse proteolysis method resulted in much lower folding and unfolding rate constants and suggested that the previous analysis may have been performed without reaching a true conformational equilibrium[224; 227].

Studies of folding, whether followed using Φ-values and traditional spectroscopic techniques or using newer techniques such as pulse proteolysis, provide an indirect view of the transition state for folding. Direct methods to probe transiently populated species have been used to study bR with some success. Hydrogen-deuterium exchange in conjunction with double mutant cycle analysis was used to probe the effect of mutations which disrupt hydrogen-bonding in bR, revealing that stabilisation from hydrogen-bonding is weak and up to 4 % of the possible hydrogen bonds in the core of the membrane-spanning region are unsatisfied[228].

One major drawback with hydrogen-deuterium exchange is the back-exchange of protons following the quenching step[229], which can be circumvented using covalent labelling techniques. Many covalent labels react too slowly to be effective in the study of transiently populated species[205], however, the use of hydroxyl radicals as described in Section 1.8 allows the study of these species. Interestingly, when bR in native purple membrane was studied using oxidative labelling, methionine residues were selectively labelled in contrast with the labelling behaviour of soluble proteins which was shown to be much less specific[209]. Three solvent-exposed methionine residues were extensively labelled, while the remaining six methionine residues showed much lower levels of oxidation consistent with the native structure of bR[209]. This analysis was then repeated under denaturing conditions using SDS, low pH or heat[210]. Under acidic conditions, the labelling pattern was unchanged from native bR, consistent with crystallography data[210]. SDS and heat denaturation produced similar results showing almost complete deprotection of helix A and partial deprotection of helix D[210]. Following hydrolysis of the chromophore in the SDS-denatured state, refolding of bO was initiated and the extent of labelling of this refolded conformation examined[210]. In refolded bO helix A is protected from oxidation as in refolded bR, however, in helix D some labelling is observed in bO suggesting a role of chromophore binding in the stabilisation of this helix[210]. While this technique has not yet been widely applied to membrane proteins, the work on bR demonstrates the potential of this method to probe structural changes within membrane proteins.

1.8.2 PagP

1.8.2.1 Structure and Catalytic Activity

PagP is a palmitoyl transferase enzyme, which transfers a palmitate chain from a phospholipid to hexa-acylated lipid A (Figure 1.23) and helps to reinforce the structure of the OM[230]. Additionally, modification of lipid A has been shown to help pathogenic bacteria evade recognition by the host immune system and attenuate the inflammatory response to endotoxin[230]. PagP is an OMP from *E. coli* that forms an 8-stranded β-barrel with a 19-residue amphipathic α-helix at its N-terminus (Figure 1.3a)[49; 231; 232]. The structure of PagP was originally solved by NMR spectroscopy in dodecylphosphocholine/n-octyl-β-D-glucoside detergent micelles showing a well-defined β-barrel with structural flexibility of the extracellular loops, particularly L1[231]. The position of the N-terminal α-helix could not be defined in this structure[231]. The subsequent crystallisation of PagP in LDAO micelles determined that the α-helix lies along the interface of the membrane with the aqueous environment and is packed against the β-barrel[49]. In both structures, the β-barrel is tilted by approximately 25° to the membrane normal[49; 231], a membrane position stabilised by the interactions of the aromatic girdles with the membrane intrafacial region (Figure 1.24a)[233].

In PagP homologues, the residues histidine-33, aspartate-76 and serine-77 are absolutely conserved and have been shown to be important for catalytic activity[231]. These residues are found at the extracellular face of the PagP β-barrel (Figure 1.24b) where they are positioned to interact with the polar head-groups of LPS molecules in the outer leaflet of the OM[231]. In both the NMR and crystal structures this region is not well resolved due to high conformational flexibility[49; 231], however, upon lowering the temperature in the NMR spectroscopy experiments a second conformer of PagP was observed in which this region becomes more structurally defined[234]. It was postulated that switching between these two states may be a feature of the catalytic cycle of PagP *in vivo*[234].

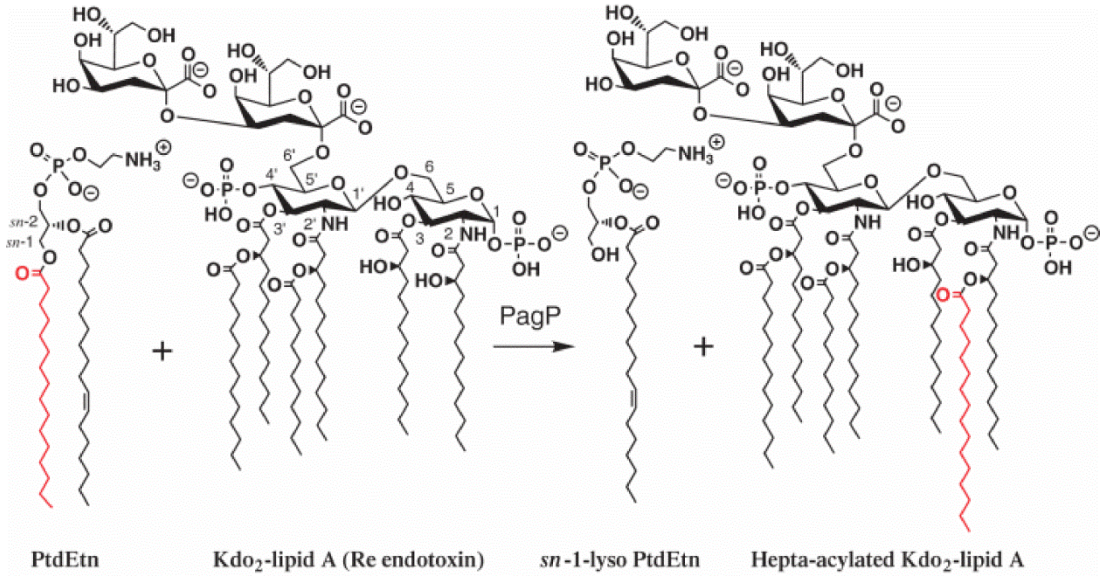

Figure 1.23 Transfer of a palmitate chain to lipid A by the acyltransferase enzyme, PagP. A palmitate chain is transferred from the *sn*-1 position of a phospholipid (phosphoethanolamine (PtdEtn) in the above depiction) to Re endotoxin, the simplest form of LPS in *E. coli*. The transferred chain is shown in red. This figure was taken from Bishop (2005)[230].

The interior of the PagP β-barrel is unusual in that the upper half of the barrel is lined with hydrophobic residues and water molecules are completely excluded while the lower half is lined with polar side chains in a more typical manner[49]. In both crystal structures reported to date, the hydrophobic pocket of the β-barrel is occupied by a single detergent molecule with the polar head-group facing the extracellular environment and the acyl tail nestled in the barrel (Figure 1.24b)[49; 232]. The hypothesis that this hydrophobic pocket represented the palmitate transfer site was confirmed by the use of LDAO to inhibit the catalytic activity of PagP[49]. In contrast, the inclusion of bulky detergents did not affect PagP catalytic activity, presumably as they cannot fit into the binding pocket[49]. Mutation of glycine-88, which forms the floor of the cavity, to the larger residues cysteine or methionine changed the specificity of PagP to preferentially transfer shorter chain lipids[49]. Using the G88C mutant of PagP, a method was developed to alkylate this residue in a site specific manner so the substrate specificity of PagP could be altered[235]. Using this technique to introduce different carbon chain lengths on cysteine-88, it was possible to create PagP variants which were highly specific for C12, C11 and C10 acyl chains, revealing that the substrate recognition mechanism of PagP is using a "hydrocarbon ruler"[236].

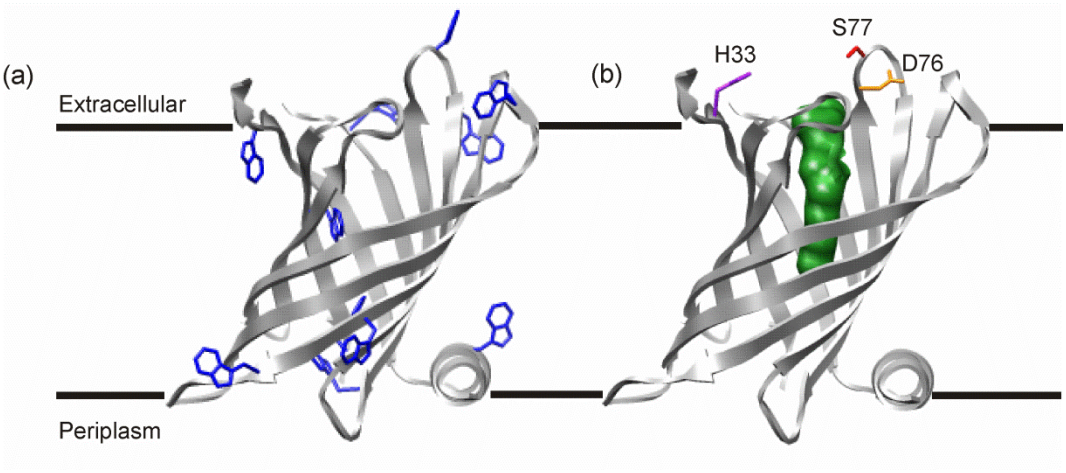

Figure 1.24 Ribbon diagrams of PagP highlighting structural features. (a) Ribbon diagram highlighting the position of tryptophan residues (blue) in PagP and (b) ribbon diagram highlighting the position of the bound LDAO molecule (green space fill) in the centre of the barrel with the conserved residues histidine-33 (purple), aspartate-76 (orange) and serine-77 (red). Side chains are shown using stick representation. The approximate position of the OM is shown by black bold lines. Images were generated from the Protein Data Bank[43] file 1THQ[49] using UCSF Chimera molecular visualisation application[44].

Structural analysis of PagP revealed conserved proline residues between strands F and G and also between strands A and B (Figure 1.25), which serve to disrupt hydrogen-bonding between these strands[49]. It was hypothesised that disruption of the hydrogen-bonding in the upper half of the PagP β-barrel could serve as a possible mechanism by which the substrate phospholipid could enter the active site[49]. MD simulations based on the more recent SDS crystal structure of PagP revealed that phospholipid access to the active site is most likely through the crenel between the F/G strands[232]. Mutation of the proline residues, which disrupt hydrogen-bonding between the F/G strands, to cysteine residues and their subsequent cross-linking was able to prevent the palmitoyl transferase reaction, indicating that this is indeed the site of substrate entry[237].

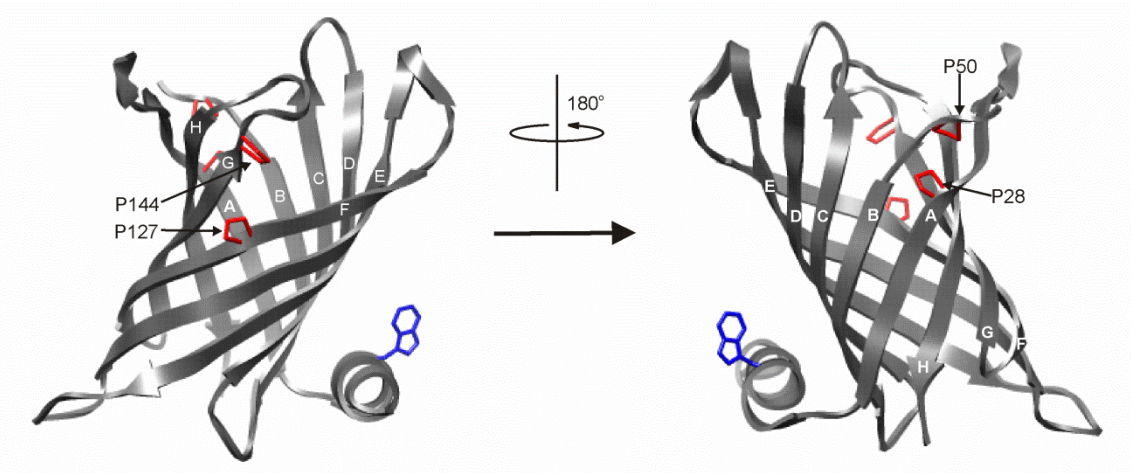

Figure 1.25 Ribbon diagrams of PagP showing conserved proline residues. Conserved proline residues (P28, P50, P127 and P144) are shown in red, while tryptophan-17 is shown in blue. The β-strands are labelled A–H beginning at the N-terminus. Side chains are shown using stick representation. Images were generated from the Protein Data Bank[43] file 1THQ[49] using UCSF Chimera molecular visualisation application[44].

1.8.2.2 Folding Mechanism In Vitro

PagP was found to be able to refold into both detergent micelles and liposomes *in vitro*, however, unlike OmpA, a high concentration of urea (7 M) was required to solubilise the protein and prevent aggregation prior to insertion[145; 238]. Far-UV CD spectroscopy was used to follow the formation of β-sheet structure at 218 nm as well as the formation of tertiary structure *via* a Cotton effect between tyrosine-26 and tryptophan-66 at 232 nm (Figure 1.26)[235]. This showed that secondary and tertiary structure form concomitantly, as had already been shown for OmpA[147]. Mutants of PagP were then created to investigate the function and role in folding of the α-helix, since the possession of a periplasmic α-helix is an unusual feature of an OMP[238]. While deletion of the helix did not prevent the refolding of the protein, analysis of the thermal stability of the PagPΔ(1-19) mutant and other mutants in which key interactions between the helix and β-barrel were deleted was carried out[238]. This study revealed that the helix greatly increases the stability of the folded protein in liposomes but this effect was not so pronounced in detergent, which illustrates the importance of the lipid bilayer to the stability of the native protein[238]. One mutant, in which the conserved residue tryptophan-17 (Figure 1.25) was replaced with alanine, unfolded fifty times more rapidly that the wild-type[238]. Moreover, in the PagPΔ1-19 construct, mutation of arginine-59 to tryptophan restored wild-type kinetics and stability in liposomes,

demonstrating the importance of the aromatic "girdle" which has been observed in OMPs and was speculated to help anchor these proteins into the membrane[45; 238; 239].

By systematically varying the protein concentration and studying folding under a range of lipid-to-protein ratios (LPRs), conditions were established under which the PagP unfolding transition is completely reversible in *di*$C_{12:0}$PC LUVs[167]. Equilibrium and kinetic analysis of PagP revealed that the protein folds *via* a two-state mechanism over the range of urea concentrations studied (7.8–10 M) (Figure 1.27), however, refolding kinetics at urea concentrations below 7.8 M showed two exponential phases[167]. Irrespective of the final urea concentration, the exponential phase was always preceded by a burst-phase in the dead-time (≈30 s) of the experiment[167]. A Φ-value analysis was undertaken on 19 mutants of PagP (Figure 1.28). The PagP variants all showed two-state folding at the urea concentrations used, and suggested a polarised transition state since mutations in the N-terminal half of the protein increased the unfolding rate constants more than mutations in the C-terminal half of the protein[167]. Residues with the highest Φ-values were observed in the C-terminal half of the protein, with lower Φ-values in the N-terminal half (Figure 1.28a)[167]. This suggests that the C-terminal residues insert into the membrane first and begin to fold, while the N-terminal region is still largely unstructured in the transition state (Figure 1.28b). Two negative Φ-values were observed, providing evidence for stabilisation of the transition state by non-native interactions[167]. This mechanism of tilted insertion is consistent with the mechanism of concerted folding and insertion suggested for OmpA[145].

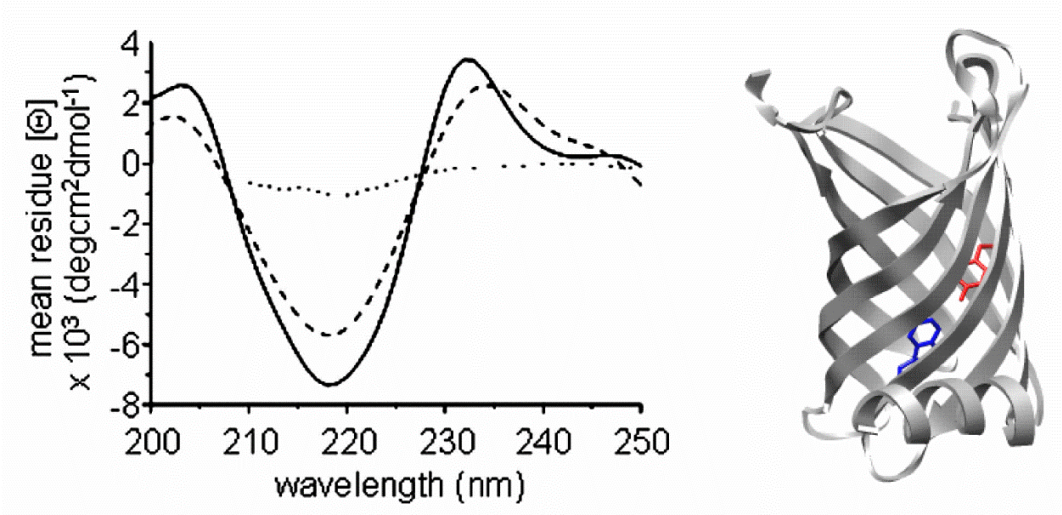

Figure 1.26 The Cotton effect observed in the far-UV CD spectrum of PagP. Far-UV CD spectra (left) of PagP folded in cyclofos-7 micelles (solid line) and *di*$C_{12:0}$PC LUVs (dashed line). The spectrum of PagP unfolded in 8 M urea is also shown (dotted line). The Cotton effect is seen as a band of positive ellipticity at 232 nm and arises from interaction of tyrosine-26 and tryptophan-66. The ribbon diagram (right) of PagP shows tyrosine-26 (red) and tryptophan-66 (blue) using stick representation. The ribbon diagram was generated from the Protein Data Bank[43] file 1THQ[49] using UCSF Chimera molecular visualisation application[44]. The far-UV CD spectra are reproduced from Huysmans *et al* (2007)[238].

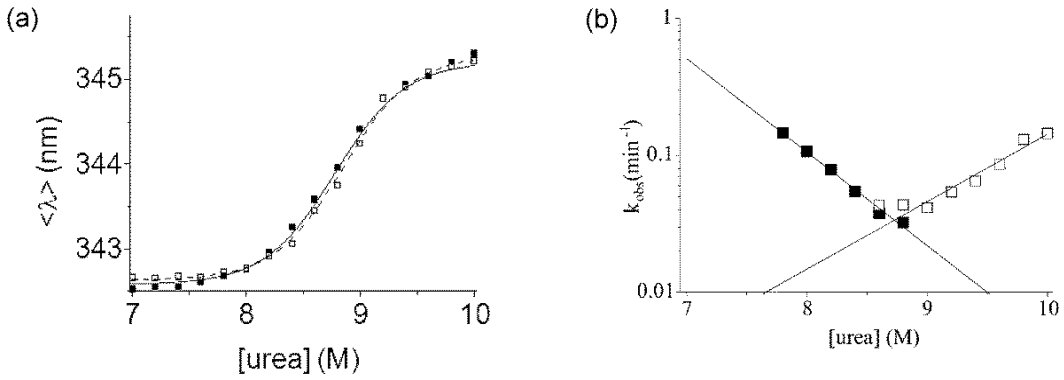

Figure 1.27 Reversible two-state unfolding of PagP. (a) Equilibrium folding (solid squares) and unfolding (open squares) of PagP. Solid and dotted black lines represent fits in the folding and unfolding direction, respectively. (b) Urea dependence of the folding (filled symbols) and unfolding (open symbols) rate constants of PagP. In both graphs, solid lines indicate appropriate fits to the data. All experiments contained 0.4 μM PagP in 100 nm *di*$C_{12:0}$PC LUVs at an LPR of 3200:1 and were performed in 50 mM sodium phosphate buffer pH 8 at 25 °C. Adapted from Huysmans (2008)[239] and Huysmans *et al* (2010)[167].

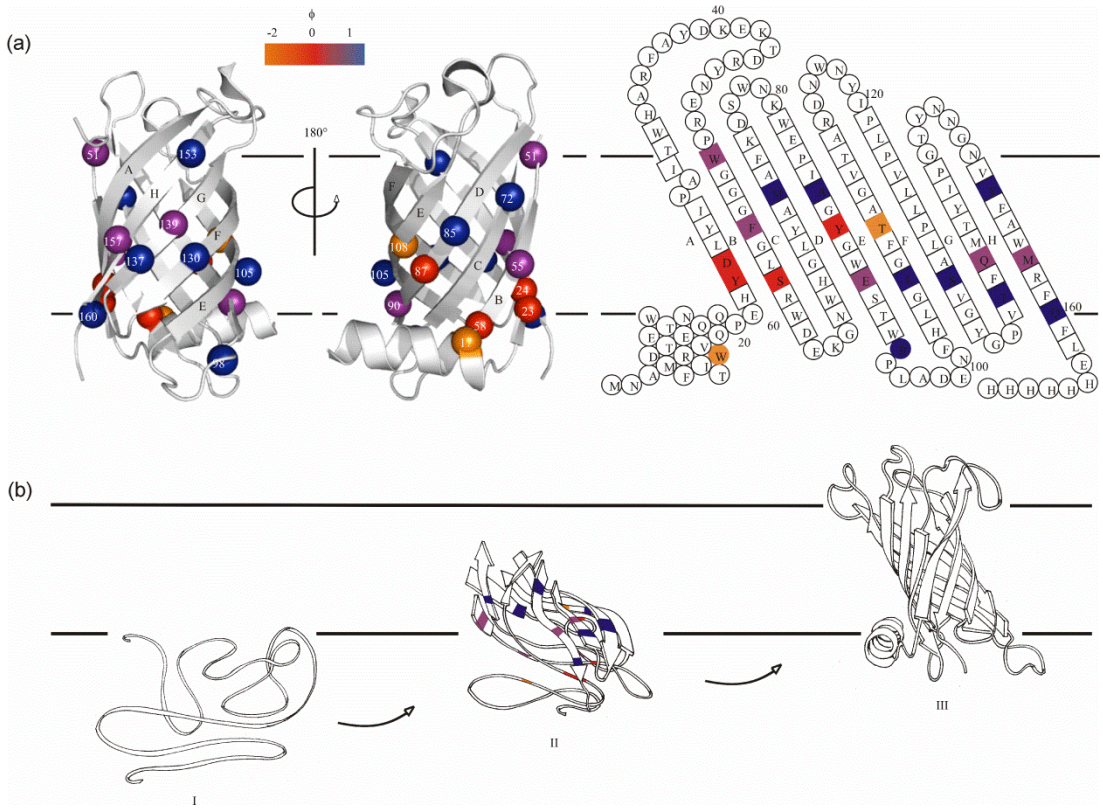

Figure 1.28 Φ-value analysis of PagP. Φ_F-values determined from kinetic analysis of PagP variants are (a) mapped onto a ribbon diagram (left) and a topology model (right) of the native structure. (b) The proposed mechanism for PagP folding into membranes. Folding is initiated from an unfolded membrane associated state (I) and proceeds *via* a tilted insertion mechanism through a highly polarised transition state (II), before assembly of the N-terminal α-helix yields native PagP (III). $\Phi_F > 0.5$ are shown in blue, $0.3 \leq \Phi_F \geq 0.5$ are in purple, $\Phi_F < 0.3$ are in red and $\Phi_F < 0$ in orange. This figure is taken from Huysmans *et al* (2010)[167].

Further kinetic analysis of PagP folding into 100 nm *di*$C_{12:0}$PC LUVs revealed that lowering the LPR to 800:1 or lowering the PagP concentration to 0.1 μM resulted in a burst phase followed by an exponential phase that was no longer well described by a single exponential function, but a second slower exponential phase was required to fit the data satisfactorily[240]. Interrupted refolding experiments, where folding is allowed to proceed for varying times before the initiation of unfolding, were carried out revealing that the rate of formation of the native state was best described by two rate constants indicative of parallel folding pathways[240]. Introduction of up to 40 % *di*$C_{12:0}$PS at an LPR of 3200:1 and PagP concentration of 4 μM (i.e. under conditions where the exponential folding phase is well described by a single exponential function) showed that *di*$C_{12:0}$PS does not alter the folding pathway of PagP[240]. By contrast, introduction of as little as 5 % *di*$C_{12:0}$PE under the same conditions caused a shift to an exponential

phase best described by a double exponential function, suggesting that this lipid can shift the preferred folding pathway of PagP[240]. The small size of the PE head-group in comparison to PC or PS head-groups (Figure 1.4) results in increased lateral pressure within the bilayer upon inclusion of PE lipids, slowing insertion of PagP[240]. While inclusion of *di*$C_{12:0}$PS did not alter the folding pathway of PagP, electrostatic effects due to the presence of this negatively charged lipid lead to smaller folding rate constants in the presence of this lipid[240]. Overall, the lipid properties were shown to be important in modulating the folding of PagP between two parallel pathways. Since the unfolding kinetics of PagP were well described by single exponential functions under all conditions tested, there was insufficient evidence for an alternative native-like state of PagP[240]. Rather, the slower pathway is most likely due to a second population of folding-competent PagP in solution ready to adsorb on to the lipid surface upon exposure of free membrane surface (Figure 1.29)[240].

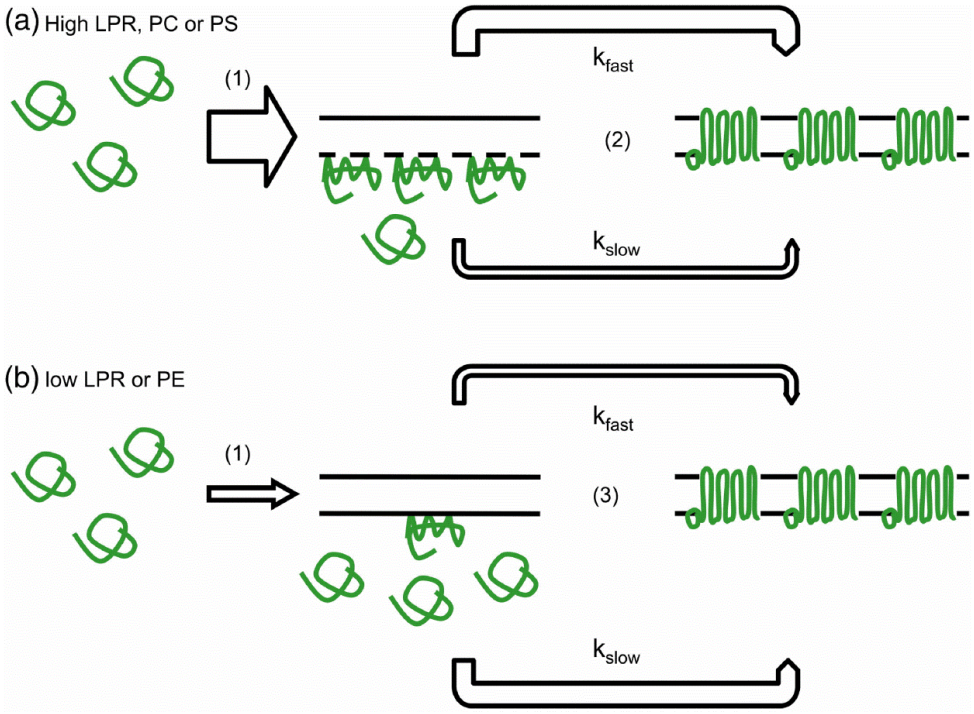

Figure 1.29 Schematic representation of the parallel folding pathways of PagP. (1) Folding is initiated by rapid adsorption on to the membrane surface. (a) The amplitude of this process is increased by high lipid-to-protein ratio (LPR) and the presence of *di*$C_{12:0}$PC and *di*$C_{12:0}$PS lipids, leading to a (2) predominantly fast folding route. (b) The amplitude of the lipid adsorption phase is reduced by low LPR or the presence of *di*$C_{12:0}$PE lipid, leading to folding by a (3) predominantly slow route. Inclusion of *di*$C_{12:0}$PS does not change the preferred folding route, but does affect the rate of membrane insertion. The thickness of the arrows indicates the relative amplitude of the folding phases. This figure was taken from Huysmans *et al* (2012)[240].

1.9 AIMS OF THIS THESIS

The overall aim of this study is to build on the previous work in the laboratory towards understand the folding mechanism of β-barrel OMPs. This problem was approached in two ways: firstly, the development of a new model folding system for OMP study with the aim of using structurally related proteins to look for generic folding rules and secondly, to continue the investigation into PagP folding by investigating the effects of periplasmic chaperones on the folding pathway.

To establish a new model system for OMP folding the OM proteases, OmpT and OmpP, of *E. coli* were chosen. Although structurally related to established OMP folding models, OmpA (see Section 1.6) and PagP (see Section 1.8.2), a Basic Local Alignment Search Tool (BLAST) alignment shows there is only ≈20 % sequence similarilty between OmpA and OmpT and no sequence similarity between OmpT and PagP[241]. The crystal structure of OmpT has been solved[51], allowing the obtained results to be structurally characterised and the previous establishment of an enzyme activity assay for OmpT allows the presence of the native state to be unequivocally detected[138; 242; 243]. While OmpP has not been crystallised, its high sequence similarity (87 %) to OmpT[244] allows a homology model to be created and thus structural rationalisation of results is possible. Additionally, the substrate specificity of OmpP is known[245] allowing activity assays to be carried out to detect the native state. The work described in Chapter 3 describes experiments that were used to clone OmpT and OmpP from *E. coli* DNA and establish protocols for the over-expression, purification and refolding of these proteins. The effect of acyl chain length and saturation on stability is investigated, revealing the ability of these proteins to assume the native state in spite of hydrophobic mis-match with the bilayer.

A recurring problem in OMP folding studies is the establishment of conditions under which completely reversible folding can be achieved[150; 165]. The work described in Chapter 4 shows that under many conditions, the unfolding and refolding denaturation curves of OmpT display hysteresis. By kinetic examination of known reversible and irreversible folding conditions of OmpT, C-terminally His-tagged PagP[167; 238; 240] and the untagged PagP construct reported by Burgess *et al*[162] a possible mechanism is revealed by which this phenomenon could arise.

Finally, the work described in Chapter 5 focuses on the effect of periplasmic chaperones on OMP folding. Burgess *et al*[162] previously reported an untagged PagP construct which remains soluble and folding-competent in as little as 1 M urea, in stark contrast with the need to use urea concentrations as high as 7 M to study the folding pathway of his-tagged PagP[167; 238; 240]. Building on previous work in the laboratory to establish purification and folding protocols of this untagged construct[246], a kinetic refolding assay of untagged PagP was developed and used to study the effects of the *E. coli* periplasmic chaperones Skp and SurA. This work lays the foundation for future studies with the aim of fully elucidating the mechanisms by which OMPs are chaperoned across the aqueous periplasm to the OM *in vivo*. The results of the work and findings of the thesis are summarised in Chapter 6 and placed in context of the field. Future avenues of research are pointed out that address key questions in this exciting and rapidly developing field.

2 MATERIALS AND METHODS

2.1 MATERIALS AND REAGENTS

2.1.1 GENERAL CHEMICALS

Greater than 15 MΩ resistance water (dH_2O) (Purite) was used in all protocols. Isopropyl β-D-1-thiogalactopyranoside (IPTG), agarose, bacto-agar, lysogeny broth (LB) granulated medium, carbenicillin disodium salt, Tris (hydroxymethyl)-aminomethane (Tris) and GuHCl were purchased from Melford Laboratories, UK. Granulated super optimal broth (SOB) was purchased from Merck, UK. Ultrapure urea was purchased from MP Biomedicals, UK. Ethylenediaminetetraacetic acid (EDTA), glycerol, sodium chloride (NaCl), sodium dodecyl sulphate (SDS), acetic acid, glucose and ethanol were purchased from Fisher Scientific, UK. 30 % (*w/v*) acrylamide: 0.8 % (*w/v*) bis-acrylamide was purchased from Severn Biotech Ltd, UK. Triton X-100 protein grade detergent was purchased from Calbiochem, UK. The fluorogenic peptide, *ortho*-aminobenzoic acid (Abz)-Ala-Arg-Arg-Ala-Tyr(NO_2)-NH_2 was purchased from Cambridge peptides, UK. Kanamycin monosulphate was purchased from ForMedium, UK. Polymyxin B sulphate was purchased from VWR International, UK. All lipids were purchased from Avanti Polar Lipids, Alabama, USA; except lipopolysaccharide (LPS, from *E. coli* strain 0111:B4) which was purchased from Sigma, UK. All other chemicals were purchased from Sigma-Aldrich, UK and were of analytical grade.

2.1.2 MOLECULAR BIOLOGY MATERIALS

Oligonucleotide primers were purchased from Eurofins MWG Operon, UK. Vent polymerase, Taq polymerase, *Nde*I, *Bam*HI and Quick Ligase enzymes, as well as bovine serum albumin (BSA) were purchased from New England Biolabs, USA. The pGEM-T Easy Vector System, 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-Gal), deoxyribonucleotide triphosphates (dNTPs), 6 × Blue/Orange Loading Dye and DNA markers for electrophoresis were purchased from Promega, UK. The

QIAprep Spin Miniprep Kit and QIAquick Gel Extraction kit were purchased from QIAGEN, UK. *E. coli* strains XL1-Blue (competent and supercompetent) and BL21 (DE3) competent were purchased from Stratagene, UK.

Plasmid pET11a (Novagen) containing the gene encoding the mature form of PagP[162] was kindly provided by Dr. Karen Fleming (T.C. Jenkins Department of Biophysics, Johns Hopkins University, USA).

Plasmid pETCrcAHΔS, which contains the gene encoding PagP with a C-terminal hexa-histidine tag in a plasmid derived from pET21a^{+} [166], was a generous gift from Prof. Russell Bishop (Department of Biochemistry and Biomedical Sciences, McMaster University, Canada).

Plasmid pET21b (Novagen) containing the gene encoding Skp was kindly provided by Dr. James Bardwell (Department of Biophysics, University of Michigan, USA).

Plasmid pSK257 containing the gene encoding the mature form of SurA with an N-terminal hexa-histidine tag[137], and a thrombin cleavage site to facilitate its removal following purification, was a kind gift from Prof. Daniel Kahne (Department of Biological Chemistry & Molecular Pharmacology, Harvard University, USA).

2.1.3 Protein Chemistry Materials

SnakeSkin dialysis tubing (3,500 Da molecular weight cut off (MWCO)) was purchased from Thermo Scientific, UK. The HiLoad 26/60 Superdex75 column, nickel Sepharose resin, HiTrap Q column and HiTrap SP column were purchased from GE Healthcare, UK. All buffers used during protein purification were filtered and degassed before use either by vacuum filtration through 0.22 μm filters purchased from Millipore, UK, or for small volumes by filtration 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.

2.2 MOLECULAR BIOLOGY METHODS

2.2.1 BACTERIAL STRAINS

E. coli XL1-Blue (*recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac* [F' *proAB lacl*^{q}*ZΔM15* Tn*10* (Tet^{r})])[247]

E. coli BL21(DE3) (F^{-} *dcm ompT hsdS*^{B} (r_B^-, m_B^-) *gal* λ(DE3))[248]

2.2.2 GROWTH MEDIA

Super optimal broth with catabolite repression (SOC) medium was used to culture XL1-blue supercompetent cells following transformation, and was made by addition of 25 g SOB granulated medium (Merck, UK) to 1 litre of deionised water. The medium was sterilised using an autoclave (121 °C, 15 psi, 20 min). Before use, filter-sterilised glucose and magnesium sulphate were added to the cooled medium to a final concentration of 20 mM.

In all other experiments, bacteria were cultured in lysogeny broth (LB)[249], which was prepared by addition of 25 g of granulated LB (Melford laboratories, UK) to 1 litre of deionised water. The medium was sterilised using an autoclave (121 °C, 15 psi, 20 min). Antibiotics were filter-sterilised and added to the cooled medium as detailed in Table 2.1.

Antibiotic	Dissolve in	Stock Conc^{n} (mg ml^{-1})	Final Conc^{n} (μg ml^{-1})
Tetracycline	EtOH	1	10
Carbenicillin	dH_2O	100	100
Kanamycin	dH_2O	20	50

Table 2.1 Summary of antibiotics used in molecular biology methods. Each antibiotic was dissolved in an appropriate solvent at the concentration (conc^{n}) listed and sterilised before use.

LB-agar plates were made by the addition of 15 g bacto-agar (Melford laboratories, UK) per litre of LB medium prior to autoclaving. Where blue/white colour screening plates were required, X-Gal (80 μg ml^{-1}) and filter-sterilised IPTG (0.5 mM) were added to the cooling agar.

2.2.3 Agarose Gel Electrophoresis

Agarose gel electrophoresis was carried out in Tris-acetate-EDTA (TAE) buffer, which was prepared as a 50 × stock as follows:

Tris	121 g
Acetic Acid (glacial)	28.55 ml
EDTA (0.5 M)	50 ml

The solution was adjusted to pH 8.0 and made up to 500 ml using deionised water.

Gels were made by dissolving 1.3 % (*w/v*) agarose in hot 1 × TAE buffer, and adding ethidium bromide (0.5 μg ml^{-1}) to the cooling solution. The gel was then poured into a 12 × 15 cm gel tray with a comb and allowed to set before use. DNA samples were mixed 5:1 with 6 × Loading Buffer (Promega, UK) before loading. 5 μl of 1 kbp and 100 bp DNA ladders (Promega, UK) mixed with 1 μl 6 × Loading Buffer were also loaded to aid size determination. Samples were resolved at a constant voltage of 90 V in 1 × TAE buffer.

After electrophoresis, gels were visualised using a transilluminator emitting ultra violet (UV) light and photographed using a Syngene InGenius gel documentation system (Syngene, UK).

2.2.4 TRANSFORMATION OF E. COLI

Plasmid DNA (50–200 ng) was added to 100 µl of competent cells. After incubation on ice for 30 min, the mixture was heat shocked at 42 °C for 45 seconds and returned to ice for 2 min. 400 µl of sterile medium (SOC for supercompetent XL1-blue cells, LB for all other cell strains) was added and the transformation mixture incubated at 37 °C, 200 rpm for 60 min. Cells were then plated in 50 µl and 200 µl aliquots onto LB-agar containing the appropriate antibiotic and incubated overnight at 37 °C.

2.2.5 PREPARATION OF PLASMIDS

A 5 ml overnight culture of XL1-blue cells (Stratagene, UK) containing the desired plasmid was grown under antibiotic selection. After the cells were pelleted, plasmid DNA was extracted using the QIAprep Spin Mini-Prep Kit (QIAGEN, UK) according to the manufacturer's instructions. The concentration of plasmid DNA was determined by measuring the absorbance of the sample at 260 nm (A_{260}) in a cuvette with a 1 cm path length. An A_{260} of 1 was assumed to be equivalent to 50 ng μl^{-1} dsDNA.

2.2.6 POLYMERASE CHAIN REACTION (PCR)

The genes coding for the OMPs, OmpT and OmpP, were amplified from the chromosomal DNA of XL1-blue cells (Stratagene, UK). A 5 ml culture of XL1-blue cells was grown in LB medium containing 10 µg ml^{-1} tetracycline (see Section 2.2.2) until an OD_{600} of 0.25 was reached. A 20 µl aliquot was removed and diluted to 1 ml in sterile dH_2O. This sample was boiled for 5 min before incubation on ice for 5 min to create a cell lysate that could be used as the DNA template. Primers were designed to amplify the desired gene, and add an N-terminal hexa-histidine and Tobacco Etch Virus protease (TEVp) cut site (Table 2.2). Polymerase chain reactions (PCR) were set up to amplify the desired gene. A typical reaction contained the following components:

Cell lysate	25 µl

Primers	100 pmol
dNTPs	250 µM
$MgSO_4$	2, 4 or 6 mM
DMSO	1 µl
Vent DNA polymerase (2000 U ml^{-1})	1 U
Vent DNA polymerase buffer	1 ×
Nuclease-free deionised water	to 100 µl

Primer	Sequence (5' to 3')
OmpT forward	GCTACATATGCATCATCACCATCACCACGAAAATCTG TACTTCCAGGGTTCTACCGAGACTTTATCGTTTA
OmpT reverse	GGCGGATCCTTAAAATGTGTACTTAAGACCAG
OmpP forward	GCTACATATGCATCATCACCATCACCACGAAAAT CTGTACTTCCAGGGTTCTGATTTCTTCGGCCCGGA
OmpP reverse	GGCGGATCCTTAAAACGTGTACTTCAGACCG

Table 2.2 Oligonucleotide primers used to amplify the OmpT and OmpP genes from the chromosomal DNA of XL1-blue cells. The restriction enzyme recognition sites are highlighted in red (*Nde*I) and blue (*Bam*HI). Forward primers include a DNA sequence encoding a hexa-histidine tag and a TEVp site (ENLYFQG).

In all experiments, an appropriate negative control sample lacking a DNA template was included. Amplification was then carried out according to the temperature cycle outlined in Table 2.3.

The PCR products were purified by agarose gel electrophoresis (see Section 2.2.3) then extracted using the QIAquick Gel Extraction Kit (QIAGEN, UK) according to the manufacturer's instructions.

PCR using plasmid DNA as a template was carried out according to the same method, however, cell lysate was replaced by 100 ng of the appropriate plasmid and the volume

of water in the reaction was adjusted accordingly. This method was used to amplify the genes encoding OmpT and OmpP from existing plasmids to remove the His-tag and the TEVp cleavage site. The primer sequences used are given in Table 2.4.

Step	Temp (°C)	Time (s)
Initial denaturation	95	300
Denaturation	95	30
Annealing	53	30
Elongation	72	90
Repeat denaturation, annealing and elongation (× 29)		
Final elongation	72	300

Table 2.3 Temperature cycle for a typical PCR reaction.

Primer	Sequence (5' to 3')
OmpT forward	GGAGACACATATACATATGTCTACCGAGACTTTATCGTTTA
OmpT reverse	GGAGACACATATACATATGTCTGATTTCTTCGGCCCGGA
pET11a_rev	TCCTTTCGGGCTTTGTTAGC

Table 2.4 Oligonucleotide primers used to amplify the OmpT and OmpP genes from plasmid DNA. These primers remove the N-terminal his-tag and TEVp cleavage site leaving just the sequence encoding the mature protein. The restriction enzyme recognition sites are highlighted in red (*Nde*I).

2.2.7 Blunt Ended Ligation into a Shuttle Vector

To ensure efficient restriction endonuclease digestion, purified PCR products were ligated directly into the pGEM-T Easy Vector (Promega, UK). PCR products were first A-tailed by incubating the following components at 70 °C for 25 min:

Taq DNA polymerase	5 U
Taq DNA polymerase buffer	1 ×

dATP	200 µM
Purified PCR fragment	to 10 µl

Taq DNA polymerase was used for this reaction as it lacks 3' to 5' exonuclease activity.

The tailed PCR product was then ligated into the pGEM-T Easy Vector. Typically, ligation reactions were incubated overnight at 4 °C and contained the following components:

10 × rapid Ligation buffer for T4 DNA Ligase	1 µl
pGEM-T Easy Vector	50 ng
T4 DNA Ligase (3 U μl^{-1})	3 U
A-tailed PCR product	to 10 µl

A control reaction containing deionised water instead of PCR product was included.

Ligation products were transformed into XL1-blue supercompetent cells (Stratagene, UK) and then plated onto colour screening plates containing 100 µg ml^{-1} carbenicillin (see Sections 2.2.2 and 2.2.4). Single white colonies were picked from the plate and grown overnight (37 °C, 200 rpm) in 5 ml LB medium containing 100 µg ml^{-1} carbenicillin. Plasmid DNA was purified and the concentration determined as described in Section 2.2.5.

2.2.8 Digestion of Plasmid DNA

Restriction digestion reactions were carried out using enzymes and buffers from New England Biolabs, USA (NEB) and typically contained the following components:

Plasmid DNA or purified PCR product	1 µg
10 × NEB buffer 3	3 µl
BSA	100 µg ml^{-1}
*Nde*I (20 U μl^{-1})	20 U

*Bam*HI (20 U μl^{-1})	20 U
Nuclease-free deionised water	to 30 µl

All reactions were accompanied by the appropriate single enzyme and enzyme-free control samples. Reactions were incubated at 37 °C for 30 min, following by enzyme inactivation at 65 °C for 20 min.

Digested DNA was purified by agarose gel electrophoresis (see Section 2.2.3) and then extracted using the QIAquick Gel Extraction Kit (QIAGEN, UK) according to the manufacturer's instructions.

2.2.9 Ligation of Digested DNA

Ligation reactions were carried out using the NEB Quick Ligation kit according to the following scheme:

Digested pET11a vector	100 ng
2 × Quick Ligation Buffer	10 µl
Quick T4 DNA Ligase (2000 U μl^{-1})	2000 U
Digested insert	to 20 µl

Appropriate control reactions containing nuclease-free deionised water instead of a digested insert were carried out. The pET11a vector was not dephosphorylated before ligation. Reactions were incubated at 25 °C for 15 min before chilling on ice prior to transformation into XL1-blue supercompetent cells (see Section 2.2.4). Plasmid DNA was purified according to the method in Section 2.2.5.

2.2.10 DNA SEQUENCING

DNA sequencing was carried out by Beckman Coulter Genomics, UK. Sequencing of genes inserted into pET vectors was carried out using the T7 promoter and T7 terminator primers. The T7 promoter primer alone was used for sequencing of pGEM-T Easy vectors.

2.2.11 SUMMARY OF CREATED PLASMIDS

The plasmids created for use in this thesis are described in Table 2.5, and were made using the techniques described in the rest of Section 2.2.

Name	Vector backbone	Insert	PCR Template used
pLMM01	pET11a	Mature OmpT with N-terminal His-tag and TEVp site	XL1-blue chromosomal DNA
pLMM02	pET11a	Mature OmpP with N-terminal His-tag and TEVp site	XL1-blue chromosomal DNA
pLMM03	pET11a	Mature OmpT	Plasmid pLMM01
pLMM04	pET11a	Mature OmpP	Plasmid pLMM02

Table 2.5 Plasmids created for use in this thesis

2.3 PROTEIN METHODS

2.3.1 SODIUM DODECYL SULPHATE POLYACRYLAMIDE GEL ELECTROPHORESIS (SDS-PAGE)

Tris-tricine buffered SDS-PAGE gels (Table 2.6) were used to monitor over-expression, purification and refolding of recombinant OMPs.

Protein samples were diluted two-fold in 2 × SDS loading buffer (50 mM Tris-HCl pH 6.8, 2 % (*w/v*) SDS, 0.1 % (*w/v*) bromophenol blue, 10 % (*w/v*) glycerol) and boiled

(5 min) prior to loading. Mark12™ Unstained Protein Standards (Invitrogen, UK) or 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 electrophoreised with the inner reservoir of the gel tank buffered with cathode buffer (100 mM Tris, 100 mM tricine, 0.1 % (*w/v*) SDS, pH 8.25) and the outer reservoir buffered with anode buffer (200 mM Tris-HCl, pH 8.9). A constant current of 30 mA was applied until the samples entered the resolving gel, and then the current was adjusted to 65 mA until the dye front reached the bottom of the gel. Gels were stained using Instant Blue stain (Expedeon, UK) and photographed using a Syngene InGenius gel documentation system (Syngene, UK).

Solution Component	Volume Added to Resolving Gel (ml)	Volume Added to Stacking Gel (ml)
30 % (*w/v*) acrylamide : 0.8 % (*w/v*) bis-acrylamide	7.50	0.83
3 M Tris-HCl, 0.3 % (*w/v*) SDS, pH 8.45	5.00	1.55
dH_2O	0.44	3.72
Glycerol	2.00	-
10 % (*w/v*) ammonium persulphate	0.05	0.10
Tetramethylethylenediamine (TEMED)	0.01	0.01

Table 2.6 Components of a Tris-tricine buffered SDS-PAGE gel. The volumes stated allow casting of two mini-gels (8 cm × 10 cm) using a 1.5 mm spacer.

2.3.2 Trichloroacetic acid (TCA) Precipitation

Trichloroacetic acid (TCA) precipitation was used to remove GuHCl from protein samples to facilitate analysis by SDS-PAGE. An equal volume of 12.5 % TCA was added to the sample, which was then incubated on ice for 20 min. The precipitated protein was collected by centrifugation in a microfuge (13000 rpm, 10 min, 4 °C) and the supernatant was discarded. The pellet was washed with 1 ml ice cold EtOH and collected by centrifugation as before. The supernatant was again discarded and any residual solvent was removed from the protein pellet by drying in a heat block at 95 °C.

The dried pellet was resuspended in SDS-PAGE loading buffer (see Section 2.3.1 for recipe) and boiled for 5 minutes before loading onto a gel.

2.3.3 WESTERN BLOTTING

Following SDS-PAGE, proteins were transferred onto polyvinylidene fluoride (PVDF) transfer membrane (GE Healthcare, UK) at 12 V for 1 h using a Trans-blot semi-dry electrophoretic transfer cell (Bio-Rad, UK) in transfer buffer (15 mM Tris-HCl, 150 mM glycine, 0.02 % (*w/v*) SDS, 20 % (*v/v*) methanol). The membrane was blocked with 2 % (*w/v*) skimmed milk powder (Marvel) in phosphate buffered saline (PBS: 137 mM NaCl, 2.7 mM potassium chloride, 10 mM dibasic sodium phosphate, 2 mM monobasic potassium phosphate, pH 7.4) containing 0.2 % (*v/v*) Tween-20 for 1 h at room temperature with shaking. The membrane was incubated with 0.2 μg ml^{-1} mouse anti-histidine tag monoclonal antibody (AbD Serotec, UK) with 2 % skimmed milk powder in PBS containing 0.2 % (*v/v*) Tween-20 for 1 h at room temperature with shaking. The membrane was then washed with 5 ml 2 % skimmed milk powder in PBS containing 0.2 % (*v/v*) Tween-20 three times before incubation with 0.2 μg ml^{-1} horseradish peroxidise-conjugated rat anti-mouse IgM monoclonal antibody (BD Pharmingen, UK) in 2 % skimmed milk powder in PBS containing 0.2 % (*v/v*) Tween-20 for 1 h at room temperature with shaking. The membrane was washed three times with 5 ml 2 % skimmed milk powder in PBS containing 0.2 % (*v/v*) Tween-20. Antibody binding was visualised using SuperSignal West Pico Chemiluminescent Substrate (Thermo Scientific) and Amersham Hyperfilm ECL (GE Healthcare) and developed.

2.4 PROTEIN EXPRESSION AND PURIFICATION

2.4.1 OVER-EXPRESSION OF OUTER MEMBRANE PROTEINS

All of the OMPs used in this thesis were over-expressed using the following method (Burgess *et al*[162]). BL21 (DE3) cells were transformed with the relevant plasmid (see Table 2.7) as described in Section 2.2.4. A single colony from a fresh transformation was used to inoculate 5 ml LB medium containing 100 μg ml^{-1} carbenicillin and incubated overnight at 37 °C with shaking at 200 rpm. The overnight culture was used to inoculate 0.5 L LB containing 100 μg ml^{-1} carbenicillin. This culture was incubated at 37 °C, 200 rpm until an approximate OD_{600} of 0.6 was reached, when protein expression was induced by addition of filter-sterilised IPTG to a final concentration of 1 mM. Cells were harvested by centrifugation (10000 *g*, 15 min, 4 °C) 4 h after induction.

Plasmid Name	Insert	Vector Backbone
pETCrcAHΔS[166]	His-tagged PagP	pET21a+
pET11a-PagP[162]	Untagged PagP	pET11a
pLMM01	His-tagged OmpT	pET11a
pLMM02	His-tagged OmpP	pET11a
pLMM03	Untagged OmpT	pET11a
pLMM04	Untagged OmpP	pET11a

Table 2.7 Plasmids used in the over-expression of OMPs in BL21 (DE3) cells.

2.4.2 ISOLATION OF OUTER MEMBRANE PROTEIN INCLUSION BODIES

Cell pellets were resuspended in 50 mM Tris-HCl, 5 mM EDTA, 1 mM phenylmethanesulfonylfluoride (PMSF), pH 8.0 and lysed by sonication (6 × 1 min, using a W-225R sonicator (Ultrasonics, Inc)). The insoluble fraction was pelleted by centrifugation (25000 *g*, 30 minutes, 4 ºC). The pellet from a 500 ml culture was resuspended in 20 ml 50 mM Tris-HCl, pH 8.0 containing 2 % (*v/v*) Triton X-100 and stirred at room temperature for 1 hour to dissolve the membranes. The inclusion bodies

were then pelleted by centrifugation as before and resuspended in 50 mM Tris-HCl, pH 8.0. The resuspended pellet was left stirring for 1 hour at room temperature to ensure removal of residual detergent. The wash step in 50 mM Tris-HCl, pH 8.0 was repeated twice, pelleting the inclusion bodies by centrifugation after each resuspension.

2.4.3 Purification of His-Tagged Constructs of OmpT and OmpP

Inclusion bodies of HT OmpT and HT OmpP were solubilised in 10 ml of 6 M GuHCl, 50 mM sodium phosphate, 250 mM NaCl, pH 8.0 at room temperature until fully dissolved and the resulting solution spun down to remove aggregates (20 min, 25000 *g*, 4 °C). Nickel Sepharose resin (GE Healthcare, UK) was pre-equilibrated with 6 M GuHCl, 50 mM sodium phosphate, 250 mM NaCl, pH 8.0 at room temperature, mixed with the dissolved inclusion bodies and incubated at room temperature for 2 h. The resin was pelleted gently (500 *g*, 1 min) and washed with 25 ml 6 M GuHCl, 50 mM sodium phosphate, 250 mM NaCl, 5 mM imidazole, pH 8.0 for 1 h at room temperature. The resin was then packed into an empty PD-10 column (Bio-Rad, USA) and the His-tagged protein eluted using 10 ml 6 M GuHCl, 50 mM sodium phosphate, 250 mM NaCl, 250 mM imidazole, pH 8.0. Purified HT OmpT/P was precipitated by dialysis (using a membrane with an MWCO of 3.5 kDa, Medicell International Ltd, UK) against deionised water (1 h, room temperature). Precipitated protein was pelleted (20 min, 25000 *g*, 4 °C) and then dissolved in 6 M GuHCl, 25 mM Tris-HCl for storage in 200 µl aliquots at a protein concentration of approximately 500 µM, at -80 °C.

2.4.4 Purification of His-Tagged PagP

Inclusion bodies of His-tagged PagP (HT PagP) were solubilised in 10 ml of 6 M GuHCl, 10 mM Tris-HCl, 250 mM NaCl, pH 8.0 at room temperature and the resulting solution spun down to remove aggregates (25000 *g*, 20 min, 4 °C). Nickel Sepharose resin (GE Healthcare, UK) were pre-equilibrated with 6 M GuHCl, 10 mM Tris-HCl, 250 mM NaCl, 5 mM imidazole, pH 8.0 at room temperature, mixed with the dissolved inclusion bodies and incubated at room temperature for 2 h. The nickel Sepharose resin

was pelleted gently (500 *g*, 1 min) and washed with 25 ml 6 M GuHCl, 10 mM Tris-HCl, 250 mM NaCl, 20 mM imidazole, pH 8.0 for 1 h at room temperature. The nickel Sepharose resin was then packed into an empty PD-10 column (Bio-Rad, USA) and the HT PagP eluted using 10 ml 6 M GuHCl, 10 mM Tris-HCl, 250 mM NaCl, 250 mM imidazole, pH 8.0. Purified HT PagP was precipitated by dialysis using a membrane with an MWCO of 3.5 kDa against deionised water (1 h, room temperature). Precipitated protein was pelleted (25000 *g*, 20 min, 4 °C) and then dissolved in 6 M GuHCl, 25 mM Tris-HCl for storage in 50 µl aliquots at a protein concentration of approximately 500 µM, at -80 °C.

2.4.5 Purification of Untagged Outer Membrane Proteins

Inclusion bodies from a 0.5 L culture were solubilised in 10 ml 6 M GuHCl, 25 mM Tris-HCl, pH 8.0. Solubilised inclusion bodies were then centrifuged (20000 *g*, 20 min, 4 ºC). After filtration through a 0.2 µM syringe filter, the OMP solution was further purified by gel filtration using a Superdex 75 HiLoad 26/60 column (GE Healthcare) on the ÄKTA Prime chromatography system. The column was washed with two column volumes of dH_2O and equilibrated by washing with two column volumes of 6 M GuHCl, 25 mM Tris-HCl, pH 8.0. For each gel filtration run, 5 ml of solubilised protein was loaded into a 5 ml loop at a concentration of up to 10 mg ml^{-1} and eluted using the program detailed in Table 2.8. Following gel filtration, OMP containing fractions were concentrated to approximately 500 µM using Vivaspin 20 concentrators (MWCO 10 kDa). Purity was assessed by SDS PAGE (see Section 2.3.1) following TCA precipitation (see Section 2.3.2). OMPs were stored in 50 µl aliquots at -80 °C.

Breakpoint (ml)	Flow rate (ml min^{-1})	Fraction size (ml)	Injection Valve Position	Auto zero
0	2	0	Load	No
10	2	0	Inject	Yes
20	2	0	Load	No
90	2	3	Load	No
320	2	0	Load	No

Table 2.8 ÄKTA programme parameters for purification of untagged 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).

2.4.6 Expression and Purification of SurA

Plasmid pSK257 containing the mature form of SurA with an N-terminal His-tag was kindly provided by Daniel Kahne (Harvard University)[137; 250] . Over expression and purification of SurA followed the method of Hagan *et al*[137]. The plasmid pET28b was transformed into BL21 (DE3) cells (Stratagene, UK) as described in Section 2.2.4. A single colony from a fresh transformation was used to inoculate 125 ml LB medium containing 50 μg ml^{-1} kanamycin and incubated overnight at 37 °C with shaking at 200 rpm. The overnight culture was used to inoculate LB, containing 50 μg ml^{-1} kanamycin, at a ratio of 1:100. This culture was incubated at 37 °C, 200 rpm until an approximate OD_{600} of 1.0 was reached. The temperature was reduced to 16 °C and protein expression was induced by addition of filter-sterilised IPTG solution to a final concentration of 0.1 mM. Following overnight expression at 16 °C, 200 rpm, cells were harvested using a Heraeus Contifuge at 15000 rpm (Rotor 8575).

The cell pellet was resuspended in 150 ml 20 mM Tris-HCl, 5 mM imidazole, pH 8.0. Protease inhibitors were excluded as they reduce the efficiency of the CleanCleave kit in subsequent steps. Cells were lysed using a cell disruptor (Constant Cell Disruption Systems) and the insoluble fraction removed by centrifugation (25000 *g*, 4 °C, 1 h). Nickel Sepharose (25 ml, GE Healthcare) packed into a XK 26 casing (GE Healthcare) was equilibrated on the ÄKTA Prime chromatography system with 20 mM Tris-HCl, 5 mM imidazole, pH 8.0 and the cell lysate loaded following filtration through a 0.2 μm

syringe filter. The column was washed with 20 mM Tris-HCl, 10 mM imidazole, pH 8.0 and bound proteins eluted using 20 mM Tris-HCl, 500 mM imidazole, pH 8.0 in 25 % steps according to the program in Table 2.9.

The N-terminal His-tag was removed using a CleanCleave kit (Sigma), following the manufacturer's recommended protocol and any residual His-tagged SurA removed using a second nickel Sepharose pull down. The cleaved SurA was concentrated to ~150 µM using Vivaspin 20 concentrators (MWCO 5 kDa) and then diluted five-fold with Tris buffered saline (TBS: 50 mM Tris-HCl, 150 mM sodum chloride, pH 7.4) to remove residual imidazole, before concentrating again to ~150 µM. Purity was assessed by SDS-PAGE (see Section 2.3.1).

Volume (ml)	Flow Rate (ml min^{-1})	Line	Valve Position	Percentage B (%)	Fraction Volume (ml)	Auto zero	Event Mark
0	10	A1	Load	0	0	No	No
100	5	A2	Load	0	0	Yes	Yes
340	10	A1	Load	0	0	No	Yes
590	10	A1	Load	0	0	No	No
590.1	10	A1	Load	25	5	No	Yes
690	10	A1	Load	25	5	No	No
690.1	10	A1	Load	50	5	No	Yes
790	10	A1	Load	50	5	No	No
790.1	10	A1	Load	75	5	No	Yes
890	10	A1	Load	75	5	No	No
890.1	10	A1	Load	100	5	No	Yes
990	10	A1	Load	100	5	No	No

Table 2.9 ÄKTA programme parameters for purification of His-tagged SurA. Volumes may be adjusted according to the volume of cleared lysate. Nickel Sepharose resin packed into XK 26 column was used for this protocol, along with the following buffers: Line A1 = 10 mM imidazole, 20 mM Tris-HCl, pH 8.0; Line A2 = cleared cell lysate; Line B = 500 mM imidazole, 20 mM Tris-HCl, pH 8.0.

2.4.7 EXPRESSION AND PURIFICATION OF SKP

Plasmid pET21b containing the Skp clone was kindly provided by James Bardwell (University of Michigan). BL21 (DE3) cells (Stratagene) were transformed with pET21b containing the Skp gene as described in Section 2.2.4. A single colony was used to inoculate 125 ml LB containing 100 μg ml^{-1} carbenicillin and incubated overnight at 37 °C with shaking at 200 rpm. The overnight culture was used to inoculate LB, containing 100 μg ml^{-1} carbenicillin, at a ratio of 1:100 and incubated at 27 °C, with shaking. Protein expression was induced by addition of sterile IPTG to give a final concentration of 25 μM once an OD_{600} of approximately 0.6 was reached. The cells were then grown overnight at 27 °C before harvesting using a Heraeus Contifuge at 15000 rpm (Rotor 8575).

The cell pellet was gently resuspended in 50 mM Tris-HCl, 5 mM EDTA, 50 mM NaCl, pH 7.5 at 4 °C using approximately 20 ml buffer per litre of cell culture. Following addition of 1 mg ml^{-1} polymyxin B sulphate, the resuspended cells were incubated at 4 °C for 1 h. The spheroplasts were sedimented by centrifugation (12000 *g*, 20 min 4 °C) and the resulting periplasmic extract was dialysed against 20 mM Tris-HCl, pH 8.0, 100 mM NaCl (Buffer A) overnight at 4 °C.

The periplasmic extract was filtered using 0.2 μm syringe filters and loaded onto a HiTrap Q (5 ml) column equilibrated in Buffer A and then washed with 3 column volumes of Buffer A. Around 80 ml of periplasmic extract was loaded per run and the column washed with 5 column volumes 20 mM Tris-HCl, pH 8.0, 750 mM NaCl (Buffer B) between runs. The flow through from this column, excluding the wash steps in Buffer B, was collected then loaded onto a HiTrap SP (5 ml) column in 100 ml aliquots, washed with 5 column volumes of Buffer A and Skp eluted with a gradient running from 0–100 % Buffer B over 15 column volumes. Purified Skp was dialysed against 20 mM Tris-HCl, pH 8.0 to remove NaCl and concentrated to ~150 μM using Vivaspin 20 concentrators (MWCO 5 kDa). Purity was assessed by SDS-PAGE (see Section 2.3.1).

2.4.8 Determination of Protein Concentration

Protein concentration of purified protein samples was determined by measuring the absorbance of the sample at 280 nm (A_{280}) in a cuvette with a 1 cm path length. The theoretical molar extinction coefficient of each protein was estimated using the ExPASy ProtParam tool[251] and then used to calculate protein concentration according to the Beer-Lambert law. The molar extinction coefficients used to calculate concentrations of proteins in this thesis are provided in Table 2.10.

Protein	ε_{280} (M^{-1} cm^{-1})
HT OmpT	79760
HT OmpP	83770
OmpT	78270
OmpP	82280
HT PagP	82390
PagP	82390
SurA	29450
Skp	1490

Table 2.10 Theoretical molar extinction coefficients of proteins used in this thesis. The theoretical molar extinction coefficient of each protein was estimated using the ExPASy ProtParam tool[251].

2.5 Outer Membrane Protein Folding Assays

2.5.1 Preparation of Liposomes

Appropriate mixtures of lipids were weighed into glass tubes and dissolved in 9:1 chloroform:methanol (*v/v*). The solvent was evaporated under a gentle stream of N_2 to create a thin lipid film, which was further dried *in vacuo* (> 2.5 h). The lipid film was resuspended in an appropriate buffer to give a lipid concentration of 40 mM and allowed to stand at room temperature for 30 minutes. Unilamellar vesicles were created by extruding the lipid suspension through polycarbonate membranes (Nuclepore, Whatman, UK) of either 0.1 μm or 0.05 μm pore size to create LUVs and SUVs,

respectively, using a mini extruder (Avanti, Alabama, USA). The lipids used are summarised in Table 2.11.

Name	Abbreviation	Supplier
1,2-dilauroyl-*sn*-glycero-3-phosphocholine	*di*$C_{12:0}$PC	Avanti
1,2-dimyristoyl-*sn*-glycero-3-phosphocholine	*di*$C_{14:0}$PC	Avanti
1,2-dipalmitoleoyl-*sn*-glycero-3-phosphocholine	*di*$C_{16:1}$PC	Avanti
1,2-dilauroyl-*sn*-glycero-3-phospho-(1'-*rac*-glycerol)	*di*$C_{12:0}$PG	Avanti
1,2-dioleoyl-*sn*-glycero-3-phospho-L-serine-N-(5-dimethylamino-1-naphthalenesulfonyl)	*di*$C_{18:1}$PS-dansyl	Avanti
Lipopolysaccharide from *E. coli* strain 0111:B4	LPS	Sigma

Table 2.11 Summary of lipids used in this thesis.

2.5.2 SDS-PAGE Refolding Assay

The chosen OMP (~500 µM) in 6 M GuHCl was diluted to a final concentration of 4.8 µM in the presence of liposomes at an LPR of 3200:1 and varying concentrations of urea. The liposome stock was prepared as described in Section 2.5.1. Samples were typically allowed to refold for 16 h at either 25 °C or 37 °C before the addition of 2 × SDS-PAGE loading buffer (see Section 2.3.1) to quench the reaction. The samples were then loaded on to an SDS-PAGE gel without boiling to resolve the folded and unfolded populations. Boiled samples were included in the gel to allow identification of the unfolded band. The folded population was estimated by comparing the relative intensities of the folded and unfolded bands by densitometry using GeneTools software (Syngene, UK) and a Syngene InGenius gel documentation system (Syngene, UK).

The following buffers were used to give the desired pH:

- 50 mM citrate, pH 3.8
- 50 mM sodium phosphate, pH 7.0 or pH 8.0
- 50 mM glycine, pH 9.0 or pH 9.5
- 50 mM sodium borate, pH 10.0

2.5.3 Tryptophan Fluorescence Emission Spectra

Spectra were acquired using a Photon Technology International fluorimieter (Ford, West Sussex, UK). Samples typically contained 0.4 μM OMP at an LPR of 3200:1 in 50 mM buffer at the appropriate pH (see Section 2.5.2 for buffers used) and were incubated overnight at either 25 °C or 37 °C before measurement. Spectra were recorded using an excitation wavelength of 280 nm, and recording emission from 300 nm to 400 nm with excitation and emission slit widths set at 3 nm. A step size of 1 nm and resolution time of 1 s were used. Corresponding blank spectra were subtracted for each sample.

2.5.4 Far-UV Circular Dichroism

Far-UV CD spectra were acquired on a Chirascan plus circular dichroism spectrometer (Applied PhotoPhysics) with a bandwidth of 1 nm, scan speed of 20 nm min^{-1}, step size of 1 nm and a pathlength of either 0.1 mm or 1 mm. The average of eight scans was taken to enhance signal to noise. Samples typically contained 10 μM OMP in either 50 mM sodium phosphate buffer, pH 8.0 or 50 mM glycine buffer, pH 9.5. An LPR of 800:1 was used to reduce light scattering. Samples were refolded for 16 h before measurement. Corresponding blank spectra were subtracted for each sample.

2.5.5 OmpT and OmpP Enzyme Activity Assays

5 μM OmpT or OmpP was refolded (16 h, 25 °C) into liposomes (prepared as described in Section 2.5.1) at an LPR of 3200:1 in 50 mM sodium phosphate, 5 mM EDTA, 0.8 mg ml^{-1} LPS, pH 7.0 containing 3 M urea (OmpT) or 5 M urea (OmpP). The folded OMP was diluted into a 50 μM solution of the internally quenched fluorogenic peptide Abz-Ala-Arg-Arg-Ala-Tyr(NO_2)-NH_2 to a final OMP concentration of 40–400 nM. The fluorescence change of the Abz group was monitored over time using a Photon Technology International fluorimeter (Ford, West Sussex, UK) using an excitation

wavelength of 325 nm and an emission wavelength of 430 nm. Both excitation and emission slit widths were fixed at 3 nm. Samples of 50 μM Abz-Ala-Arg-Arg-Ala-Tyr(NO_2)-NH_2 in the presence of unfolded OMP or folded OMP in the absence of LPS were measured as controls. The fluorescence of 50 μM Abz-Ala-Arg-Arg-Ala-Tyr(NO_2)-NH_2 was subtracted from each trace.

Initial enzyme activity was calculated as shown in Equation 2.1 using only the initial, linear part of the fluorescence trace.

$$v_0 = \frac{dI}{dt} \cdot \frac{S}{[OMP]} \qquad \text{Equation 2.1}$$

where v_0 is the initial activity of the enzyme in nmol μM^{-1} min^{-1} (nmol of substrate processed per μM enzyme per minute), $\frac{dI}{dt}$ is the change in relative fluorescence intensity with respect to time (where the relative intensity is the intensity at time *t* divided by the maximum intensity at the endpoint of the reaction), *S* is the total amount of substrate in nmol, *[OMP]* is the concentration of enzyme in μM.

The initial activity of the enzyme was calculated over a range of concentrations then averaged, and the standard error of the mean (SEM, Equation 2.2) determined.

$$SEM = \frac{\sigma}{\sqrt{N}} \qquad \text{Equation 2.2}$$

where σ is the standard deviation (calculated using Equation 2.3) of the rates and *N* is the number of liposome batch replicates.

$$\sigma = \sqrt{\frac{\Sigma_i \, (x_i - \bar{x})^2}{N}} \qquad \text{Equation 2.3}$$

where x_i is the initial activity of the enzyme at each enzyme concentration, $\bar{x}$ is the mean initial activity of the enzyme and *N* is the number of replicates.

2.5.6 Equilibrium Denaturation Assays

To measure equilibrium curves in the unfolding direction, the chosen OMP (~500 μM) in 6 M GuHCl was diluted to 1.2 μM in an appropriate concentration of urea containing liposomes (prepared as described in Section 2.5.1) at an LPR of 3200:1 and refolded overnight at 25 °C. This stock was then diluted to give samples containing 0.4 μM OMP (LPR 3200:1) with final urea concentrations in the range 3–10 M urea, as appropriate, and incubated for a minimum of 8 h at 25 °C before measuring. To measure equilibrium curves in the refolding direction, the chosen OMP (~500 μM) in 6 M GuHCl was diluted to 3.6 μM in an appropriate concentration of urea at an LPR of 3200:1 and refolded overnight at 25 °C. This stock was then diluted to 1.2 μM OMP in 10 M urea at an LPR of 3200:1 and unfolded for a minimum of 8 h at 25 °C. Samples were then prepared from the unfolded stock which contained 0.4 μM OMP (LPR 3200:1) with final urea concentrations in the range 3–10 M urea and incubated overnight at 25 °C before measuring.

The fluorescence emission spectrum of each sample was then measured as described in Section 2.5.3. Corresponding blank spectra were subtracted from each spectrum and the average wavelength (< λ >) of each resulting spectrum was calculated according to Equation 2.4:

$$< \lambda > = \frac{\Sigma_i \lambda_i \cdot I_i}{\Sigma_i I_i} \qquad \text{Equation 2.4}$$

where λ_i is the wavelength and I_i is the fluorescence intensity at that wavelength.

< λ > was calculated over the range i = 320–370 nm in order to abolish the variation in absolute intensity which arises from pipetting liposomes[167]. < λ > does not vary linearly with the fraction of folded protein present in the solution[252], however, it can be normalised using the ratio of the quantum yields of the folded and unfolded conformations over the same range of wavelengths used to calculate < λ > as shown in Equation 2.5[149; 252]:

$$Q_R = \frac{\Sigma_i I_i (F)}{\Sigma_i I_i (U)} \qquad \text{Equation 2.5}$$

where $\Sigma_i I_i (F)$ is the sum of the intensities of the folded state spectrum over the range of wavelengths used to calculate $< \lambda >$, $\Sigma_i I_i (U)$ is the sum of the intensities of the unfolded state spectrum over the range of wavelengths used to calculate $< \lambda >$ and Q_R is the quantum ratio correction.

Where the OMP equilibrium curves in the folding and unfolding directions could be overlaid, indicating folding reversibility, the equilibrium stability and denaturant dependence of the OMP were calculated by fitting to Equation 2.6, which defines a two-state transition including a correction for quantum yield, using Igor Pro (Wavemetrics).

$$S_{obs} = \frac{\frac{1}{Q_R}[(a[D] + b)\exp((\Delta G^{\circ H_2O}_{UN} - M_{UN}[D])/RT)] + (c[D] + d)}{1 + \frac{1}{Q_R}[\exp((\Delta G^{\circ H_2O}_{UN} - M_{UN}[D])/RT)]} \qquad \text{Equation 2.6}$$

where S_{obs} is the observed signal, a and c are the signals of the native and denatured states, respectively, in the absence of denaturant, b and d are the denaturant dependence of the signal of the native and denatured states, respectively, $[D]$ is the denaturant concentration, $\Delta G^{\circ H_2O}_{UN}$ is the free energy of unfolding in the absence of denaturant, M_{UN} is the m-value (which reflects the denaturant dependence of $\Delta G^{\circ H_2O}_{UN}$) and Q_R is the quantum yield ratio as defined in Equation 2.5.

Where curves were not reversible, the apparent denaturation midpoint ($[D]_{50\%\ app}$) was estimated by fitting the unfolding transition to Equation 2.6 and using the obtained apparent values of $\Delta G^{\circ H_2O}_{UN\ app}$ and $M_{UN\ app}$ according to Equation 2.7.

$$\Delta G^{\circ H_2O}_{UN\ app} = M_{UN\ app} \times [D]_{50\%\ app} \qquad \text{Equation 2.7}$$

2.5.7 KINETIC UNFOLDING ASSAYS

Changes in Trp fluorescence emission measured at 335 nm upon excitation at 280 nm were used to monitor OMP unfolding using a Photon Technology International Fluorimeter (Ford, West Sussex, UK) equipped with a thermally controlled four-cell changer. The temperature was maintained using a circulating water bath.

The chosen OMP (~500 µM) in 6 M GuHCl was diluted to 4 µM in buffer containing liposomes (prepared as described in Section 2.5.1) at an LPR of 3200:1 and urea (3 M for OmpT, 7 M for PagP and HT PagP) and allowed to refold overnight. Unfolding was initiated by rapid dilution into buffer containing high concentrations of urea (9–10 M). Samples were mixed manually and contained a final concentration of 0.4 µM OMP, 1.28 mM liposomes (LPR 3200:1) and either 9 M urea (OmpT) or 10 M urea (PagP and HT PagP). Kinetic transients were followed for up to 2 h and the kinetic traces were fitted to a single exponential function (Equation 2.8).

$$y = A \cdot e^{-kt} + c \qquad \text{Equation 2.8}$$

where *A* is the signal change upon unfolding, *k* is the unfolding rate constant and *c* is the fluorescence end point.

If a single exponential function did not satisfactorily fit the data as determined by the residuals of the fit, a double exponential equation was instead used (Equation 2.9).

$$y = \left(A_1 \cdot e^{-k_1 t}\right) + \left(A_2 \cdot e^{-k_2 t}\right) + c \qquad \text{Equation 2.9}$$

where *A* is the signal change upon unfolding, *k* is the unfolding rate constant and *c* is the fluorescence end point.

2.5.8 KINETIC UNFOLDING ASSAYS USING DANSYL-LABELLED LIPID

The chosen OMP (~500 µM) in 6 M GuHCl was diluted to 4 µM in buffer containing liposomes doped with 1 % *di*$C_{18:1}$PS-dansyl (prepared as described in Section 2.5.1) at an LPR of 3200:1 and 7 M urea and allowed to refold overnight. Unfolding was initiated by rapid dilution into buffer containing 10 M urea. Samples were mixed manually and contained a final concentration of 0.4 µM OMP, 1.28 mM liposomes (LPR 3200:1) and 10 M urea.

OMP unfolding was monitored using a Photon Technology International Fluorimeter (Ford, West Sussex, UK) equipped with a thermally controlled four-cell changer. Samples were excited at 280 nm and emission was measured between 300–550 nm at regular intervals for up to 2 h. Trp emission (300–400 nm) was used to follow OMP unfolding and the $< \lambda >$ calculated as described in Equation 2.4. Dansyl emission by FRET from nearby Trp residues was measured between 450–550 nm and used as a measure of the relative distance between the OMP and the liposome surface.

The approximate area under the dansyl emission curve at each time point was estimated by calculating the sum of the fluorescence intensities between 450–550 nm. The area was then normalised according to Equation 2.10 to give a measure of the proportion of FRET occurring relative to that of folded OMP in dansyl-doped liposomes.

$$Normalised\ area = \frac{A_{sample} - A_{blank}}{A_{max} - A_{blank}} \qquad \text{Equation 2.10}$$

where A_{sample} is the area under the curve of the sample, A_{blank} is the area under the curve of the dansyl-doped liposomes in the absence of OMP and A_{max} is the area under the curve of the OMP folded in dansyl-doped liposomes.

2.6 FOLDING ASSAYS IN THE PRESENCE OF PERIPLASMIC CHAPERONES

2.6.1 KINETIC REFOLDING ASSAYS

Changes in Trp fluorescence emission measured at 335 nm upon excitation at 280 nm were used to monitor PagP folding using a Photon Technology International fluorimeter (Ford, West Sussex, UK) equipped with a thermally controlled four-cell changer. The temperature was maintained at 37 °C using a circulating water bath. A stock of 100 µM PagP unfolded in 10 M urea was rapidly diluted into buffer containing liposomes (prepared as described in Section 2.5.1) and urea to initiate folding. Samples were mixed manually and contained a final concentration of 0.4 µM PagP, 1.28 mM liposomes (LPR 3200:1) and varying concentrations of urea. Kinetic transients were followed for up to 2 h. For the kinetic experiments in the presence of periplasmic folding factors or lysozyme, PagP was first incubated in the presence of an excess of chaperone or lysozyme (three-fold molar excess of SurA or lysozyme; or a two-fold molar excess of Skp trimers) for 5 min at room temperature in 50 mM glycine buffer, pH 9.5, 0.24 M urea before being diluted six-fold into 50 mM glycine buffer, pH 9.5 containing liposomes and varying concentrations of urea (2–4 M). Samples contained a final concentration of 0.4 µM PagP, 1.28 mM liposomes (LPR 3200:1). Control experiments were carried out to ensure that this protocol did not alter the observed rate of PagP folding in the absence of Skp or SurA.

For each sample four or more replicate traces were fitted using the global fit package in Igor Pro 6.0 (Wavemetrics) sharing the rate constants. Kinetic traces were fitted to a single exponential function (Equation 2.8), using a negative value for the amplitude (A) to reflect the increase in fluorescence signal upon folding. Each experiment was measured using three different batches of liposomes, and the rate constant obtained from the global fits for each batch averaged. The standard error of the mean was then calculated (Equation 2.2).

2.6.2 SDS-PAGE Analysis of Folding Yield

An unfolded PagP stock was diluted to a concentration of 4 μM in the presence of liposomes (prepared as described in Section 2.5.1) at an LPR of 3200:1 and an excess of chaperone or lysozyme (three-fold molar excess of SurA or lysozyme; or a two-fold molar excess of Skp trimers), in 50 mM glycine buffer, pH 9.5 containing the specified concentration of urea. Typically samples were allowed to refold for 16 h at 37 °C before folding was quenched by the addition of 2 × SDS-PAGE loading buffer (see Section 2.3.1). The samples were then loaded onto a Tris-tricine SDS-PAGE gel without boiling and electrophoreised (see Section 2.3.1)

2.6.3 Analytical Gel Filtration

To analyse the interaction between PagP and SurA, a stock solution of PagP unfolded in 50 mM glycine, pH 9.5 containing 10 M urea was diluted to 10 μM in 50 mM glycine, pH 9.5 in 0.24 M urea and allowed to incubate at room temperature for 5 min in the presence of 60 μM SurA. To analyse the interaction between lysozyme and PagP, 2 μM PagP was added to 6 μM lysozyme in 50 mM glycine, pH 9.5 containing 0.24 M urea and incubated for 5 min at room temperature. In all experiments, 200 μL of the sample was then injected onto a Superdex 75 10/300 GL column (GE Healthcare, UK) equilibrated in 50 mM glycine, pH 9.5 and controlled by an ÄKTA prime purification system. The protein was eluted at a flow rate of 0.5 ml min^{-1} and 0.5 ml fractions were collected for analysis by SDS-PAGE (see Section 2.3.1). Under these conditions, lysozyme interacted with the Superdex resin hence the requirement to use very low protein concentrations, which prevented the analysis of the eluted fractions by SDS-PAGE. Appropriate PagP only, SurA only, and lysozyme only samples were also analysed.

2.6.4 ANS BINDING

Fluorescence emission spectra of 8-anilinonaphthalene-1-sulfonic acid (ANS) were acquired using a Photon Technology International fluorimieter (Ford, West Sussex, UK) in the presence or absence of PagP. 100 μM PagP in 50 mM glycine buffer, pH 9.5 was diluted to a final concentration of 1 μM PagP in buffer containing 250 μM ANS, 0.24 M urea, 50 mM glycine, pH 9.5 and allowed to incubate for 5 min at 25 °C before measurement. The spectrum of 1 μM PagP in buffer containing 250 μM ANS, 10 M urea, 50 mM glycine, pH 9.5 was also measured as a non-binding control, along with the spectra of 250 μM ANS in both buffer conditions in the absence of PagP.

Spectra were acquired by excitation of ANS at 389 nm and measurement of the emission between 400 nm and 600 nm with 1 nm resolution and 1 s integration time. The excitation and emission slit widths were set at 4 nm. Three spectra were recorded for each sample and averaged.

2.6.5 CHAPERONE BINDING TO NICKEL-SEPHAROSE-IMMOBILISED HT PAGP

10 μM HT PagP was bound to nickel Sepharose resin (GE Healthcare) equilibrated in 6 M GuHCl, 250 mM NaCl, 10 mM Tris, pH 8.0 containing 5 mM imidiazole for 2 h at 4 °C. The resin was then washed with deionised water followed by equilibration in 50 mM glycine, pH 9.5. 10 μM lysozyme, SurA or Skp was then added and allowed to bind for 1 h in 50 mM glycine buffer, pH 9.5 containing 5 mM imidazole. The resin was then washed with 20 mM imidazole in 50 mM glycine, pH 9.5. Elution of bound protein was carried out at room temperature using 50 mM glycine, pH 9.5 containing 5 M urea and 500 mM imidazole. The final eluent and the unbound fractions were analysed by SDS-PAGE (see Section 2.3.1).

2.6.6 Skp Holdase Assay

HT PagP (21.6 μM) was incubated in 50 mM glycine, pH 9.5, 1 M urea in the presence of a two-fold molar excess of Skp trimers or a six-fold molar excess of hen egg white lysozyme. After 5 min, the samples were diluted six-fold into *di*$C_{12:0}$PC liposomes (prepared as described in Section 2.5.1) at an LPR of 3200:1 and allowed to equilibrate overnight at 37 °C. The supernatant of each sample after centrifugation in a microfuge (13000 rpm, 30 min) was analysed by SDS-PAGE (see Section 2.3.1), with and without boiling, and by Western blotting (see Section 2.3.3). As a control, each protein was incubated in 1 M urea then mixed with liposomes individually. In order to estimate the amount of HT PagP rescued from aggregation by Skp, 3.6 μM HT PagP solubilised in 50 mM glycine, pH 9.5 containing 10 M urea was included on the gels as a loading control (data not shown). Densitometry was carried out using GeneTools image analysis software (Syngene) to estimate the intensity of the relevant bands and the fraction rescued calculated by dividing the intensity of the test band by the intensity of the loading control.

2.6.7 Equilibrium Denaturation of Periplasmic Chaperones

SurA was dissolved in 50 mM glycine buffer, pH 9.5 and diluted into separate aliquots of 50 mM glycine buffer containing different concentrations of urea (0–8 M in 0.2 M increments). The final concentration of protein was 2.5 μM. The aliquots were equilibrated overnight (16 h) at 37 °C before measurement. The fluorescence emission intensity was measured at 335 nm (following excitation at 280 nm), due to the large intensity difference between the folded and unfolded states at this wavelength, for 60 s and the average signal calculated.

Equilibrium denaturation of Skp was also carried out in 50 mM glycine, pH 9.5 and followed a similar protocol, however, the final concentration of Skp in each aliquot was increased to 15 μM. Following overnight equilibration at 37 °C, the far-UV CD signal at 222 nm of each sample was recorded for 60 s and averaged. Far-UV CD was used as

Skp contains no Trp residues and thus cannot easily be studied using tryptophan fluorescence emission.

3 OMPT AND OMPP: ESTABLISHING A NEW MODEL FOLDING SYSTEM

3.1 INTRODUCTION

The first detailed investigation on the folding mechanism of an OMP was that on OmpA, an ion channel from *E. coli*. These studies revealed that folding and membrane insertion occurs as a concerted process that is highly dependent on bilayer properties (see Section 1.6)[144; 145; 147; 153; 161]. Only one other OMP, a C-terminally his-tagged construct of the *E. coli* acyltransferease PagP, has been the subject of a detailed folding analysis (described in Section 1.8.2)[167; 238; 240]. The work of Huysmans *et al*[167; 238; 240] revealed that folding of HT PagP is also highly influenced by membrane properties[240] and proceeds *via* a tilted insertion mechanism, with concerted folding and membrane insertion[167]. From these data it can be concluded that folding and membrane insertion of bacterial OMPs is likely to be a concerted process, however, the polarised transition state for HT PagP folding[167] could arise from the presence of an N-terminal α-helix (Figure 1.3a)[49; 231; 232] in the structure of this protein which is not conserved in the OMP family (Figure 1.3). The N-terminal α-helix has been shown to increase the stability of folded PagP as its deletion causes an increased rate of PagP unfolding[238], suggesting that this structural feature does indeed influence the folding pathway. While similarities and differences in the folding mechanisms of OmpA and HT PagP are apparent, it is clear that detailed folding information on a larger subset of OMPs will be required to fully understand the underlying principles of OMP folding.

The use of the "fold approach", that is, the study of structurally related proteins in order to find generic folding principles, has been long established in the field of soluble protein folding (see Section 1.7)[175; 177]. Recently, the work of the Fleming group has begun to apply this approach to OMP folding by conducting folding screens of nine different OMPs in order to find the optimal folding conditions for maximum folding efficiency[162; 164; 174]. These studies have so far been unable to achieve the level of mechanistic detail reported for OmpA[144; 145; 147; 153; 161] or HT PagP[167; 238; 240], resulting in limited opportunities to compare and contrast OMP folding pathways. In order to expand current knowledge in this area, the OM proteases, OmpT and OmpP, were

selected as the subject of preliminary folding studies to assess their suitability as OMP folding models.

3.1.1 The Outer Membrane Proteases OmpT and OmpP

The OM proteases, OmpT and OmpP, of *E. coli* have been linked to bacterial pathogenicity, due to their ability to cleave antimicrobial peptides such as colicins secreted by competing bacteria and proteamines secreted by mammalian urinary tract epithelial cells[253; 254]. The crystal structure of OmpT (Figure 1.3c) revealed it to be a 10-stranded, antiparallel β-barrel with long loops protruding from the extracellular face of the OM, which form the active site[51]. Determination of the inhibitor profile of OmpT led to its classification as a serine protease[255] and mutagenesis studies were later undertaken using a fluorogenic peptide to monitor peptidase activity to test this classification with the results suggesting that histidine-212 and serine-99 are active site residues[242]. Unlike a conventional serine protease, however, the mutation of the serine did not completely abolish enzymatic activity[242]. Analysis of the crystal structure showed that serine-99 is approximately 9 Å away from histidine-212, suggesting that this serine residue was unlikely to be directly involved in the catalytic mechanism (Figure 3.1)[51].

Further mutagenesis work was undertaken to elucidate the OmpT catalytic mechanism[256]. The very negative environment of the active site was suggested to promote the specific cleavage between basic residues, with aspartate-210 and histidine-212 making up a novel catalytic dyad[51; 256]. Activation of a water molecule by histidine-212 creates the nucleophile for the reaction, with aspartate-210 acting to stabilise the build up of positive charge on histidine-212 during catalysis[51; 256]. Serine-99 was proposed to orient the substrate in the active site, explaining why mutation of this residue causes lowered peptidase activity[256]. These results were supported by microsecond-long MD simulations of OmpT, which suggested that mutation of serine-99 caused the loss of a hydrogen bond with the substrate, leading to its dissociation from the enzyme[257]. Mutation of either aspartate-83 or aspartate-85 to alanine also abolished the protease activity of OmpT, suggesting these residues play an important

role in the catalytic mechanism, perhaps by coordinating the water molecule and activating it as a nucleophile[256].

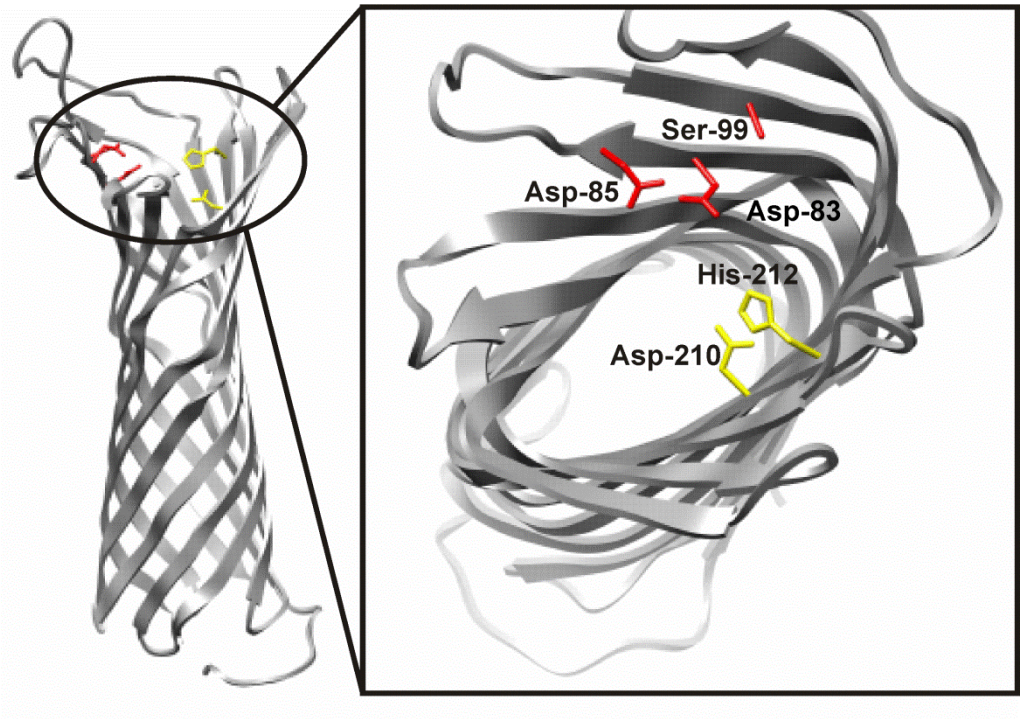

Figure 3.1 The active site residues of OmpT. Ribbon diagram showing the location of the active site residues aspartate-210 and histidine-212 in yellow and aspartate-83 and aspartate-85, which coordinate the water nucleophile, in red (left). The enlargement shows the same residues viewed from above (right). Side chains of active site residues are shown using stick representation. The position of serine-99 is also shown in red; however crystallisation was carried out on the S99A variant of the protein. Amino acids are labelled using the standard three-letter codes. This image was generated from Protein Data Bank (PDB)[43] file (1I78[51]) using UCSF Chimera molecular visualization application[44].

OmpT refolding *in vitro* (monitored by cold SDS-PAGE assays) was first reported into micelles of *N*-dodecyl-*N*,*N*-dimethyl-1-ammonio-3-propanesulphonate[138]. Acquisition of far-UV CD spectra of OmpT in these detergent micelles confirmed the high β-sheet content of the sample, suggesting that it is indeed folded in this environment. Despite successful refolding, the enzyme was found to be inactive until LPS was added indicating that although this molecule is not required for folding of OmpT, it is required for its activity[138]. Further investigation revealed that other membrane lipids are unable

to activate the enzyme[258]. It was proposed that this is a safety mechanism to prevent activation of OmpT during trafficking to the OM and that LPS may allosterically activate the enzyme *via* a subtle conformational change in the active site since the far-UV CD spectrum of the LPS-OmpT complex was not significantly different from the spectrum of OmpT alone[258].

More recently, cold SDS-PAGE has been used to monitor the folding of OmpT into lipid vesicles alongside eight other OMPs[162]. The results suggest that OmpT folding yield is highest in liposomes composed of PC lipids with shorter acyl chains at high pH but PE and PG are tolerated as guest lipids[162]. In liposomes composed of lipids with longer acyl chains, folding yield of OmpT was improved by decreasing vesicle diameter and hence increasing membrane curvature[162]. The rate of OmpT folding and insertion into SUVs of native *E. coli* lipids was shown to be unaffected by the presence of the molecular crowding agent Ficoll 70, but with a reduced overall folding efficiency[259]. Analysis of analytical ultracentrifugation data allowed the aggregation propensity of unfolded OMPs to be assessed and revealed that OmpT is particularly aggregation prone below 3 M urea or in the presence of salt but that higher pH (8–10) can reduce the extent of aggregation[164].

In contrast with OmpT, little is known about OmpP; a crystal structure has not been obtained and refolding *in vitro* has not been reported. The substrate specificity of OmpP, however, has been studied and is very similar to OmpT[243; 245]. Both proteins cleave peptides between basic residues with arginine being preferred in the P1 and P1' positions, but OmpP showed a greater ability to accept lysine in either position[243; 245]. The similar but non-identical substrate specificities of OmpT and OmpP, and their co-expression *in vivo*, suggests that these proteins work together to provide protection against a range of antimicrobial peptides[51; 245; 254]. Disregarding the N-terminal signal sequences, OmpT shares 87 % sequence similarity (and 71 % sequence identity) with OmpP[244; 245] allowing a homology model of OmpP to be created from the crystal structure of OmpT (Figure 3.2). While the residues responsible for the catalytic mechanism are conserved, the active site cleft of OmpP contains approximately half of the non-conserved residues and these differences could explain how the differences in substrate specificity of the two proteins arise[243; 245].

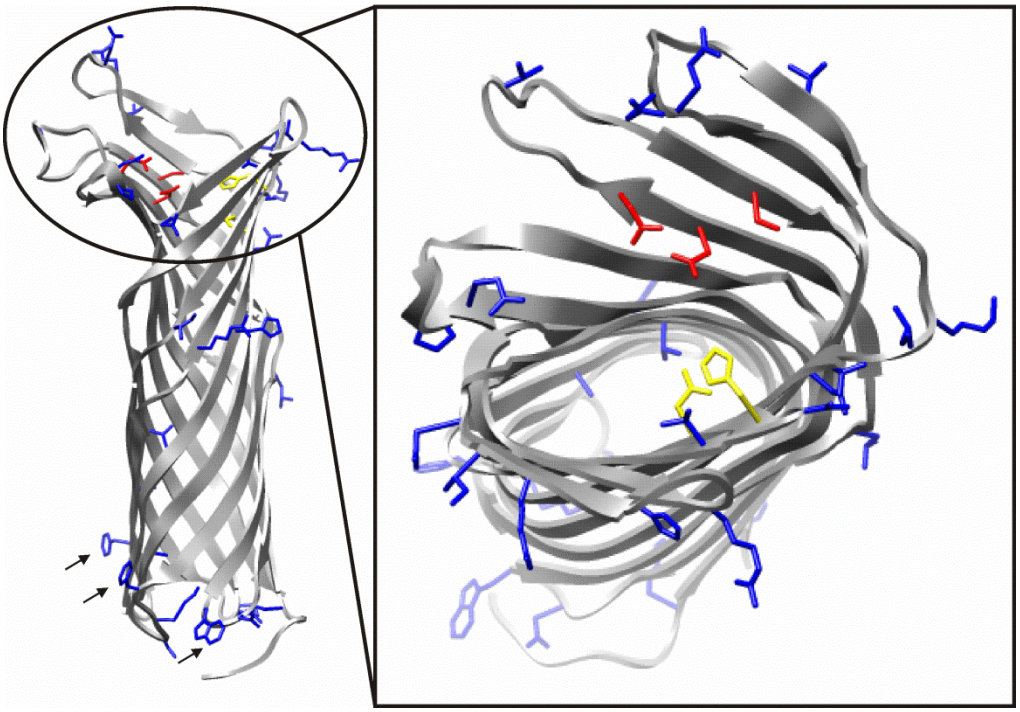

Figure 3.2 Homology model of OmpP showing non-conserved residues. Ribbon diagram showing the location of the active site residues aspartate-207 and histidine-209 in yellow and aspartate-80, aspartate-82 and serine-96, which are also involved in catalysis, in red (left). Residues which are not conserved in OmpT are shown in blue. Extra aromatic residues in the periplasmic aromatic girdle which are not found in OmpT (tryptophan-60, phenylalanine-117 and phenylalanine-191) are indicated by black arrows. The enlargement shows the same residues viewed from above (right). Side chains of coloured residues are shown using stick representation. This model was created from the Protein Data Bank (PDB)[43] file (1I78[51]) of the OmpT crystal structure using MODELLER[260] with help from Theodoros Karamanos (University of Leeds) and the image was generated using UCSF Chimera molecular visualization application[44].

The previous work on OmpT demonstrates the viability of this protein as a subject for *in vitro* folding studies as it can be over-expressed and purified in *E. coli*, and refolded in the absence of chaperones[138; 162; 259]. The protease activity of OmpT has also been shown to be a useful reporter of correct folding of this protein *in vitro*[137]. The parallel study of the homologous protein, OmpP, aims to identify the similarities and differences between the folding of OMPs which are similar in structure and sequence. Both OmpT and OmpP were cloned from *E. coli* DNA, and protocols were established for the over-expression and purification of both His-tagged and untagged constructs. Folding screens were then carried out on the untagged construct of both proteins into *di*$C_{12:0}$PC, *di*$C_{14:0}$PC and *di*$C_{16:1}$PC liposomes (Figure 3.3), revealing that these OMPs are able to

fold with high efficiency to the native state in a range of lipid conditions. Under all conditions tested, OmpT is seen to require lower urea concentrations than OmpP to reduce aggregation and promote folding. Additionally, once inserted, OmpP requires lower concentrations of denaturant to induce unfolding. Given that many of the non-conserved residues in OmpP are localised in the active site, these results suggest that the reduced stability of OmpP and increased aggregation propensity result from the need to sacrifice stability and folding efficiency in order to optimise function.

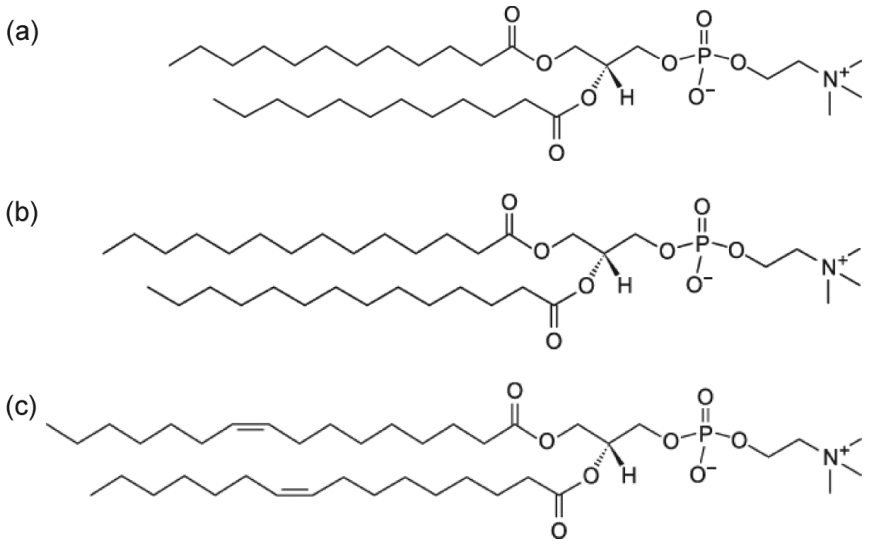

Figure 3.3 Structures of PC lipids used in this chapter. (a) 1,2-dilauroyl-*sn*-glycero-3-phosphocholine (*di*$C_{12:0}$PC); (b) 1,2-dimyristoyl-*sn*-glycero-3-phosphocholine (*di*$C_{14:0}$PC); and (c) 1,2-dipalmitoleoyl-*sn*-glycero-3-phosphocholine (*di*$C_{16:1}$PC).

Following the development of folding and functional assays to study OmpT and OmpP insertion into liposomes, these experiments were adapted in collaboration with Aneika C. Leney (University of Leeds) to fold OmpT into the Apol, A8-35. OmpT folded with almost 100 % efficiency to its native state and study of the protein by electrospray ionisation-mass spectrometry (ESI-MS) was successful.

The work described in Section 3.2.6 has been published in: Leney, A. C., McMorran, L. M., Radford, S. E. & Ashcroft, A. E. (2012). Amphipathic polymers enable the study of functional membrane proteins in the gas phase. *Anal. Chem.* **84**, 9841-9847.

3.2 RESULTS

3.2.1 CLONING, OVER-EXPRESSION AND PURIFICATION OF HIS-TAGGED CONSTRUCTS

3.2.1.1 Gene Cloning

The genes encoding the OmpT and OmpP proteins were amplified from *E. coli* strain XL1-blue (Stratagene, UK)[247] as described in Section 2.2. Primers were designed to allow amplification of the DNA encoding the mature protein sequences, that is, the proteins without the N-terminal signal sequences that result in targeting to the OM. Removal of the signal sequence results in cytoplasmic protein expression as insoluble inclusion bodies[162]. Primers included an *Nde*I restriction site 5' to the gene as this introduces an ATG start codon to the sequence and a *Bam*HI restriction site 3' to the gene. Digestion of the amplified genes with these enzymes creates overhangs that are complementary to those within a pET11a plasmid digested with the same enzymes, allowing the ligation of the genes into the desired vector. Additionally, a sequence encoding a hexa-histidine tag (His-tag) was included on the N-terminal primers between the *Nde*I restriction site and the sequence coding for the mature proteins to facilitate purification. In order to remove the His-tag following protein purification, a recognition site for Tobacco Etch Virus protease (TEVp) was encoded between the gene sequence and the His-tag on each N-terminal primer. The sequences of the primers used are provided in Table 2.2 and a schematic of the designed primers is shown in Figure 3.4. The full DNA and protein sequences for His-tagged OmpT (HT OmpT) and His-tagged OmpP (HT OmpP) are provided in the Appendix.

Bacterial chromosomal DNA was extracted from *E. coli* XL1-blue cells by diluting a liquid culture of the bacteria with deionised water and boiling to lyse the cells. This solution was then used directly as a DNA template in a series of polymerase chain reactions (PCR) containing different concentrations of Mg^{2+} to find the optimum amplification conditions as described in Section 2.2.6. Amplification of the HT OmpT and HT OmpP genes from the bacterial DNA was successful (Figure 3.5) and following gel electrophoresis the desired bands were excised from the gel and purified by gel extraction (see Section 2.2.6). The amplified genes were "A-tailed" by incubation with dATP and Taq polymerase before ligation into a shuttle vector, pGEM-T Easy vector

(Promega, UK), which is supplied as a linearised plasmid with a single T overhang at the 5' end of each DNA strand (Section 2.2.7). The ligation products were transformed into *E. coli* XL1-blue cells and colonies containing putative inserts identified using blue/white colour screening (Section 2.2.7). Following ligation of both the genes, the transformation plates contained around 200 colonies, of which half were white. Six white colonies were selected from each plate and grown in a 5 ml overnight liquid culture before the plasmid DNA was extracted and then digested using the restriction enzymes *Bam*HI and *Nde*I as described in Section 2.2.8. Plasmids that were found to contain an insert of the expected size (939 bp and 924 bp for HT OmpT and HT OmpP, respectively) were sent for DNA sequencing. Once the correct sequence had been confirmed, inserts were purified by restriction digestion followed by agarose gel electrophoresis and gel extraction.

Ligation into the pET11a vector, which had been digested with the same restriction enzymes as the HT OmpT and HT OmpP genes, was carried out as described in Section 2.2.9. These reactions led to successful transformation of *E. coli* XL1-blue cells, with around 150 colonies on each transformation plate. Six colonies were picked and grown as liquid cultures from which plasmid DNA was purified. Samples of DNA from each culture were digested using the restriction enzymes *Bam*HI and *Nde*I to confirm the presence of inserted genes (Figure 3.6). Plasmids which contained digested inserts of the correct size (939 bp and 924 bp for HT OmpT and HT OmpP, respectively) were sent for DNA sequencing and confirmed to contain the correct sequences for HT OmpT (plasmid pLMM01) and HT OmpP (plasmid pLMM02) (see Appendix).

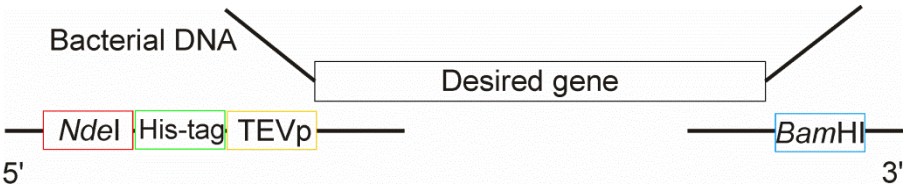

Figure 3.4 Schematic of primers designed to create HT OmpT and HT OmpP constructs. The location of the *Nde*I restriction site is shown in red, the sequence coding for a Tobacco Etch Virus protease (TEVp) cut site shown in yellow, the sequence coding for the hexa-histidine tag (His-tag) shown in green and the *Bam*HI restriction site shown in blue.

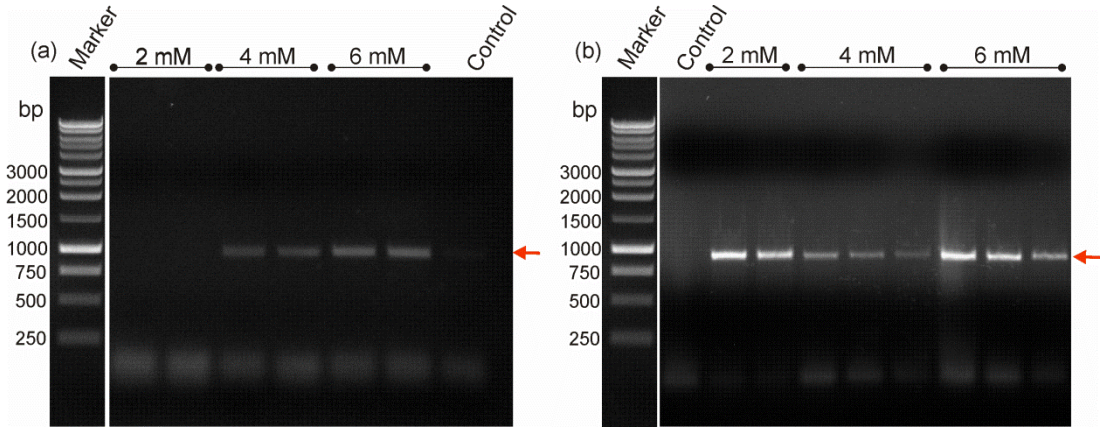

Figure 3.5 Agraose gel electrophoresis following PCR to amplify genes encoding HT OmpT and HT OmpP from *E. coli* chromosomal DNA. PCR was carried out to amplify the genes encoding (a) OmpT and (b) OmpP from the DNA of XL1-blue *E. coli* cells. Lanes show replicate samples and are labelled according to the final concentration of Mg^{2+} in each sample. Control reactions contained all components except the DNA template. The number of base pairs (bp) corresponding to bands in the DNA marker are indicated. DNA fragments of correct expected size are labelled with red arrows. Note that the amplified OmpT is expected to contain 952 bp and the amplified OmpP is expected to contain 937 bp.

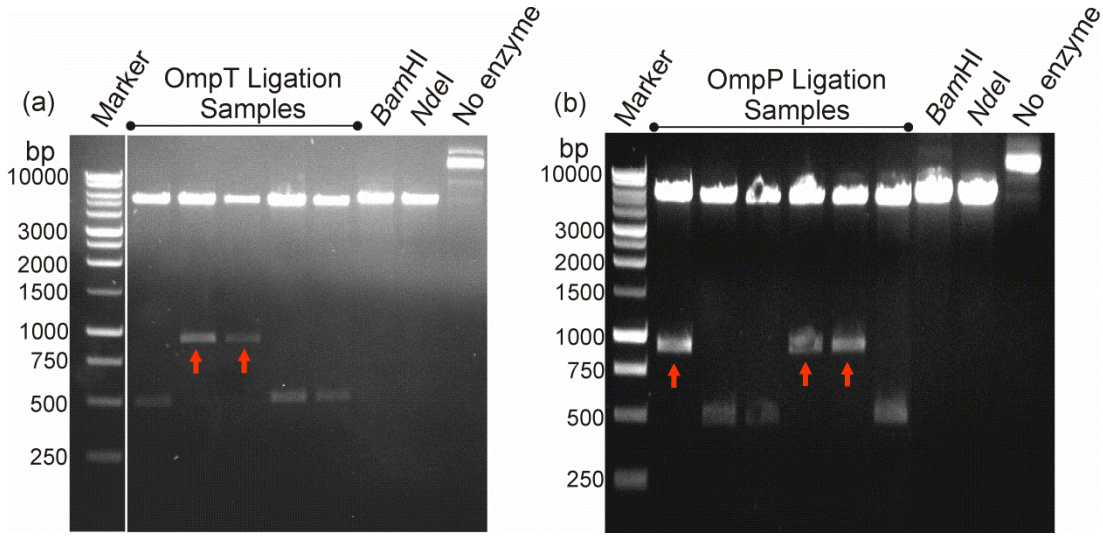

Figure 3.6 Analytical restriction digests following ligation of the genes encoding HT OmpT and HT OmpP into the pET11a plasmid vector. Sample lanes contain the products of digestion reactions of DNA taken from different single colonies transformed with (a) HT OmpT gene and (b) HT OmpP gene ligation mixtures. DNA fragments of correct expected size are labelled with red arrows. Single enzyme digestion reactions and enzyme-free reactions of pET11a are included. The number of base pairs (bp) corresponding to bands in the DNA marker are indicated.

3.2.1.2 Over-Expression

The plasmid vectors containing the genes for HT OmpT (pLMM01) or HT OmpP (pLMM02) were transformed into *E. coli* strain BL21 (DE3)[248] (Stratagene, UK) and grown on agar under antibiotic selection. Single colonies were used to inoculate 5 ml overnight cultures, which in turn were used to inoculate 100 ml lysogeny broth (LB) cultures at a ratio of 1:100. Protein production was induced by addition of isopropyl-β-D-thiogalactoside (IPTG) to a final concentration of 1 mM upon the culture reaching an OD_{600} of 0.6 (see Section 2.4.1). After induction of protein expression, 1 ml samples were removed from the cultures every two hours to monitor protein over-expression. To ensure comparative loadings at each of the time points, the OD_{600} of each sample was measured, the cells harvested by centrifugation and then resuspended in 100 µl SDS-loading buffer for each OD_{600} unit. Bands corresponding to proteins of the expected molecular masses (35.3 kDa and 34.9 kDa for HT OmpT and HT OmpP, respectively) were observed in the cell lysate of whole cell boils by SDS-PAGE (Figure 3.7). Protein expression reached a maximum four hours after induction, indicating that this is a sufficient amount of time for protein production to occur before the cells are harvested.

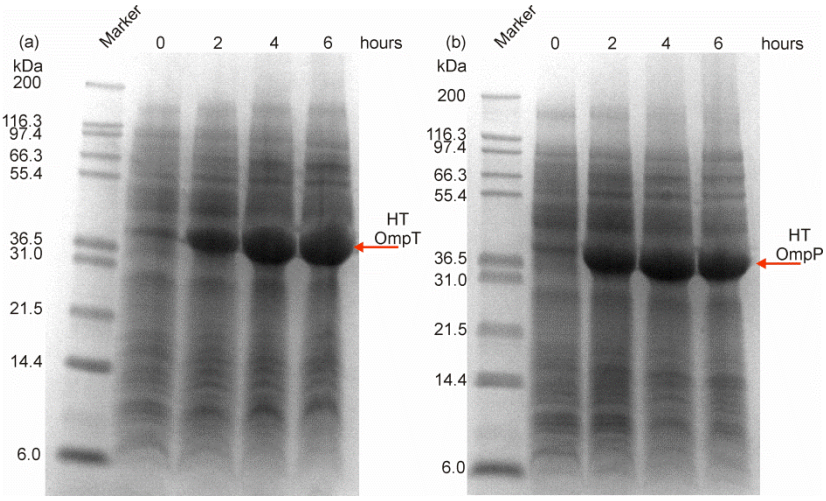

Figure 3.7 Small-scale expression trials of HT OmpT and HT OmpP. 100 ml *E. coli* BL21 (DE3) cultures were transformed with (a) pLMM01 encoding HT OmpT and (b) pLMM02 encoding HT OmpP and grown at 37 °C with shaking. Protein expression was induced at an OD_{600} of 0.6 with 1 mM IPTG. Samples are labelled according to the time after induction at which they were removed. The size in kiloDaltons (kDa) of the protein markers is indicated.

Given the high levels of over-expression observed for both HT OmpT and HT OmpP in small scale expression trials, the use of large volumes of bacterial culture were deemed unnecessary for the production of these proteins. Instead, a single 1000 ml culture or two 500 ml cultures of BL21 (DE3) cells transformed with pLMM01 and pLMM02 were grown. Expression was induced with 1 mM IPTG at an OD_{600} of 0.6 and cell growth continued for a further four hours (37 ^{o}C, 200 rpm). Cells were harvested by centrifugation and the cell pellets weighed (Table 3.1). The higher yield of bacterial cells from the two 500 ml cultures grown in two-litre conical flasks compared with one 1000 ml culture grown in a two-litre conical flask suggests that growing the bacteria in smaller volumes relative to the total volume of the conical flask will increase the final protein yield. This procedure was adopted for all further experiments.

Culture volume (ml)	Mass of HT OmpT cell pellet (g)	Mass of HT OmpP cell pellet (g)
1 × 1000	2.76	2.74
2 × 500	3.46	3.57

Table 3.1 Cell pellet yields for BL21 (DE3) cultures transformed with either plasmid pLMM01 (encoding HT OmpT) or pLMM02 (encoding HT OmpP). All cultures were grown in conical flasks with a capacity of two litres.

3.2.1.3 Purification

To enable purification of HT OmpT and HT OmpP, two 500 ml cultures of BL21 (DE3) cells transformed with the appropriate plasmid were grown as described above (Section 3.2.1.2). The cells were then harvested by centrifugation before lysis by sonication (Section 2.4.2). Lysed cells were treated with a 2 % (*v/v*) solution of Triton X-100 (Calbiochem, Germany) to solubilise the membranes. The insoluble fraction was then collected by centrifugation and washed with 25 mM Tris-HCl, pH 8.0 to remove any residual detergent (Section 2.4.2). Many of the impurities present are removed with the soluble fraction and only a small amount of the desired protein is lost (Figure 3.8).

The isolated inclusion bodies were next solubilised and HT OmpT and HT OmpP purified further by nickel affinity chromatography (see Section 2.4.3). Briefly, the inclusion body was solubilised in 6 M GuHCl, 50 mM sodium phosphate, 250 mM NaCl, pH 8.0 and bound to 4 ml of nickel Sepharose resin (GE Healthcare, UK). The resin was washed with buffer containing 5 mM imidazole, before elution in buffer containing 250 mM imidazole. The protein was then dialysed against dH_2O and the precipitate redissolved in 6 M GuHCl, 25 mM Tris-HCl, pH 8.0 and stored in 200 µl aliquots at -80 °C. Protein greater than 95 % in purity was obtained for both HT OmpT and HT OmpP with yields of approximately 80 mg pure protein per litre of culture (Figure 3.9).

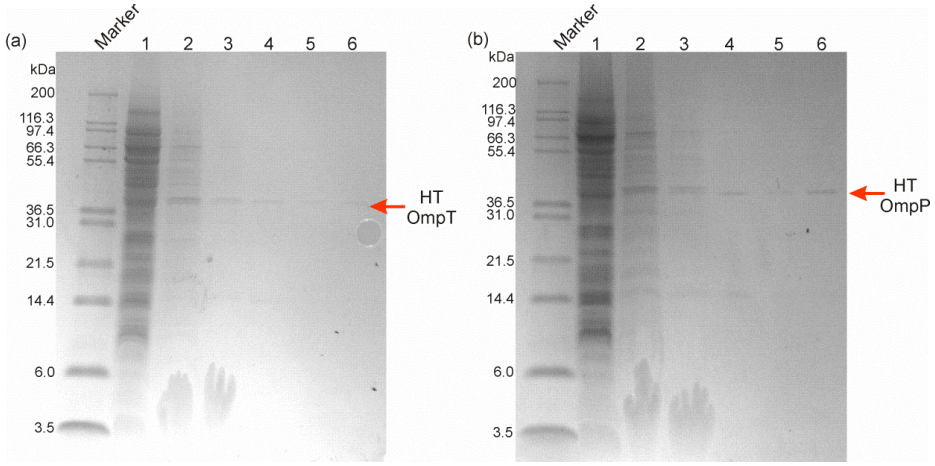

Figure 3.8 SDS-PAGE gels showing the protein content of the discarded soluble fractions during inclusion body isolation. During isolation of (a) HT OmpT and (b) HT OmpP inclusion bodies, small amounts of the desired protein are lost. Lanes are labelled as follows: lane 1 contains the soluble fraction of the whole cell lysate, lanes 2 and 3 contain the supernatant from subsequent detergent treatments after the insoluble fraction was collected by centrifugation and lanes 4–6 contain the supernatant from subsequent buffer washes after the insoluble fraction was collected by centrifugation. The size in kiloDaltons (kDa) of the protein markers is indicated.

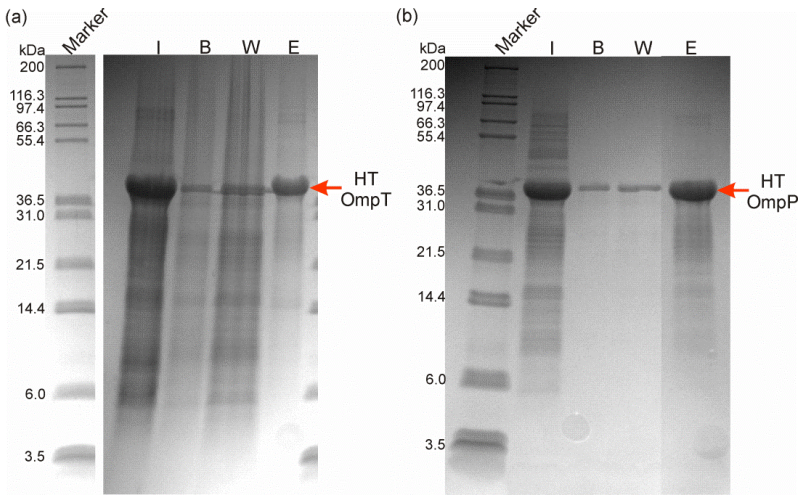

Figure 3.9 Purification of HT OmpT and HT OmpP using nickel affinity chromatography. SDS-PAGE gels showing (I) impure solubilised inclusion bodies, (B) buffer containing unbound protein after addition to nickel Sepharose resin, (W) wash buffer containing 5 mM imidazole and (E) eluted protein after addition of 250 mM imidazole of both (a) HT OmpT and (b) HT OmpP. Purification of both proteins was carried out in 50 mM sodium phosphate, 6 M GuHCl, 250 mM NaCl, pH 8.0. The size in kiloDaltons (kDa) of the protein markers is indicated.

3.2.1.4 Tobacco Etch Virus Protease (TEVp) Cleavage

Following purification of HT OmpT and HT OmpP, removal of the His-tag was attempted using TEVp (kindly provided by Alessandro Sicorello, University of Leeds). His-tagged OMPs in 50 mM sodium phosphate, pH 8.0 containing 10 M urea were diluted into low urea (final concentration of 2, 3 or 4 M) or added to 12 mM Tween-20 in the presence of 1 M urea. TEVp was added at an OMP:TEVp ratio of 1:5 and incubated for 3 h at either 25 °C or 30 °C. In all experiments, the OMP concentration was 5 μM. Following analysis of the reactions by SDS-PAGE, it was shown that for HT OmpT a small amount of cleaved product was observable in 2 M urea or in Tween-20 and 1 M urea, while for HT OmpP only 2 M urea gave any observable cleaved product (Figure 3.10). This may be due to unfolding of TEVp in the presence of urea or Tween-20. Alternatively, inaccessibility of the TEVp cut site due to collapse of the OMPs in the presence of low urea or burial of the cut site within the Tween-20 micelle. Since no condition tested gave a high yield of cleaved product, it was instead decided to use a purification strategy based on untagged constructs in order to improve the final yield of protein obtained.

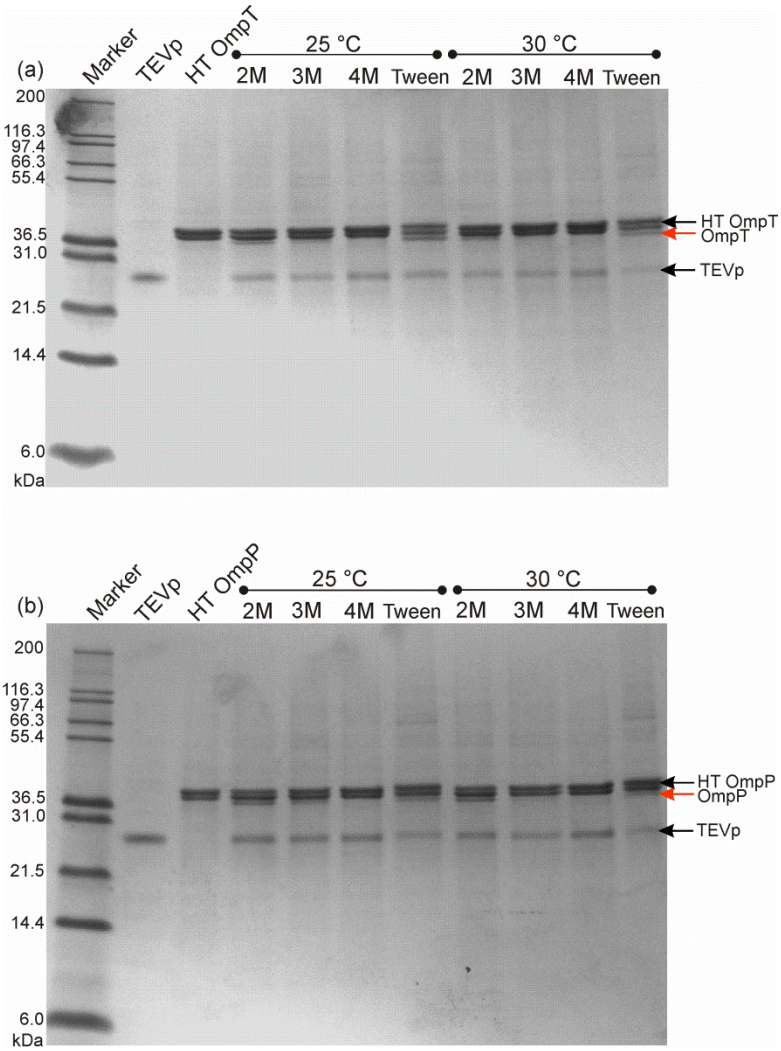

Figure 3.10 Cleaving the His-tag from OmpT and OmpP using Tobacco Etch Virus protease (TEVp). (a) HT OmpT and (b) HT OmpP were diluted to 2, 3 or 4 M urea, or 1 M urea in the presence of 12 mM Tween-20 (indicated at 2M, 3M, 4M and Tween, respectively) and incubated for 3 h in the presence of a 1:5 excess of TEVp. Reactions were carried out at either 25 °C or 30 °C. Black arrows indicate the bands corresponding to uncleaved OmpT/P and TEVp, while the red arrow indicates the band corresponding to cleaved OmpT/P. All reactions were carried out on 5 μM protein in 50 mM sodium phosphate, pH 8.0. The size in kiloDaltons (kDa) of the protein markers is indicated. TEVp was kindly supplied by Alessandro Sicorello, University of Leeds.

3.2.2 Cloning, Over-Expression and Purification of Untagged Constructs

3.2.2.1 Cloning

To create constructs for the over-expression of untagged variants of OmpT and OmpP, primers were designed to amplify the genes encoding the mature sequences of the proteins using the plasmids pLMM01 and pLMM02 as templates. Forward primers were designed to include an *Nde*I restriction site 5' to the gene as this introduces an ATG start codon to the sequence. The reverse primer was designed to adhere to the pET11a plasmid 3' to the end of the gene encoding the desired protein, no restriction site was included in this primer, as the original pLMM01 and pLMM02 plasmids include a BamHI restriction site 3' to the gene. The sequences of the primers used are provided in Table 2.4 and a schematic of the designed primers is shown in Figure 3.11. The full gene and protein sequences for untagged variants of OmpT and OmpP are provided in the Appendix.

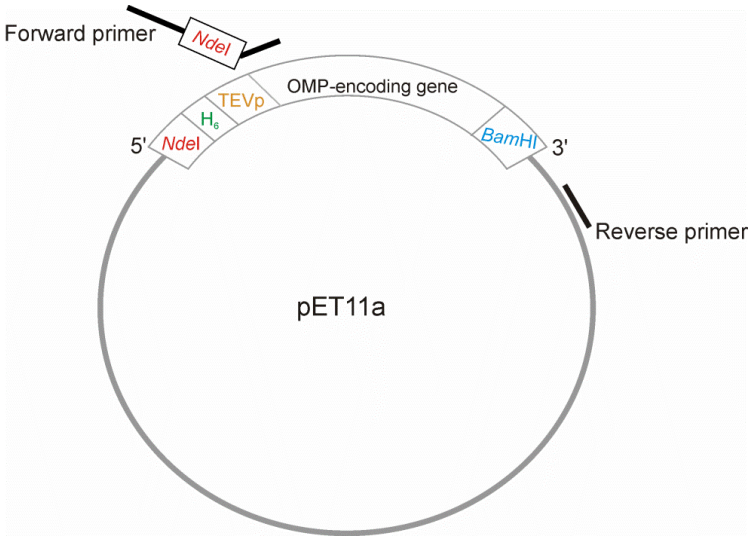

Figure 3.11 Schematic of the cloning strategy used to create untagged OmpT and OmpP constructs. The location of *Nde*I restriction sites are shown in red, the sequence coding for a Tobacco Etch Virus protease (TEVp) cut site shown in yellow, the sequence coding for the hexa-histidine tag (His-tag) shown in green and the *Bam*HI cut site shown in blue. The designed primers are shown as black lines at the approximate locations where they bind to the plasmid template.

Amplification of the genes encoding OmpT and OmpP was successful using a series of PCR reactions containing different concentrations of Mg^{2+} to find the optimum amplification conditions (Figure 3.12) as described in Section 2.2.6. Following gel electrophoresis the desired bands were excised from the gel and purified by gel extraction (see Section 2.2.6). The expected sizes of the fragments were 937 bp for OmpT and 922 bp for OmpP.

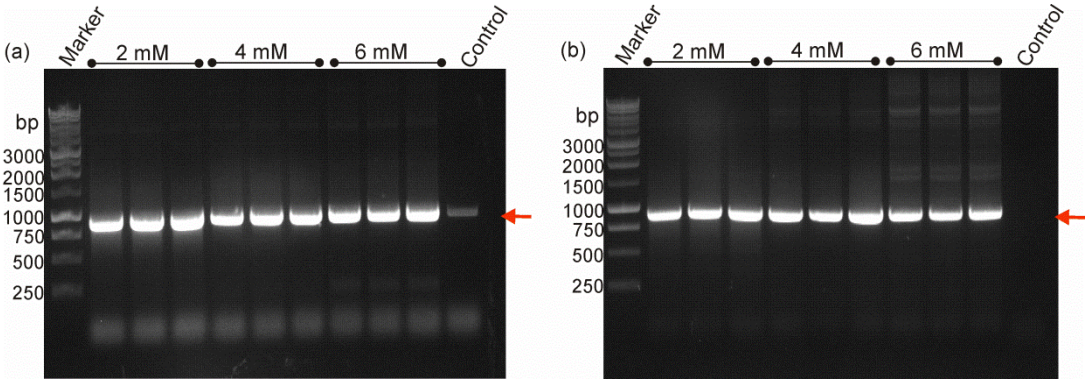

Figure 3.12 Agarose gel electrophoresis following PCR to amplify genes encoding OmpT and OmpP. PCR was carried out to amplify the genes encoding (a) OmpT and (b) OmpP using plasmid templates. Lanes show replicate samples and are labelled according to the final concentration of Mg^{2+} in each sample. Control reactions contained all components except the DNA template. The number of base pairs (bp) corresponding to bands in the DNA marker are indicated. DNA fragments of correct expected size are labelled with red arrows. Note that the amplified OmpT fragment is expected to contain 937 bp and the amplified OmpP fragment is expected to contain 922 bp.

The cloning strategy employed resulted in amplified fragments with a minimum of 12 bp on either side of the restriction enzyme binding sites, so the use of a shuttle vector was deemed unnecessary. Following restriction digestion reactions with *Nde*I and *Bam*HI enzymes as described in Section 2.2.8, inserts were ligated directly into pET11a digested with *Nde*I and *Bam*HI as described in Section 2.2.9. These reactions led to successful transformation of *E. coli* XL1-blue cells, with around 50 colonies on each transformation plate. Six colonies were picked and grown as liquid cultures from which plasmid DNA was purified. Samples of DNA from each culture were digested using the restriction enzymes *Bam*HI and *Nde*I to confirm the presence of inserted sequences (Figure 3.13). Plasmids which contained digested inserts of the correct size

(900 bp and 885 bp for OmpT and OmpP, respectively) were sent for DNA sequencing and confirmed to contain the correct sequences for OmpT (plasmid pLMM03) and OmpP (plasmid pLMM04) (see Appendix).

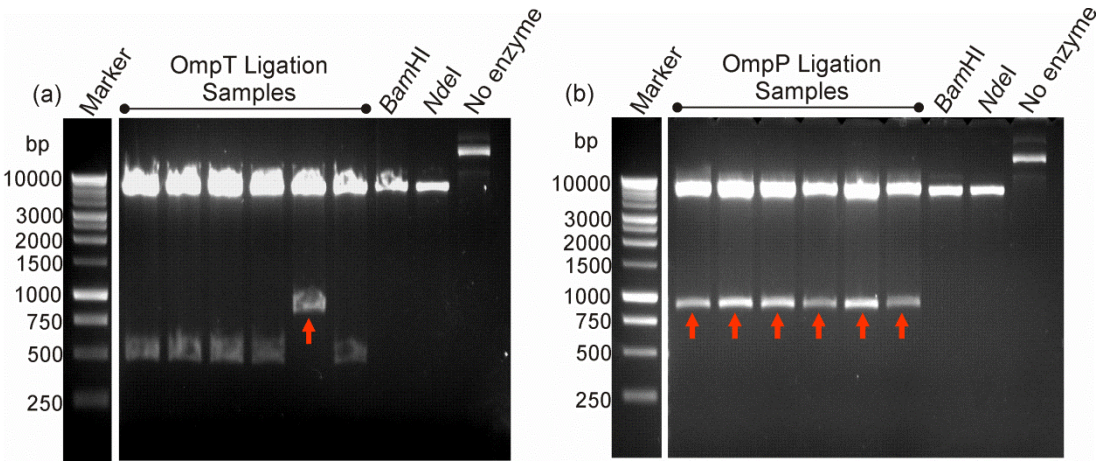

Figure 3.13 Analytical restriction digests following ligation of the genes encoding OmpT and OmpP into the pET11a plasmid vector. Sample lanes contain the products of digestion reactions of DNA taken from different single colonies transformed with ligation mixtures containing (a) the gene encoding OmpT and (b) the gene encoding OmpP. DNA fragments of correct expected size are labelled with red arrows. Single enzyme digestion reactions and enzyme-free reactions of pET11a are included. The number of base pairs (bp) corresponding to bands in the DNA marker are indicated.

3.2.2.2 Over-Expression

The plasmid vectors containing the genes encoding OmpT (pLMM03) or OmpP (pLMM04) were transformed into *E. coli* strain BL21 (DE3) (Stratagene, UK) and grown on agar under antibiotic selection. Single colonies were used to inoculate 5 ml overnight cultures, which in turn were used to inoculate 100 ml LB cultures at a ratio of 1:100. Protein production was induced by addition of IPTG to a final concentration of 1 mM upon the culture reaching an OD_{600} of 0.6 (see Section 2.4.1). Protein over-expression was monitored as described for the His-tagged proteins in Section 3.2.1.2. Bands corresponding to proteins of the expected molecular masses (33.6 kDa and 33.2 kDa for OmpT and OmpP, respectively) were observed in the cell lysate of whole cell boils by SDS-PAGE (Figure 3.14). Protein expression reached a maximum four hours

after induction, indicating that this is a sufficient amount of time for protein production to occur before the cells are harvested.

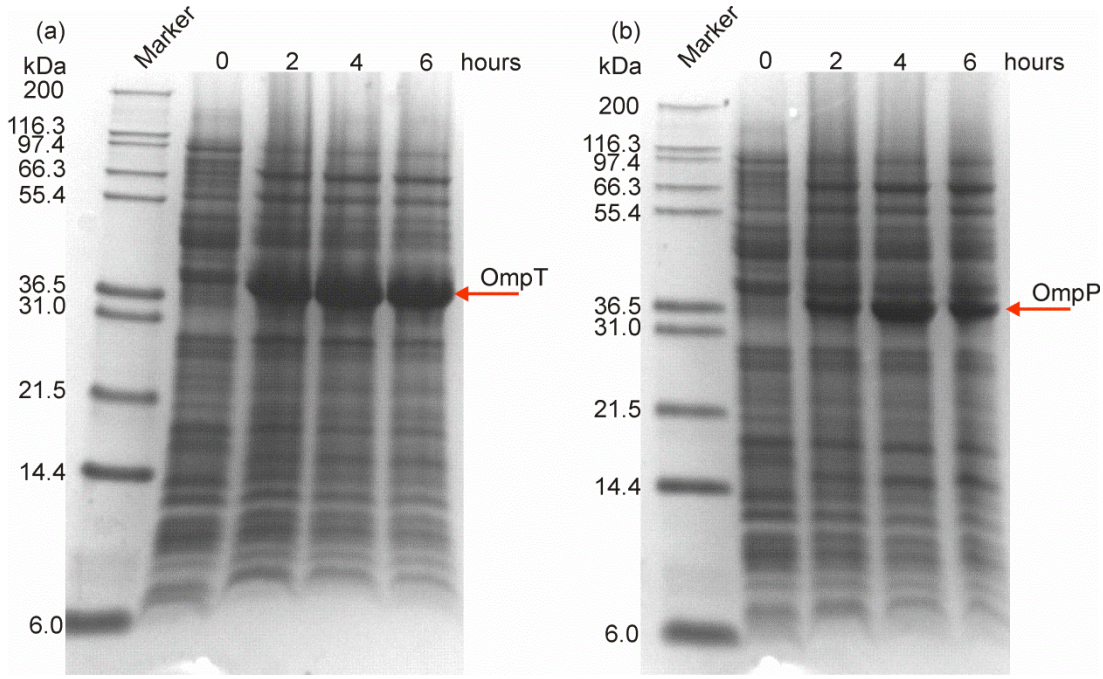

Figure 3.14 Small-scale expression trials of OmpT and OmpP. 100 ml *E. coli* BL21 (DE3) cultures were transformed with plasmids (a) pLMM03 encoding OmpT and (b) pLMM04 encoding OmpP and grown at 37 °C with shaking. Protein expression was induced at an OD_{600} of 0.6 with 1 mM IPTG. Samples are labelled according to the time after induction at which they were removed. The size in kiloDaltons (kDa) of the protein markers is indicated.

High levels of over-expression were observed for both OmpT and OmpP and so the protocol used for all subsequent large scale expression experiments was identical to that used for the His-tagged constructs (see Section 3.2.1.2). Briefly, two 500 ml cultures of BL21 (DE3) cells transformed with the desired plasmid were grown in LB medium in two-litre conical flasks. Protein expression was induced with 1 mM IPTG when the cultures reached an OD_{600} of 0.6 and expression was allowed to continue for a further 4 h (37 ^{o}C, 200 rpm) before cells were harvested by centrifugation.

3.2.2.3 Purification

To enable purification of OmpT and OmpP, two 500 ml cultures of BL21 (DE3) cells transformed with the appropriate plasmid were grown as described above (Section 3.2.2.2). The cells were then harvested by centrifugation before lysis by sonication and isolation of inclusion bodies was carried out as previously described (Section 2.4.2). The isolated inclusion bodies were next solubilised in 6M GuHCl, 25mM Tris, pH 8.0 and OmpT and OmpP purified further by size exclusion chromatography (see Section 2.4.5) using a Superdex 75 HiLoad 26/60 column (GE Healthcare). OmpT or OmpP containing fractions were pooled and concentrated to approximately 500 μM and stored in aliquots at -80 ^{o}C. Protein greater than 90 % in purity was obtained for both OmpT and OmpP with yields of approximately 65 mg pure protein per litre of culture (Figure 3.15).

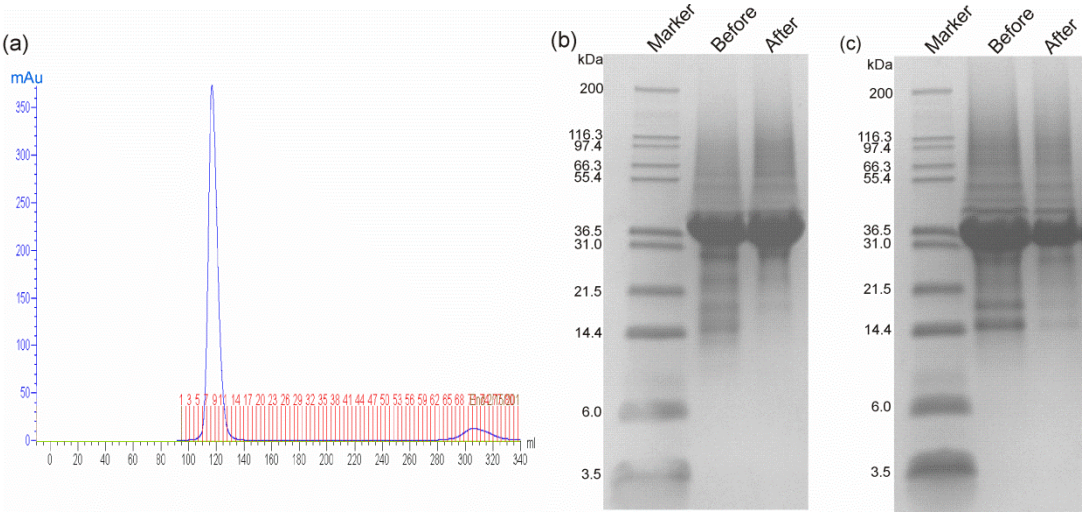

Figure 3.15 Purification of OmpT and OmpP. (a) Typical OmpT elution profile from the size exclusion column. The elution profile of OmpP is nearly identical (data not shown). SDS-PAGE gels of (b) OmpT and (c) OmpP showing protein samples before size exclusion and after. The size in kiloDaltons (kDa) of the protein markers is indicated.

3.2.3 *Folding Screens of OmpT and OmpP Using Cold SDS-PAGE*

Following purification, OmpT and OmpP were stored in an unfolded form in 6 M GuHCl at -80 °C. Therefore, conditions for the refolding of these proteins must be optimised so that a high yield of native protein can be obtained for further study. The work of Huysmans *et al*[167; 240] showed that the choice of LPR is important in influencing the folding pathway of HT PagP and also modulates the reversibility of folding. For these reasons a high LPR of 3200:1 was chosen for all subsequent experiments. For both OmpT and OmpP, the final urea concentration was varied from 2–6 M during all experiments.

Based on the protocol described by Burgess *et al*[162], 4.8 μM OMP was diluted into liposomes at an LPR of 3200:1 in varying buffer conditions and allowed to refold for 16 h. The reaction was then quenched by the addition of 2 × SDS-PAGE loading buffer and loaded on to an SDS-PAGE gel with and without prior boiling (Section 2.5.2). Due to the high stability of OMPs, the folded protein is often resistant to SDS-induced denaturation if samples are not boiled and migrates at a different apparent molecular weight to the unfolded conformation[140]. The fraction of folded OMP was estimated using densitometry and averaged over three experiments to reduce error. In most of the experiments, some higher molecular weight species are present indicating that under all conditions tested the competing process of aggregation is occurring alongside folding. Since these aggregated species can be depolymerised by boiling in SDS-PAGE loading buffer (Figure 3.16a), the fraction folded was calculated by comparing the intensity of the folded band in the unboiled sample with the intensity of the single, unfolded band observed in the boiled fraction so that the effects of aggregation are considered. Values of folding efficiency were quite reproducible between replicate samples (± 10 %). Possible sources of error could include differing extents of aggregation between samples, differences in the efficiency of staining of individual gels and disruption of the folded OMP by the presence of SDS in the SDS-PAGE loading buffer. The latter was reduced by running gels immediately after addition of SDS-PAGE loading buffer.

Using 100 nm LUVs of *di*$C_{12:0}$PC, the pH of the reaction buffer was varied from 7–10 while the temperature was fixed at either 25 °C or 37 °C. For both OmpT and OmpP, the highest yields (70–80 % folded) were seen at pH 8.0, 25 °C (Figure 3.16b, c; Table

3.2) so these conditions were kept constant throughout the remaining experiments. The effect of vesicle diameter was next examined using 50 nm SUVs of *di*$C_{12:0}$PC as this has previously been suggested to improve the folding yield of OmpA in some lipids[147]. For OmpT the decrease in liposome diameter caused little change in the folded fraction, however, for OmpP the folded yield was seen to decrease from 69 % to 46 % at 5 M urea and so it was decided not to use SUVs in subsequent experiments (Table 3.2).

Finally, the effect of acyl chain length and saturation on OmpT and OmpP folding efficiency was investigated. The hydrophobic thickness of a *di*$C_{12:0}$PC bilayer is 19.5 ± 1 Å, while that of *di*$C_{14:0}$PC is 23.0 ± 1 Å[261]. Comparing the folded yields of OmpT and OmpP at a urea concentration which gave a high yield in *di*$C_{12:0}$PC (3 M for OmpT and 5 M for OmpT), increasing the hydrophobic thickness of the membrane by using 100 nm LUVs of *di*$C_{14:0}$PC for refolding had little effect on yield (Table 3.2). Conversely, using LUVs of *di*$C_{16:1}$PC lowered the folding efficiency of both proteins, from 72 % to 56 % for OmpT in 3 M urea and from 69 % to 32 % for OmpP in 5 M urea (Table 3.2), an effect most likely due to the degree of saturation of the acyl chain since the hydrophobic thickness of *di*$C_{16:1}$PC bilayers is estimated to be 23.5 ± 1 Å[261] and hence is not significantly different to that of *di*$C_{14:0}$PC.

Interestingly, the effect of changing the concentration of urea in any given reaction is dependent on the other experimental conditions and the OMP. For example, for OmpT folding into 100 nm LUVs of *di*$C_{12:0}$PC at pH 8.0, 25 °C the folding yield is increased from 46 % in 2 M urea to 82 % in 6 M urea, however, at pH 10.0 the folding yield decreases from 62 % in 2 M urea to 48 % in 6 M urea. Conversely, OmpP folding yield in 100 nm *di*$C_{12:0}$PC LUVs at either pH 8.0 or pH 10.0 is higher when urea concentration is increased from 3 M to 6 M. The varying effects of urea concentration on the folding yield of OmpT and OmpP may be attributable to the delicate balance which exists between folding, unfolding and aggregation. If the urea concentration present is too low, insoluble OMPs will aggregate; however, if the urea concentration is too high, the unfolded, urea-solubilised state becomes more favourable than the native state. Therefore, optimising the urea concentration for folding in any given set of conditions is as important as optimising the membrane and buffer properties to ensure a high yield of folded OMP.

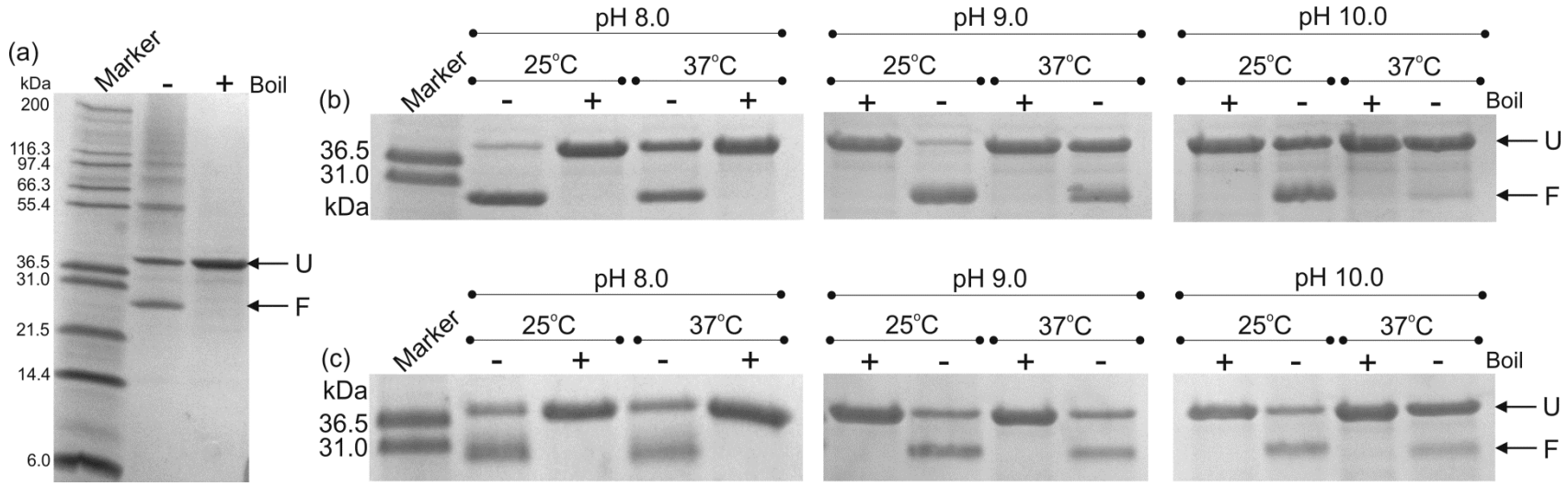

Figure 3.16 Temperature and pH dependence of folding into 100 nm *di*$C_{12:0}$PC LUVs. (a) Example gel showing higher molecular weight OmpT aggregates in the unboiled sample, which are depolymerised upon boiling. Samples contained 4.8 μM (b) OmpT or (c) OmpP at an LPR of 3200:1 and were allowed to refold in the presence of 5 M urea for 16 hours prior to 1:1 dilution with 2 × SDS-PAGE loading buffer. The relative intensities of the folded and unfolded monomer bands (indicated by F and U, respectively, in the figure) are given in Table 3.2. The following buffers were used in these experiments: 50 mM sodium phosphate, pH 8.0; 50 mM glycine, pH 9.0 and 50 mM sodium borate, pH 10.0.

		Lipid	***di*C$_{12:0}$PC**								***di*C$_{14:0}$PC**	***di*C$_{16:1}$PC**
		Vesicle Size	**50 nm**	**100 nm**								
		Temperature	**25 °C**					**37 °C**			**25 °C**	
		pH	**8.0**	**7.0**	**8.0**	**9.0**	**10.0**	**8.0**	**9.0**	**10.0**	**8.0**	**8.0**
OmpT	[Urea] (M)	**2**	43	43	46	55	62	49	53	63	75	ND
		3	54	56	72	63	69	66	74	64	73	56
		4	72	75	67	61	57	61	62	60	69	67
		5	71	61	75	73	41	51	47	22	53	60
		6	ND	ND	82	53	48	49	51	28	54	50
OmpP		**2**	16	16	–	–	–	9	9	11	9	ND
		3	26	33	44	40	48	52	54	51	37	11
		4	47	43	63	55	52	59	56	56	50	–
		5	46	61	69	58	48	63	58	28	60	32
		6	ND	ND	74	66	68	56	62	38	54	36

Table 3.2 Estimated percentage folded yield of OmpT and OmpP into PC liposomes *in vitro*. Yield was determined by using densitometry to calculate the relative intensities of the folded and unfolded monomer bands following cold SDS-PAGE analysis of the folding reaction. Yields are colour coded as follows: 0–24 % in very pale blue, 25–49 % in pale blue, 50–74 % in mid-blue and ≥ 75 % in deep blue. – denotes the absence of any detectable folded yield, while ND indicates folding conditions which were not tested.

3.2.4 OmpT and OmpP Fold to the Native State in Synthetic Liposomes

While shifts in apparent molecular weight observed in cold SDS-PAGE assays are often indicative of the presence of folded OMPs, these assays provide no structural information and do not confirm the presence of correctly folded OmpT or OmpP. To determine if natively folded OMPs are present, tryptophan fluorescence emission spectra were first measured for both OmpT and OmpP after folding for 16 h in 100 nm LUVs of *di*$C_{12:0}$PC, *di*$C_{14:0}$PC and *di*$C_{16:1}$PC at an LPR of 3200:1 in 50 mM sodium phosphate, pH 8.0, 25 °C (Figure 3.17). Upon folding into PC liposomes, the tryptophan fluorescence emission intensity increases in yield and the wavelength corresponding to maximal intensity (λ_{max}) is blue shifted, indicating protection of tryptophan residues from exposure to solvent.

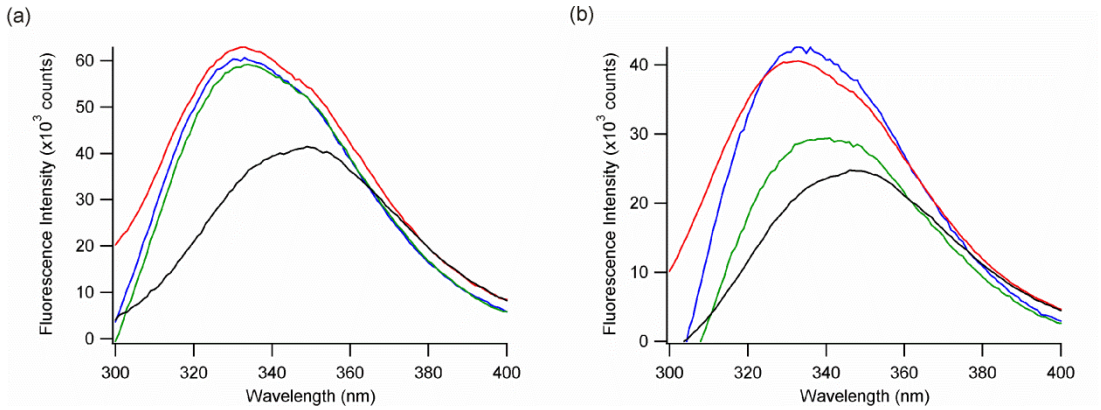

Figure 3.17 Fluorescence emission spectra of OmpT and OmpP folded in PC liposomes. Tryptophan emission spectra of (a) OmpT and (b) OmpP folded into 100 nm LUVs of *di*$C_{12:0}$PC (red line), *di*$C_{14:0}$PC (blue line) and *di*$C_{16:1}$PC (green line) or unfolded in 10 M urea in the presence of 100 nm LUVs of *di*$C_{12:0}$PC (black line). All spectra were acquired at a protein concentration of 0.4 μM an LPR of 3200:1 in 50 mM sodium phosphate, pH 8.0, 25 °C. OmpT samples were refolded in buffer containing 3 M urea and OmpP samples were refolded in buffer containing 5 M urea.

In order to remove the variation in spectral intensity that is inherent to pipetting liposomes[167], the intensity averaged wavelength, $\langle\lambda\rangle_{320-370}$ (Equation 2.4), was calculated for each spectrum (Table 3.3). For OmpT, the spectral $\langle\lambda\rangle$ is blue shifted upon folding into liposomes compared with the unfolded spectrum with all three lipids giving similar values of approximately 342.5 nm. OmpP folded into either *di*$C_{12:0}$PC or

di$C_{14:0}$PC has a similarly low $<\lambda>$, however, in *di*$C_{16:1}$PC the $<\lambda>$ is increased to 343.8 nm and the spectral intensity also appears significantly lower than in the other lipids tested. This could be reflective of the low yield of folded protein in *di*$C_{16:1}$PC (Table 3.2).

Lipid	OmpT $<\lambda>_{320–370}$ (nm)	OmpP $<\lambda>_{320–370}$ (nm)
di$C_{12:0}$PC	342.6	342.4
di$C_{14:0}$PC	342.3	342.1
di$C_{16:1}$PC	342.7	343.8
Unfolded	345.9	345.8

Table 3.3 Intensity averaged wavelengths ($<\lambda>_{320–370}$) of tryptophan fluorescence emission spectra of OmpT and OmpP folded in PC liposomes. All spectra were acquired in 100 nm LUVs at an LPR of 3200:1 in 50 mM sodium phosphate, pH 8.0, 25 °C. OmpT samples were refolded in buffer containing 3 M urea and OmpP samples were refolded in buffer containing 5 M urea. Unfolded spectra were acquired in 10 M urea in the presence of *di*$C_{12:0}$PC liposomes. $<\lambda>_{320–370}$ was calculated according to Equation 2.4.

Following the conclusion that upon folding into PC liposomes, tryptophan residues become protected from exposure to solvent, far-UV CD spectra were next measured of both OmpT and OmpP after folding for 16 h in 100 nm LUVs of *di*$C_{12:0}$PC, *di*$C_{14:0}$PC and *di*$C_{16:1}$PC in 50 mM sodium phosphate, pH 8.0, 25 °C (Figure 3.18). In all experiments, a protein concentration of 40 μM was used with an LPR of 400:1 in order to reduce the light scattering caused by the presence of liposomes and maximise the signal-to-noise ratio. The spectra of both OmpT and OmpP in all lipids shows a peak of negative ellipticity with a minimum at approximately 218 nm, indicative of β-sheet secondary structure. This peak is not observed in the spectrum of the unfolded state of either protein. In the case of OmpP folded into *di*$C_{16:1}$PC LUVs, the negative peak observed at 218 nm had a low intensity (Figure 3.18b), which may again be attributed to the low folding efficiency of OmpP under these conditions (Table 3.2). The far-UV CD spectra of OmpP were acquired by Remco N. P. Rodenburg (a visiting M.Sc. student from the University of Utrecht under the supervision of Lindsay M. McMorran).

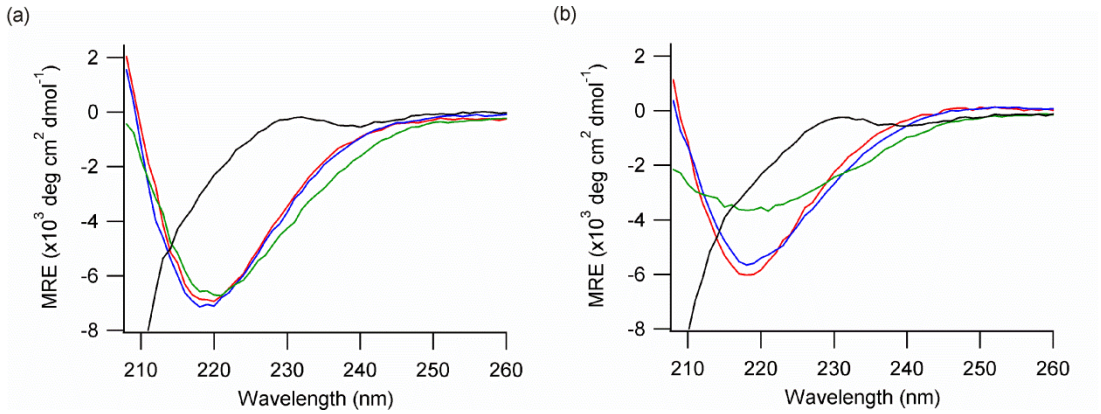

Figure 3.18 Far-UV CD spectra of OmpT and OmpP folded in PC liposomes. Spectra of (a) OmpT and (b) OmpP folded into 100 nm LUVs of *di*$C_{12:0}$PC (red line), *di*$C_{14:0}$PC (blue line) and *di*$C_{16:1}$PC (green line) or unfolded in 10 M urea in the absence of lipid (black line). All spectra were acquired at a protein concentration of 40 μM and an LPR of 400:1 in 50 mM sodium phosphate, pH 8.0, 25 °C. OmpT samples were refolded in buffer containing 3 M urea and OmpP samples were refolded in buffer containing 5 M urea. OmpP spectra were acquired by Remco N. P. Rodenburg (University of Utrecht).

The spectral data suggest that both OmpT and OmpP can fold to their native states in synthetic PC vesicles *in vitro*, and this was confirmed using an enzyme assay that was designed based on functional assays reported by Kramer *et al*[138] and Hagan *et al*[137]. Briefly, OmpT and OmpP were refolded into 100 nm LUVs of *di*$C_{12:0}$PC, *di*$C_{14:0}$PC and *di*$C_{16:1}$PC containing 0.8 mg ml^{-1} LPS at an LPR of 3200:1 in 50 mM sodium phosphate, pH 8.0, 25 °C. The protein was rapidly diluted into a 100 μM solution of the internally quenched fluorogenic peptide Abz-Ala-Arg-Arg-Ala-Tyr(NO_2)-NH_2 in 50 mM sodium phosphate, pH 7.0. The concentration of the peptide stock was measured using absorbance at 428 nm (A_{428}) with a molar extinction coefficient of 4200 M^{-1} cm^{-1}. Both OmpT and OmpP should cleave the peptide between the two arginine residues[245; 262], releasing the *ortho*-aminobenzoic acid (Abz) group from the nitrotyrosine quenching group. The reaction progress can be monitored by exciting the Abz group at 325 nm and measuring the increase in intensity of emission at 430 nm over time.

The increase in fluorescence intensity over time was measured for a variety of OmpT and OmpP concentrations using a large excess of substrate (Figure 3.19). The slope of the linear part of the line (0–30 s for OmpT and 0–300 s for OmpP) was used to

estimate the initial rate of the reaction and this was used to calculate the specific enzyme activity (nmol of substrate processed per μM enzyme per minute, $nmol.\mu M^{-1}.min^{-1}$) (see Equation 2.1) of each protein in each lipid (Figure 3.20). OmpP activity assays in *di*$C_{14:0}$PC and *di*$C_{16:1}$PC liposomes were conducted by Remco N. P. Rodenburg (University of Utrecht).

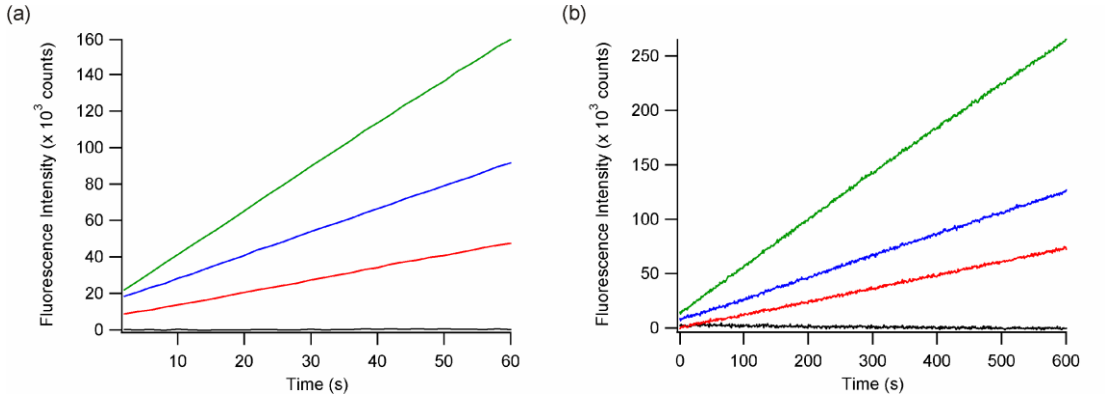

Figure 3.19 Enzyme activity assays of OmpT and OmpP in 100 nm *di*$C_{12:0}$PC LUVs. The fluorescence emission increase over time of Abz within an internally quenched peptide at 430 nm following excitation at 325 nm was monitored over time in the presence of (a) OmpT: 5 nM (red line), 10 nM (blue line), 20 nM (green line) or 25 nM in the absence of LPS (black line) and (b) OmpP: 180 nM (red line), 240 nM (blue line), 320 nM (green line) or 320 nM in the absence of LPS (black line). Both OmpT and OmpP were refolded for 16 h into 100 nm LUVs of *di*$C_{12:0}$PC containing 0.8 mg ml^{-1} LPS (where appropriate) at an LPR of 3200:1 in 50 mM sodium phosphate, pH 8.0, 25 °C before the reaction was initiated by rapid dilution into a 100 μM solution of Abz-Ala-Arg-Arg-Ala-Tyr(NO_2)-NH_2 in 50 mM sodium phosphate, pH 7.0.

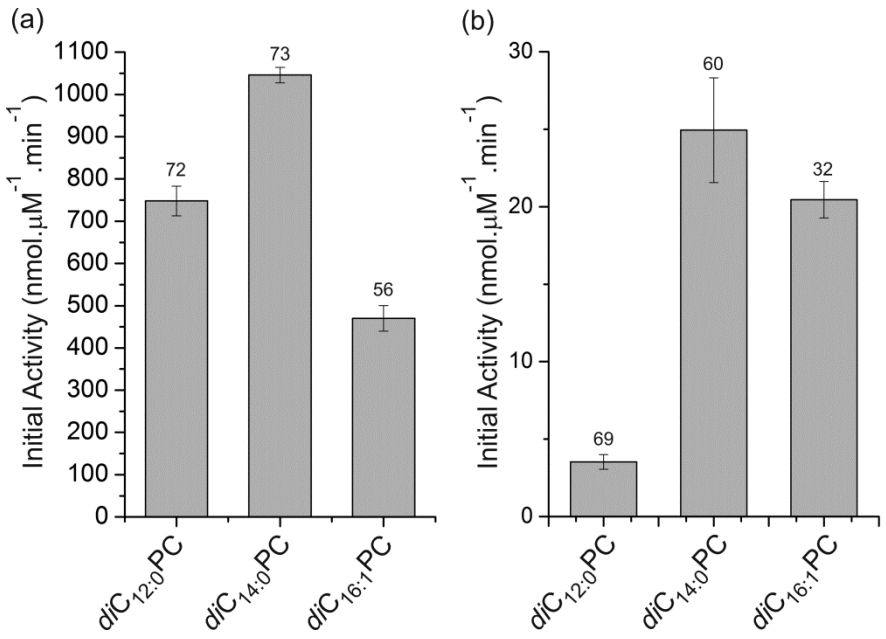

Figure 3.20 Initial enzymatic rates of OmpT and OmpP folded in PC liposomes. The fluorescence emission increase over time of Abz upon cleavage from an internally quenched peptide at 430 nm following excitation at 325 nm was monitored over time in the presence of (a) OmpT and (b) OmpP. The initial rate was calculated from the slope of the fluorescence traces according to Equation 2.1. Error bars indicate the standard error of the mean of the enzymatic activity measured over a range of enzyme concentrations. The number above each bar indicates the percentange of folded OMP estimated using cold SDS-PAGE. Both OmpT and OmpP were refolded for 16 h into 100 nm LUVs of *di*$C_{12:0}$PC, *di*$C_{14:0}$PC or *di*$C_{16:1}$PC containing 0.8 mg ml^{-1} LPS at an LPR of 3200:1 in 50 mM sodium phosphate, pH 8.0, 25 °C before the reaction was initiated by rapid dilution into a 100 μM solution of Abz-Ala-Arg-Arg-Ala-Tyr(NO_2)-NH_2 in 50 mM sodium phosphate, pH 7.0. The data for OmpP in *di*$C_{14:0}$PC and *di*$C_{16:1}$PC liposomes were acquired by Remco N. P. Rodenburg (University of Utrecht).

OmpT shows similarly high specific activities in the three lipids tested, however, OmpP activity was much lower (Figure 3.19b, Figure 3.20b), which is in part due to the slightly different substrate specificity of this enzyme[245; 262]. Comparing the activity of OmpP upon changing the liposome environment demonstrates that OmpP shows ten-fold higher activity in the thicker *di*$C_{14:0}$PC and *di*$C_{16:1}$PC bilayers (Figure 3.20b), in spite of the lower folding efficiency observed in cold SDS-PAGE assays compared with *di*$C_{12:0}$PC. These data demonstrate the ability of OmpT and OmpP to fold to their native state in synthetic liposomes of varying acyl chain length and saturation. Increasing the bilayer thickness (from *di*$C_{12:0}$PC to *di*$C_{14:0}$PC) does not affect the folding yield of either protein under the buffer conditions tested for these assays. Additionally, bilayer thickness does not seem to significantly affect OmpT enzyme

activity, however, the activity of OmpP increases significantly upon folding into thicker bilayers. Using lipids with unsaturated acyl chains (*di*$C_{16:1}$PC) causes a decrease in folding yield of both proteins, but has little effect on activity when compared with saturated lipids of similar hydrophobic thickness (*di*$C_{14:0}$PC).

3.2.5 Relative Stability of OmpT and OmpP in Synthetic Liposomes

The relative stabilities of OmpT and OmpP in 100 nm LUVs of *di*$C_{12:0}$PC, *di*$C_{14:0}$PC and *di*$C_{16:1}$PC were determined by using urea denaturation based on the method described by Huysmans *et al*[167]. Briefly, 1.2 μM OmpT or OmpP was allowed to refold for 16 h in 50 mM sodium phosphate, pH 8.0, 25 °C at an LPR of 3200:1 before a three-fold dilution into aliquots of increasing urea concentration (3–10 M). Refolding was carried out in an appropriate concentration of urea, as determined from the results of the cold SDS-PAGE screens (Table 3.2). These samples were allowed to equilibrate for a minimum of 8 h before the tryptophan emission spectrum of each sample was measured. The $\langle\lambda\rangle_{320\text{-}370}$ nm of each spectrum was calculated and plotted against the final urea concentration of that sample to monitor the unfolding transition (Figure 3.21). Urea denaturation experiments for OmpP were carried out by Remco N. P. Rodenburg (University of Utrecht).

Unfolding of both proteins from the native state in *di*$C_{12:0}$PC and *di*$C_{14:0}$PC liposomes does not go to completion, even after incubation in 10 M urea, and no post-transition baseline is observed (Figure 3.21). This suggests that these proteins are very stable under these conditions, even though the hydrophobic thickness of the two membranes differs by approximately 4 Å[261]. In contrast, unfolding from *di*$C_{16:1}$PC liposomes occurs much more readily: OmpP is fully unfolded at 5 M urea and OmpT at approximately 9 M urea. OmpP stability in *di*$C_{16:1}$PC is so low that no pre-transition baseline is observed and coupled with the low folding yields observed in the cold SDS-PAGE folding screens, these data suggest that although enzyme activity is detectable in this lipid only a small proportion of the OmpP in the folding reaction reaches and maintains the native state.

As OmpT unfolding from *di*$C_{16:1}$PC liposomes shows a full unfolding transition over the accessible range of urea concentrations, the denaturation midpoint ($[D]_{50\%}$) could be estimated. It should be noted that urea denaturation has thus far only been measured in the unfolding direction and it is therefore unknown if OmpT unfolding is fully reversible under these conditions. For this reason, a full thermodynamic analysis of OmpT unfolding is not possible; however, this will be addressed in Chapter 4. The apparent denaturant midpoint ($[D]_{50\%\ app}$) was estimated by first fitting the data to Equation 2.6 using a quantum yield correction (Q_R, Equation 2.5) of 1.25. The apparent values of $\Delta G^{\circ H_2O}_{UN\ app}$ and $M_{UN\ app}$ obtained from this fit can be used to calculate the apparent $[D]_{50\%}$ according to Equation 2.7. The apparent $[D]_{50\%}$ of OmpT unfolding from *di*$C_{16:1}$PC liposomes is 6.24 M, demonstrating that even though OmpT is destabilised in this lipid compared to *di*$C_{12:0}$PC or *di*$C_{14:0}$PC liposomes it is still relatively resistant to unfolding (Figure 3.21a). Due to the similar hydrophobic thickness of *di*$C_{14:0}$PC and *di*$C_{16:1}$PC bilayers[261], the markedly lower stability of both OmpT and OmpP in this lipid is most likely an effect of the difference in acyl chain saturation between these two lipids.

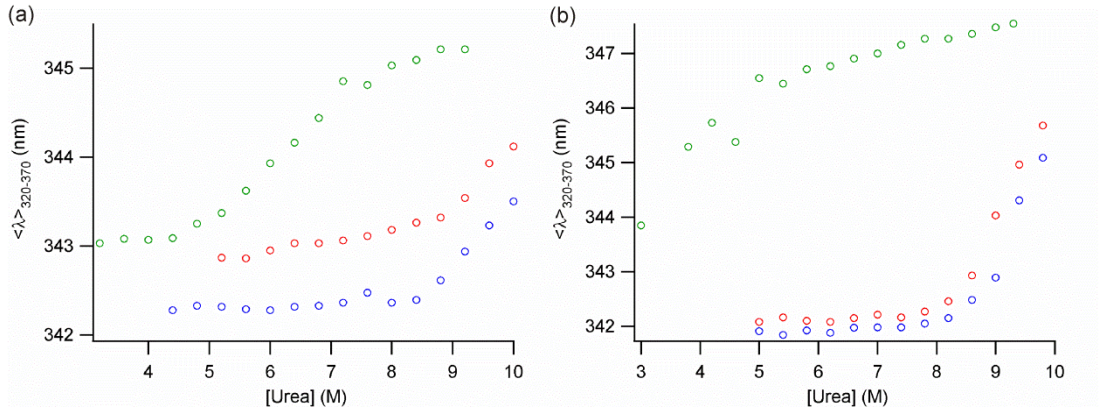

Figure 3.21 Urea denaturation curves of OmpT and OmpP in PC lipids. Urea denaturation of (a) OmpT and (b) OmpP in *di*$C_{12:0}$PC (red circles), *di*$C_{14:0}$PC (blue circles) and *di*$C_{16:1}$PC (green circles). All experiments were carried out at a final OMP concentration of 0.4 µM, in 100 nm liposomes, LPR 3200:1, 25 ºC, 50 mM sodium phosphate, pH 8.0. OmpP data were acquired by Remco N. P. Rodenburg (University of Utrecht).

Interestingly, the pre-transition baselines in *di*$C_{12:0}$PC and *di*$C_{14:0}$PC liposomes overlay for OmpP but are offset by approximately 0.4 nm for OmpT. The baseline for OmpT in *di*$C_{14:0}$PC liposomes occurs at a lower $<\lambda>_{320\text{-}370}$ than the other two lipids (Figure 3.21a), which would likely be caused by lower aggregation or a higher folded yield under these conditions. Interestingly, the results of the cold SDS-PAGE suggest that the folding yield of OmpT would be lower in *di*$C_{14:0}$PC than *di*$C_{12:0}$PC (Table 3.2), perhaps indicating that the ability of OMPs to withstand denaturation by SDS when the samples are unboiled could vary between individual OMPs and buffer conditions.

3.2.6 OmpT Folds to the Native State in an Amphipathic Polymer (Apol), A8-35

The use of liposomes provides a convenient method to examine the folding and stability of OMPs in a native-like lipid bilayer. Liposomes are, however, unsuitable for some biophysical techniques including mass spectrometry. Currently, the most commonly used method for the study of membrane proteins by mass spectrometry is solubilisation in detergent micelles, but this can lead to inactivation, unfolding and aggregation of the membrane protein under study[263; 264]. To overcome these problems, the use of Apols that bind strongly yet non-covalently to the transmembrane region of membrane proteins has become more prevalent[264]. Apol-bound membrane proteins have been shown to be more stable than micelle-encapsulated membrane proteins[264] and recently matrix-assisted laser desorption ionisation-mass spectrometry (MALDI-MS) was successfully used to determine the molecular mass of bR, OmpA, cytochrome b_6f and cytochrome bc_1 in the Apol A8-35 (Figure 1.18c)[265]. This approach, however, is limited as conformational information cannot be ascertained using this ionisation technique. The work described in this section was conducted in collaboration with Aneika C. Leney (University of Leeds) and aimed to develop methods for the study of Apol-bound membrane proteins by mass spectrometry.

Unfolded OmpT in 8 M urea was folded into A8-35 by dilution into a solution of Apol at a ratio of 1:5 OmpT:A8-35, and refolding was initiated by dialysis against 100 mM ammonium bicarbonate, pH 8.0. After folding had been allowed to proceed for 24 h, 4 °C, cold SDS-PAGE analysis was used to determine the folding efficiency of OmpT in A8-35, revealing that the protein migrates as a single, folded band (Figure 3.22a).

This was confirmed using analytical size exclusion chromatography, which also shows elution of the OmpT:A8-35 complex as a single peak, demonstrating that a single, folded species is present in solution (Figure 3.22b). Far-UV CD of the complex shows a peak of negative ellipticity with a minimum at approximately 218 nm (Figure 3.22c), comparable in magnitude to that of OmpT folded in PC liposomes (Figure 3.18a), and indicative of the present of β-sheet secondary structure. Acquisition of the native state of OmpT in A8-35 was confirmed using the fluorogenic peptide Abz-Ala-Arg-Arg-Ala-Tyr(NO_2)-NH_2 as described in Section 3.2.4, revealing that OmpT specific activity is approximately 700 nmol.μM.min^{-1} (Figure 3.22d) in the presence of LPS, comparable with the specific activity measured in PC liposomes (Figure 3.20a). Interestingly, a lower specific activity of approximately 100 nmol.μM.min^{-1} was measured in the absence of LPS (Figure 3.22d), which is not seen in PC liposomes (Figure 3.19a) and suggests that the interaction with A8-35 is slightly activating as well as stabilising.

Analysis of the OmpT:A8-35 complex by ESI-MS coupled with ion mobility spectrometry (IMS) by Aneika C. Leney (University of Leeds) revealed that OmpT is released from the amphipol complex and can be detected as multiply charged ions (Figure 3.23). A narrow charge state distribution of the 6+, 7+, 8+ and 9+ ions was observed, corresponding to a compact structure with a calculated mass of 33,462 ± 5 Da, which is within 0.01 % of the theoretical mass based on the amino acid sequence (33,460 Da). Analysis of the arrival time distribution of OmpT from ESI-IMS-MS revealed that the 6+ ion has a collisional cross section of 2601 $Å^2$, which is consistent with the value of 2718 $Å^2$ predicted from the PDB structure using the projected superposition approximation method. The agreement of the experimentally derived collisional cross section with the predicted value allows the conclusion that OmpT remains in a native-like conformation in the gas phase[266] and demonstrates the utility of amphipols as an alternative to a membrane mimetic for the study of native membrane proteins by a variety of techniques.

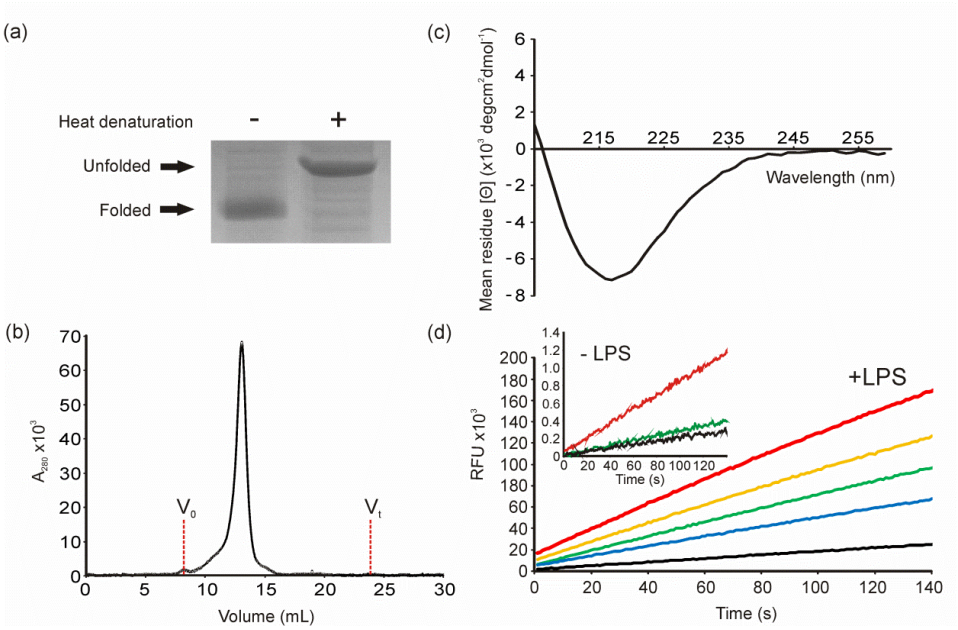

Figure 3.22 Folding of OmpT into A8:35. (a) SDS-PAGE of OmpT:A8-35 complex with and without heat denaturation; (b) size exclusion chromatogram showing a single peak corresponding to the OmpT:A8-35 complex. The void (V_o) and total column volume (V_t) are highlighted; (c) far-UV CD spectrum of the OmpT:A8-35 complex and (d) functional assay showing the fluorescence increase (relative fluorescence units) on enzymatic cleavage of the peptide Abz-Ala-Arg-Arg-Ala-Tyr-(NO_2)-NH_2 on addition of 0.05 µM (black), 0.10 µM (blue), 0.15 µM (green), 0.20 µM (yellow), and 0.30 µM (red) OmpT:A8-35 complex in the presence of LPS. The inset shows the weak catalytic activity of OmpT:A8-35 without LPS at OmpT:A8-35 concentrations of 0.05 µM (black), 0.15 µM (green) and 0.30 µM (red). This figure is taken from Leney *et al*, (2012)[267]. These data were acquired in collaboration with Aneika C. Leney (University of Leeds).

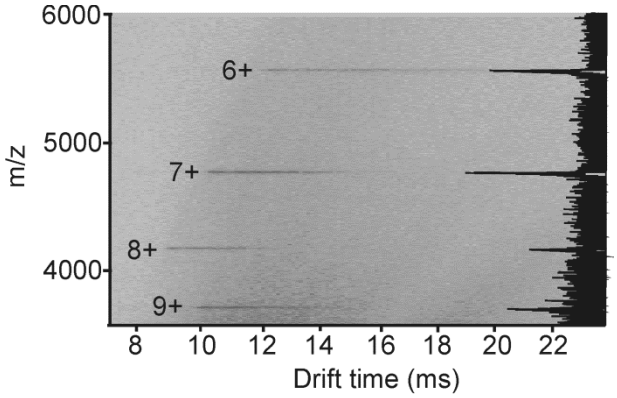

Figure 3.23 ESI-IMS-MS driftscope plot of the OmpT:A8-35 complex. The charge state of each ion is labelled and the summed *m/z* spectrum is displayed on the right-hand side. This figure is adapted from Leney *et al*, (2012)[267]. These data were acquired by Aneika C. Leney (University of Leeds).

3.3 DISCUSSION

Detailed analysis of OMP folding pathways has previously only been achieved for OmpA[144; 145; 147; 153; 161] and PagP[167; 238; 240], resulting in limited opportunities to compare and contrast the folding of different OMPs in the search for generic folding principles. In this study, the OM proteases, OmpT and OmpP, from *E. coli* have been successfully cloned, over-expressed, purified and refolded into liposomes composed of PC lipids of varying chain length and saturation. Additionally, protocols for the folding of OmpT into the Apol A8-35 have been developed in order to improve methodology for the study of functional membrane proteins by mass spectrometry. The results demonstrate the ability of OmpT and OmpP to fold to and maintain their native structure in changing membrane environments and the different stabilities of the two proteins suggest a role of competing evolutionary constraints in the folding of these two proteins.

3.3.1 ADAPTATION OF OMPT AND OMPP TO DIFFERENT LIPID ENVIRONMENTS

Cold SDS-PAGE was utilised in order to screen a variety of folding conditions assessing the effects of temperature and pH, vesicle size, as well as lipid acyl chain length and saturation on the folding efficiencies of OmpT and OmpP. The highest yield of folded protein obtained in *di*$C_{12:0}$PC LUVs for both proteins was in 50 mM sodium phosphate, pH 8.0, 25 °C, so these conditions were kept constant while the liposome composition was varied. The results demonstrate the ability of these OMPs to fold to their native, functional state in all three lipids tested (*di*$C_{12:0}$PC, *di*$C_{14:0}$PC and *di*$C_{16:1}$PC) suggesting an ability of the protein to adapt to a changing lipid environment. Such adaptation has been observed previously: a combination of X-ray crystallography data and MD simulations has shown that during the conformational changes required for the function of the sarco(endo)plasmic reticulum Ca^{2+}–ATPase, a hydrophobic mismatch occurs between the protein and the membrane[268]. This mismatch can be overcome by local deformation in the lipid bilayer around the protein as well as small rearrangements of the amino acid side chains and helix tilts of the protein[268]. These, or similar effects, could explain how OmpT and OmpP remain functional in synthetic liposomes with varying hydrophobic thicknesses.

Interactions between membrane proteins and the lipid bilayer have been well characterised, demonstrating the importance of individual amino acid side chains in stabilising the membrane-spanning regions in the bilayer. The requirement of LPS in the lipid membrane to activate OmpT protease activity is one example[258]. In general, the interactions of the protein, lipid and solvent in the complex interfacial region of the membrane are most important for determining placement of transmembrane segments[156; 269]. The effect of different residues on the positioning of α-helices in the membrane has been studied extensively using the "glycosylation mapping technique" (Figure 3.24), in which the active site of a membrane-bound endoplasmic-reticulum enzyme, oligosaccharyl transferase, is used as a point of reference to determine the relative position of an adjacent transmembrane helix[270]. Mutations in the transmembrane helix which change its position in the membrane will change the number of residues between the membrane surface and the glycosylation site of the oligosaccharyl enzyme, known as the minimal glycosylation distance[270]. The glycosylation mapping technique showed that the basic residues, arginine and lysine, which have long hydrophobic side chains, could reach along the face of the helix to position their positively charged groups into the negatively charged head-group region of the lipid, an effect called snorkelling[270]. This could lead to dramatic changes in the membrane positioning of the helix, and hence, the minimal glycoslyation distance. Aspartate and glutamate residues did not have such a marked effect on glycosylation distance, which was suggested to arise from their acidic side chains preventing snorkelling though unfavourable electrostatic interactions with the lipid head-groups. Instead, the relatively low pH at the membrane interface was proposed to promote the protonation of the side chains, allowing them to be buried in the transmembrane region[270].

The intrafacial region of both α-helical and β-barrel membrane proteins has been observed to be enriched with tryptophan and tyrosine residues, and has been dubbed the "aromatic girdle"[271; 272]. By contrast, phenylalanine prefers to be buried in the membrane core[271]. This was also demonstrated by glycosylation mapping experiments, which showed that both tryptophan and phenylalanine would push a transmembrane helix further into the membrane if they were positioned in a region normally outside the membrane[271]. When placed in the core of the membrane phenylalanine had no effect on membrane position, whereas tryptophan residues repositioned themselves in the

intrafacial region, causing the helix to be pushed out of the membrane[271]. The effect of tryptophan and tyrosine on minimal glycosylation distances was observed to be similar to that of arginine, suggesting these residues also exhibit snorkelling behaviour[271]. The preference of the polar aromatic residues for the membrane-water interface was studied using indole analogues and it was shown that hydrogen bonding and dipole interactions had little effect on membrane positioning[272]. Instead, it was suggested that the quadrupole interactions of the aromatic system would give rise to complex electrostatic interactions which favour the intrafacial region. The difference in behaviour of phenylalanine compared with tyrosine and tryptophan was attributed to the polar nature of the aromatic systems in the latter residues[272]. While these experiments elegantly demonstrate the ability of individual helices to move with respect to the membrane, and perhaps each other, the extensive lateral hydrogen bonding network present within the β-barrel of OMPs may prevent repositioning of individual secondary structural elements.

The effects of the positioning of tryptophan residues in the β-barrel of OmpA have been investigated more recently and begun to shed light on the interactions of individual amino acids with the lipid environment within the context of a β-barrel structure[273]. All residues in strand 1 of the OmpA β-barrel were first mutated to alanine in order to minimise interactions between side chains and increase the effect of lipid-exposure of the inserted tryptophan residues[273]. Each residue was in turn mutated to tryptophan and the effect on stability measured by carrying out urea denaturation under conditions which have previously been reported to yield reversible folding of OmpA[149; 273]. These mutations had little effect on folding efficiency of OmpA but mutants with tryptophan residues placed in the centre of the bilayer were shown to be slightly stabilised while those with tryptophan residues placed towards the edge of the bilayer were slightly destabilised[273]. The results suggest that while tryptophan residues show a preference for the intrafacial region when contained in a transmembrane segment which is free to move with respect to the bilayer, the more rigid structure of the OmpA β-barrel may limit the ability of tryptophan to reposition favourably within the complex electrostatic environment of the intrafacial region.

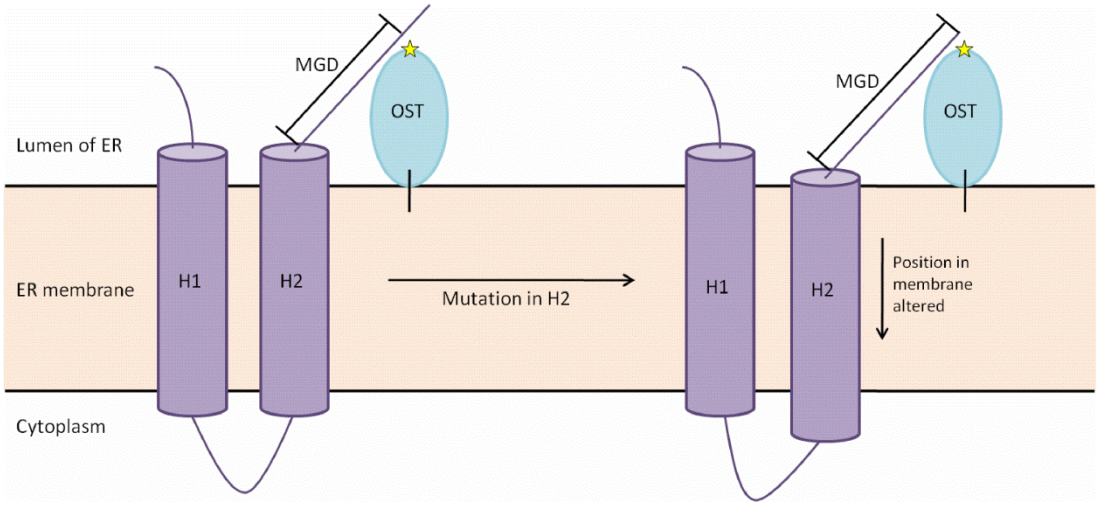

Figure 3.24 The "gylcosylation mapping technique". The endopasmic reticulum (ER) enzyme, oligosaccharide transferase (OST) is membrane-anchored. The number of residues between the membrane and the active site of OST (yellow star) is the minimum glycosylation distance (MGD). Helix 1 (H1) is a reference transmembrane helix, which is unchanged in the experiments. Mutations in Helix 2 (H2) may alter the position of H2 in the membrane, relative to H1, and this will change the MGD.

One mutation which did not follow the observed trend in the tryptophan mutants of OmpA was positioned in the periplasmic aromatic girdle of OmpA, and the anomalous stability of this mutant was thought to arise from stabilising interactions with nearby aromatic residues[273]. Indeed, the stabilising effects of aromatic residues in the OmpA periplasmic aromatic girdle has previously been measured using the same method. The study concluded that individual aromatic residues contribute much less to OmpA stability than clusters of residues[155]. Furthermore, OmpA stability depended more strongly on clustering of nearby aromatics than the type of aromatic residues present[155]. The clustering of aromatic residues at the intrafacial regions of β-barrel proteins may therefore arise not because positioning in this region is particularly stable, but from the highly stabilising nature of the side chain interactions of these residues. The ability of these residues to exist at the membrane-solvent interface or to be buried within the membrane core may help to explain the adaptability of OmpT and OmpP to membrane environments of varying hydrophobic thickness without losing functionality. Indeed, this ability to adapt to changing environments is so pronounced that OmpT can fold to and maintain its native state in the non-membrane mimetic environment of an amphipol as demonstrated in this study.

3.3.2 Differences in OmpT and OmpP Stability: Evidence of Fold Versus Function?

The urea denaturation experiments of OmpT and OmpP demonstrate that both of these OMPs are very resistant to urea denaturation in unsaturated *di*$C_{12:0}$PC or *di*$C_{14:0}$PC liposomes, in spite of the 4 Å difference in hydrophobic thickness of these bilayers[261]. In contrast, both OmpT and OmpP are destabilised in *di*$C_{16:1}$PC bilayers and since the hydrophobic thickness of this bilayer is not significantly different to that of *di*$C_{14:0}$PC[261], the destabilisation observed is most likely due to the unsaturated acyl chains in *di*$C_{16:1}$PC bilayers. Unsaturated lipids have kinks in their hydrocarbon chains causing the acyl chains to require more lateral space than the head-group of the lipid[274]. Within a bilayer, it is not possible for the two layers to bend away from each other, but the tendency for this to occur gives rise to the stored curvature stress of the membrane and is the origin for increased lateral pressure in membranes containing unsaturated lipids[274]. The effects of lateral pressure on the kinetics and thermodynamics of bR folding have been investigated and revealed that increased lateral pressure can improve the packing of tertiary structure and increase stability of bR, however, the kinetics of folding are slowed[275]. In the case of OmpT and OmpP, the increased lateral pressure in *di*$C_{16:1}$PC bilayers appears to have the opposite effect and is destabilising. This is in agreement with work that showed increasing lateral pressure in the bilayer by inclusion of PE lipids caused a decrease in stability of OmpA[149]. These differences in the response of α-helical and β-barrel membrane proteins to changing lateral pressure most likely arise due to the different topologies of these families. The β-barrel forms a very rigid structure, particularly in the region at the bilayer centre[45], and is unlikely to be able to conformationally adapt to relieve lateral pressure. In contrast, helix packing appears to be reinforced in the presence of increased lateral pressure[275] and the low lateral pressure in detergent micelles may explain the reported low stability of α-helical proteins in micellar systems[29]. Accordingly, the bacterial OM has been shown to be enriched in lipids with saturated acyl chains[276] and is thus expected to have a lower lateral pressure than the IM.

Interestingly, OmpP is much more destabilised in *di*$C_{16:1}$PC bilayers than OmpT and this could be a consequence of the differences in sequence of these homologous proteins. As illustrated in Figure 3.2, residues in OmpP which are not conserved in OmpT are clustered in two areas: around the active site and at the intrafacial regions, particularly in the periplasmic aromatic girdle. The differences in amino acid sequence

around the active site are presumably necessary in order to achieve the slightly different substrate specificities of OmpT and OmpP[245; 262]. Assuming these mutations cause destabilisation of OmpP, the three extra aromatic residues, tryptophan-60, phenylalanine-117 and phenylalanine-191 in the periplasmic aromatic girdle which are not present in OmpT may have evolved to offset the destabilisation due the mutations in the active site residues. The fact that OmpP remains less stable than OmpT in spite of the reinforcement of the periplasmic aromatic girdle highlights the competing evolutionary constraints of maintaining a stable fold while optimising function in these homologous proteins.

The concept of "fold versus function" has not previously been highlighted in the context of membrane proteins, perhaps due to the lack of detailed information on the folding pathways of these proteins or the few comparative studies available. It is well established, however, in the field of soluble protein folding. One example of this is the bacterial immunity protein, Im7, in which formation of non-native interactions has been observed in each of the two folding transition states *en route* to the native state[201; 277]. Many of the residues which form these non-native contacts have been shown to be functionally important for colicin binding, and so the unusually rugged folding landscape arises due to functional constraints[201]. Another study which uncovered the competing relationship of protein fold and function looked at the conserved glycine-48 residue in the SH3 domain (a domain found in a large number of eukaryotic signalling proteins)[278]. All mutations at residue 48 led to destabilisation of the folded protein but also caused a large increase in the folding rate of the protein[278]. In addition, many of the mutants showed decreased binding affinity to a target peptide[278]. Overall, this suggests that glycine-48 is a highly conserved functional residue that actually slows the folding of the SH3 domain. The data presented in this chapter alongside the homology model of OmpP suggest that a "fold versus function" tension may affect the folding and stability of this protein. Further work will be required, however, in order to conclusively show whether the larger decrease in stability of OmpP in *di*$C_{16:0}$PC liposomes compared with that of OmpT is a direct consequence of functional constraints.

In this chapter, relative stability of OmpT and OmpP was examined by comparing urea denaturation curves in the unfolding direction. In order to conduct a more rigorous thermodynamic analysis and determine the free energy of unfolding of a protein, it is

required that the protein must fold reversibly. Determining conditions under which folding of these OMPs is reversible would allow further examination of the "fold versus function" tension in OmpT and OmpP and also allow information to be gained on the folding mechanism of these proteins. Accordingly, the work in the next chapter focuses on OMP reversibly and the origins of irreversible folding.

4 Is There a Link Between Folding Reversibility and Lipid Adhesion in the Unfolded State?

4.1 Introduction

4.1.1 Measuring Protein Stability

The relative folding stabilities of OmpT and OmpP were measured in Section 3.2.5 by urea titration in the unfolding direction only, that is, folded OMP in liposomes was unfolded by adding increasing concentrations of urea. While this comparative approach allowed conclusions on the relative stability of these proteins in liposomes of differing composition to be drawn, a more rigorous approach is required in order to determine the standard free energy of unfolding (ΔG^o_{UN}) of a protein. In a simple, two-state folding mechanism where only the unfolded (U) and native (N) states of the protein are significantly populated, the folding and unfolding transitions can be characterised by the folding rate and unfolding constants k_f and k_u, respectively[279], as shown in the scheme below:

$$U \underset{k_u}{\overset{k_f}{\rightleftharpoons}} N$$

When the system is at equilibrium, then the concentration of N and U is described by:

$$k_u[N] = k_f[U] \qquad \text{Equation 4.1}$$

And the equilibrium constant for folding, K_{UN}, can be calculated by:

$$K_{UN} = \frac{[N]}{[U]} = \frac{k_f}{k_u} \qquad \text{Equation 4.2}$$

According to the van 't Hoff equation, the free energy of folding (ΔG^o_{UN}) can be connected to the equilibrium constant as follows:

$$\Delta G^o_{UN} = -RT \ln K_{UN} = -RT \ln\left(\frac{k_f}{k_u}\right) \qquad \text{Equation 4.3}$$

Thus, in a two-state system protein stability can be determined by either kinetic or thermodynamic analysis since both approaches yield the same free energy.

The free energy of unfolding is linearly proportional to the denaturant concentration, and so the free energy of unfolding in the absence of denaturant ($\Delta G^{\circ H_2O}_{UN}$) can be calculated[280] according to:

$$\Delta G^{o}_{UN} = \Delta G^{\circ H_2O}_{UN} - M_{UN}[D] \qquad \text{Equation 4.4}$$

where M_{UN} is the m-value for the equilibrium folding transition and [D] is the concentration of denaturant.

Equation 4.3 and Equation 4.4 can be combined with Equation 2.5 to yield Equation 2.6, which allows the fitting of spectroscopic data describing the unfolding of the protein of interest with urea to determine the standard free energy of unfolding.

4.1.2 Overcoming Hysteresis in OMP Equilibrium Studies

The approach outlined in Section 4.1.1 relies on the reversible folding of the protein of interest in order to satisfy the definition of a two-state folding mechanism and this has resulted in few reports to date on OMP stability (Table 4.1)[149; 150; 151; 165; 167; 174]. The low number of successful applications of this approach to membrane protein folding has been attributed to the difficulties in handling unfolded membrane proteins *in vitro*[152; 164]. Hysteresis of denaturation curves in the folding and unfolding directions is commonly observed for OMPs (Figure 4.1), which may arise from aggregation funnelling some of the protein of interest off the folding pathway and preventing its return to a folding competent state as demonstrated by Moon *et al*[165]. Another possible explanation is that folded OMPs are subject to a very high kinetic barrier to unfolding. If this is the case, hysteresis of the folding and unfolding curves should disappear if the system is given sufficient time to equilibrate, as demonstrated for OmpA folding into both LDAO micelles and the amphipol, A8-35[151]. The concept of "infinite kinetic stability" is not unique to OMPs and has previously been observed as a barrier to dissociation of the subunits of the *E. coli* Type 1 pilus[281] and as a barrier to unfolding in the meta-stable

α-lytic protease[202; 203]. Given the range of essential functions carried out by OMPs *in vivo*[46; 64; 282] it is not unreasonable to assume they are subject to similarly high kinetic barriers to unfolding. From the few successful reports of reversible OMP folding, it is clear that no consensus has been reached on conditions which reduce aggregation and promote reversibility, reflecting the current lack of understanding of these processes.

Previously, conditions were found under which a full unfolding transition of OmpT could be measured in *di*$C_{16:1}$PC liposomes (LPR 3200:1, 50 mM sodium phosphate, pH 8.0, 25 °C, see Section 3.2.5). Here, conditions were screened to attempt to establish reversible folding of OmpT, however, this was unsuccessful. In the process of this investigation, a link was established between the unfolding kinetics and two-state reversible folding of OMPs, leading to the hypothesis that folding reversibility may be achieved where there is adhesion of the OMP to the liposome in the unfolded state. This relationship was shown to be upheld for OmpT, PagP and HT PagP under a variety of conditions. The validity of the hypothesis was tested by incorporating dansyl-labelled guest lipid into liposomes of *di*$C_{12:0}$PC and *di*$C_{16:1}$PC (Figure 4.2) and using the resulting FRET to tryptophan residues in PagP and HT PagP to follow the unfolding reaction. This analysis revealed that residual FRET is observed at the end of a reversible folding reaction that is not present at the end of an irreversible folding reaction, providing evidence that lipid adhesion in the unfolded state may be one mechanism by which folding reversibility is mediated.

Protein	Membrane Mimetic	Buffer Conditions	Free Energy of Folding (kJ mol^{-1})	M-value of Folding (kJ mol^{-1} M^{-1})	Reference
OmpA	30 nm SUVs, 92.5:7.5 $C_{16:0}C_{18:1}PC$:$C_{16:0}C_{18:1}PG$ (800:1 excess)	10 mM glycine, 2 mM EDTA, pH 10.0, 37.5 °C Urea denaturant, 16 h	-14.2	4.60	Hong and Tamm (2004)[149]
OmpA	Octyl maltoside micelles (100,000:1 excess)	10 mM glycine, 2 mM EDTA, pH 10.0, 25 °C GuHCl denaturant, 4 days	-65.2	12.5	Andersen *et al* (2012)[150]
OmpA	LDAO micelles (800:1 excess)	10 mM sodium borate, 2 mM EDTA, pH 10.0, 40 °C Urea denaturant, 52 days	-59.8	Not Reported	Pocanschi *et al* (2013)[151]
OmpA	Amphipol A8-35 (65:1 excess)	10 mM sodium borate, 2 mM EDTA, pH 10.0, 40 °C Urea denaturant, 25 days	-7.94	Not Reported	Pocanschi *et al* (2013)[151]
HT PagP	100 nm LUVs, *di*$C_{12:0}PC$ (3200:1 excess)	50 mM sodium phosphate, pH 8.0, 25 °C Urea denaturant, 16 h	-60.2	6.86	Huysmans *et al* (2010)[167]
OmpLa	100 nm LUVs, *di*$C_{12:0}PC$* (2000:1 excess)	100 mM sodium citrate, 2 mM EDTA, pH 3.8, 37 °C GuHCl denaturant, 5 h	-136	38.5	Moon *et al* (2011)[165]
PagP	100 nm LUVs, *di*$C_{12:0}PC$* (2000:1 excess)	100 mM sodium citrate, 2 mM EDTA, pH 3.8, 37 °C GuHCl denaturant, 5 h	-102	22.6	Moon *et al* (2013)[174]
OmpW	100 nm LUVs, *di*$C_{12:0}PC$* (2000:1 excess)	100 mM sodium citrate, 2 mM EDTA, pH 3.8, 37 °C GuHCl denaturant, 5 h	-76.5	18.8	Moon *et al* (2013)[174]

Table 4.1 Reported free energies of folding of OMPs to date. The excess of membrane mimetic in each reaction is expressed as the mol:mol ratio of mimetic:OMP. *Indicates the presence of 1.5 mM 3-(N,N-dimethylmyristyl-ammonio)propanesulfonate during initial dilution from the GuHCl unfolded state.

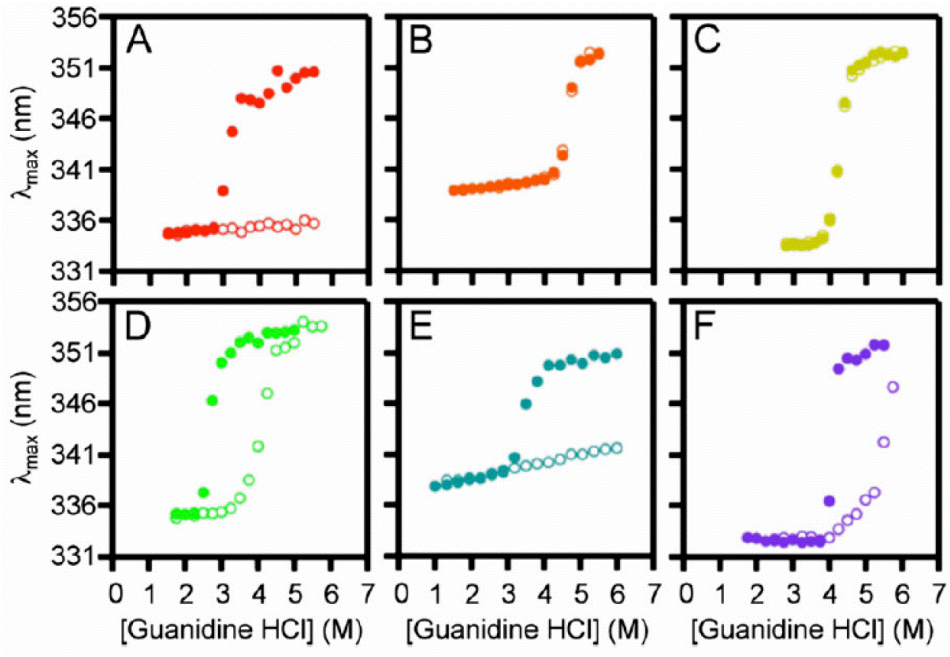

Figure 4.1 OMP denaturation curves in the folding and unfolding directions often display hysteresis. Denaturation of (A) OmpX, (B) PagP, (C) OmpW, (D) OmpA, (E) OmpT and (F) FadL in the unfolding (open circles) and folding (filled circles) directions. All experiments were performed in 100 mM citrate, 2 mM EDTA, pH 3.8, 37 °C for 40 h in 100 nm LUVs of *di*$C_{12:0}$PC at an LPR of 2000:1 and contained 1.5 mM 3-(N,N-dimethylmyristyl-ammonio)propanesulfonate during initial dilution from the GuHCl unfolded state λ_{max} corresponds to the wavelength at which maximum tryptophan fluorescence emission is observed. This figure is taken from Moon *et al* (2013)[174].

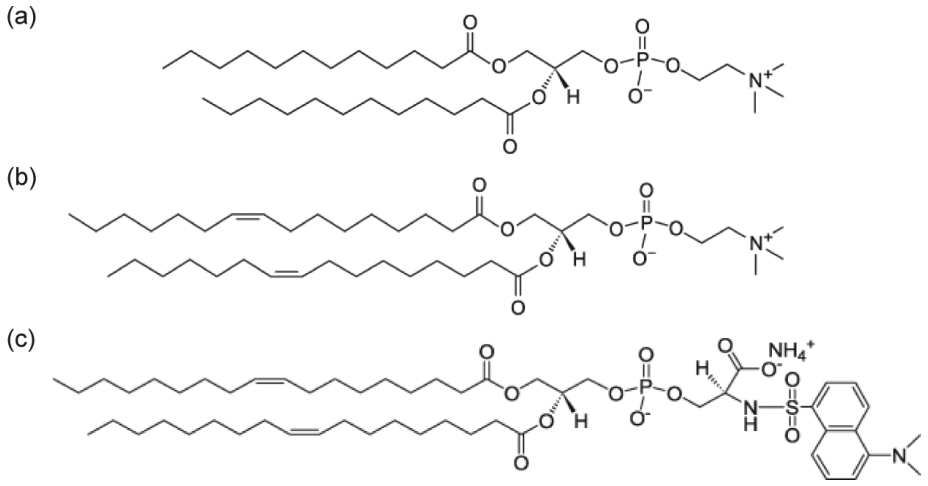

Figure 4.2 Structures of lipids used in this chapter. (a) 1,2-dilauroyl-*sn*-glycero-3-phosphocholine (*di*$C_{12:0}$PC); (b) 1,2-dipalmitoleoyl-*sn*-glycero-3-phosphocholine (*di*$C_{16:1}$PC) and (c) 1,2-dioleoyl-*sn*-glycero-3-phospho-L-serine-N-(5-dimethylamino-1-naphthalenesulfonyl) (*di*$C_{18:1}$PS-dansyl).

4.2 RESULTS

4.2.1 TESTING OMPT FOLDING REVERSIBILITY

The urea denaturation experiments of OmpT and OmpP in the unfolding direction (Section 3.2.5) revealed that in 100 nm LUVs of *di*$C_{12:0}$PC and *di*$C_{14:0}$PC both proteins are so stable that they do not fully unfold. In contrast, in 100 nm *di*$C_{16:1}$PC LUVs both proteins are destabilised. The folded yield of OmpP is so low that no pre-transition baseline is observed in the denaturation experiments and so its folding reversibility could not be analysed under these conditions. A full transition was observed upon unfolding of OmpT from 100 nm *di*$C_{16:1}$PC LUVs and so folding reversibility was investigated with the aim of calculating the free energy of unfolding.

Reversibility experiments were conducted according to the method described by Huysmans *et al*[167]. For the unfolding direction, 1.2 μM OmpT was allowed to fold for 16 h in 3 M urea, 50 mM sodium phosphate, pH 8.0, 25 °C at an LPR of 3200:1 before a three-fold dilution into aliquots of increasing urea concentration (3.2–9.2 M). These samples were allowed to equilibrate for a minimum of 8 h before the tryptophan emission spectrum of each sample was measured. Samples in the refolding direction were made by folding 4.8 μM OmpT for 16 h under identical buffer conditions before a four-fold dilution into 9.2 M urea for a minimum of 8 h to unfold the protein. A three-fold dilution was then carried out to refold samples at urea concentrations of 3.2–9.2 M, which were allowed to equilibrate for a minimum of 8 h before the tryptophan emission spectrum of each sample was measured. The $\langle\lambda\rangle_{320\text{-}370}$ of each spectrum was calculated and plotted against the final urea concentration of that sample, revealing that under these conditions OmpT denaturation curves display hysteresis (Figure 4.3a). The observed hysteresis shows that in the refolding direction, many samples have a higher value of $\langle\lambda\rangle_{320\text{-}370}$ than the corresponding sample in the unfolding direction at the same concentration of urea.

OmpT aggregation upon dilution of urea in the refolding step could skew the value of $\langle\lambda\rangle_{320\text{-}370}$ to higher wavelengths and seems a likely source of hysteresis. To

attempt to reduce aggregation and hence promote reversibility without hysteresis, the urea denaturation scheme was amended so that final samples contained 0.1 μM OmpT while the LPR was fixed at 3200:1. All equilibration times and buffer conditions were unchanged from the initial experiment. The resulting denaturation curves also exhibit hysteresis and show no improvement on the denaturation experiments performed at higher OmpT concentration (Figure 4.3b) suggesting that if hysteresis does occur due to aggregation that this effect is not dependent on protein concentration in the range of concentrations tested.

Aggregation may also occur when the unfolded OmpT stock is diluted from 6 M GuHCl into 3 M urea to initiate folding at the beginning of the experiments, resulting in a population of aggregated protein present in all samples which could affect the calculated value of $\langle\lambda\rangle_{320\text{-}370}$ in a urea-dependent manner. A new scheme was employed where OmpT unfolded in 6 M GuHCl was first diluted into 10 M urea in the presence 100 nm *di*$C_{16:1}$PC liposomes at an LPR of 3200:1 and allowed to equilibrate overnight. A four-fold dilution into aliquots of decreasing urea concentration (3.2–9.2 M) was then carried out to create samples in the refolding direction. For the unfolding direction, the initial equilibration at 10 M urea in the presence of liposomes was identical. A four-fold dilution into 3 M urea was then used to initiate refolding, and following equilibration for at least 8 h, a second four-fold dilution into aliquots of increasing urea concentration (3.2–9.2 M) was carried out. Buffer conditions were unchanged (50 mM sodium phosphate, pH 8.0) from previous experiments and all incubations were at 25 °C. The resulting curves demonstrated the largest discrepancy observed thus far between the folding and unfolding directions (Figure 4.3c). Note that in the folding direction, changing urea concentration has little effect on $\langle\lambda\rangle_{320\text{-}370}$, in contrast with other conditions. The reasons for this are unclear and it is not known which conformation OmpT is in at the start of this reaction.

Previously, samples made in both the refolding and unfolding directions were subjected to an initial folding step such that the starting point for both sets of experiments is folded OmpT in *di*$C_{16:1}$PC liposomes. In this scheme, this step was omitted and the equilibrium between the initial folding step and unfolding measured. Based on the poor results obtained, it was decided to return to the previous

experimental method to include the initial folding step, however, the unfolded stock of OmpT was first buffer exchanged into 10 M urea by gel filtration using a Zeba Spin Column (Thermo Scientific). All other experimental details were unchanged. This method also resulted in hysteresis of the folding and unfolding denaturation curves, suggesting that reversibility without hysteresis may not be achievable under these conditions.

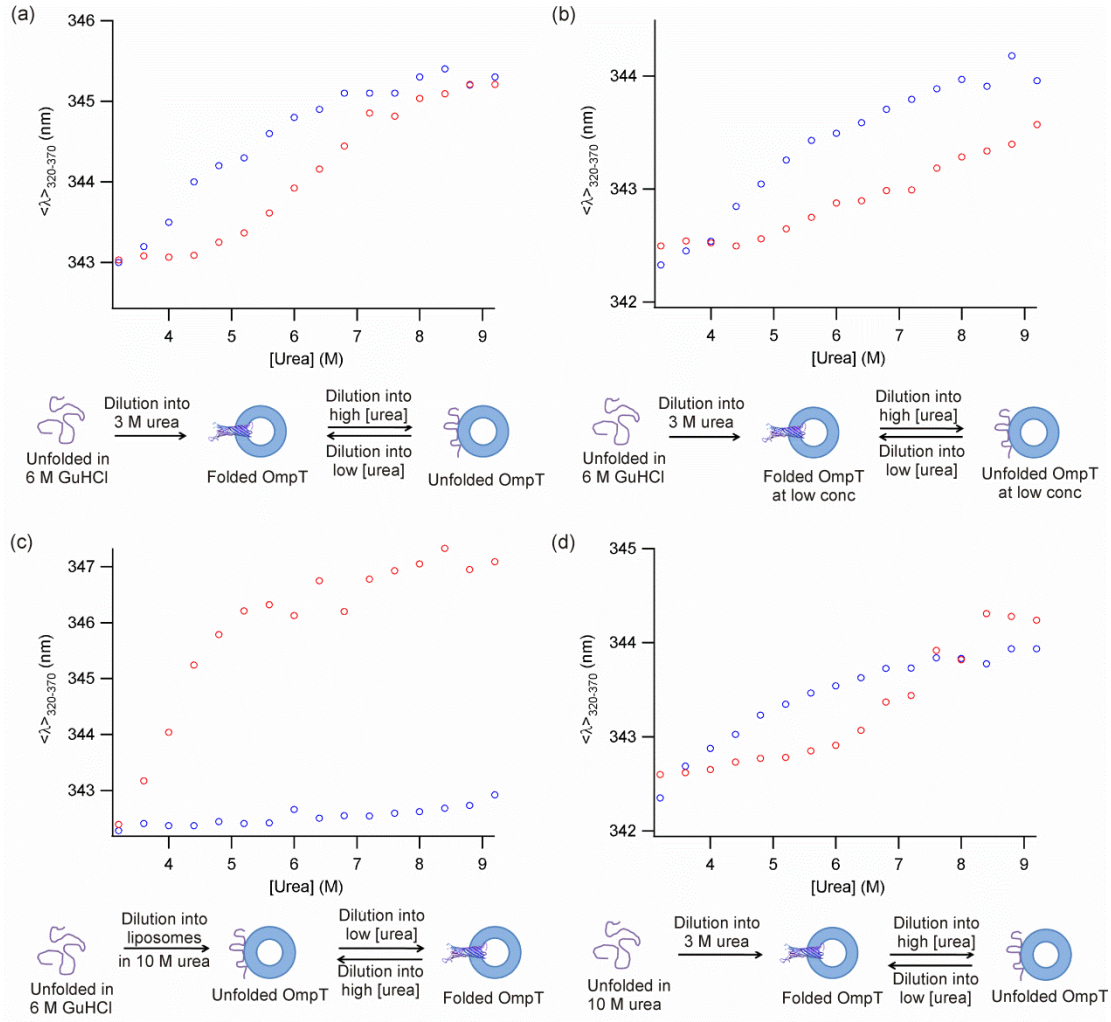

Figure 4.3 OmpT denaturation curves measured in the folding and unfolding directions exhibit hysteresis. Denaturation curves of (a) 0.4 μM OmpT from an unfolded stock in 6 M GuHCl, (b) 0.1 μM OmpT from an unfolded stock in 6 M GuHCl, (c) 0.1 μM OmpT from an unfolded stock in 6 M GuHCl that was diluted into 10 M urea before folding was initiated and (d) 0.1 μM OmpT from an unfolded stock in 10 M urea. The unfolding curve is shown in red and the folding curve is shown in blue. Dilution schemes for each experiment are shown below the appropriate panel. All experiments were carried out in 100 nm *di*$C_{16:1}$PC LUVs at an LPR of 3200:1, 50 mM sodium phosphate, pH 8.0, 25 °C.

Following the study by Moon *et al*[165], which reported reversible folding of OmpLa without hysteresis at acidic pH, it was decided to test a similar method to see if OmpT folding is also reversible without hysteresis under these conditions. First, a cold SDS-PAGE folding screen was carried out to examine the ability of OmpT to fold at pH 3.8 in 100 nm *di*$C_{16:1}$PC LUVs containing different concentrations of GuHCl. Following 16 h refolding time at 37 °C, the samples were analysed revealing folding yields of approximately 50 % under these conditions (Figure 4.4a, Table 4.2). Equilibrium denaturation experiments were then conducted to determine whether folding reversibility could be achieved under these conditions. Unfolded OmpT in 6 M GuHCl was mixed with 100 nm *di*$C_{16:1}$PC LUVs in 50 mM sodium citrate, pH 3.8 in either 6 M GuHCl (refolding direction) or 1.5 M GuHCl (unfolding direction) and allowed to equilibrate for 5 h at 37 °C. These reactions were then diluted four-fold into aliquots containing different GuHCl concentrations (1.5–6.0 M) and allowed to equilibrate for 36 h at 37 °C before the tryptophan fluorescence spectrum of each sample was measured. The resulting denaturation curves (Figure 4.4b) demonstrate that this method does not result in folding reversibility without hysteresis for OmpT and suggests the observed hysteresis is not a pH-dependent effect. Subsequent, similar experiments by the Fleming group concur with this observation (Figure 4.1E)[174]

[GuHCl] (M)	Folding Yield (%)
1.0	40
1.5	49
2.0	53
2.5	50

Table 4.2 Folding yield of OmpT at acidic pH determined using cold SDS-PAGE. All samples contained 100 nm *di*$C_{12:0}$PC LUVs at an LPR of 3200:1 and 50 mM sodium citrate buffer, pH 3.8, 37 °C.

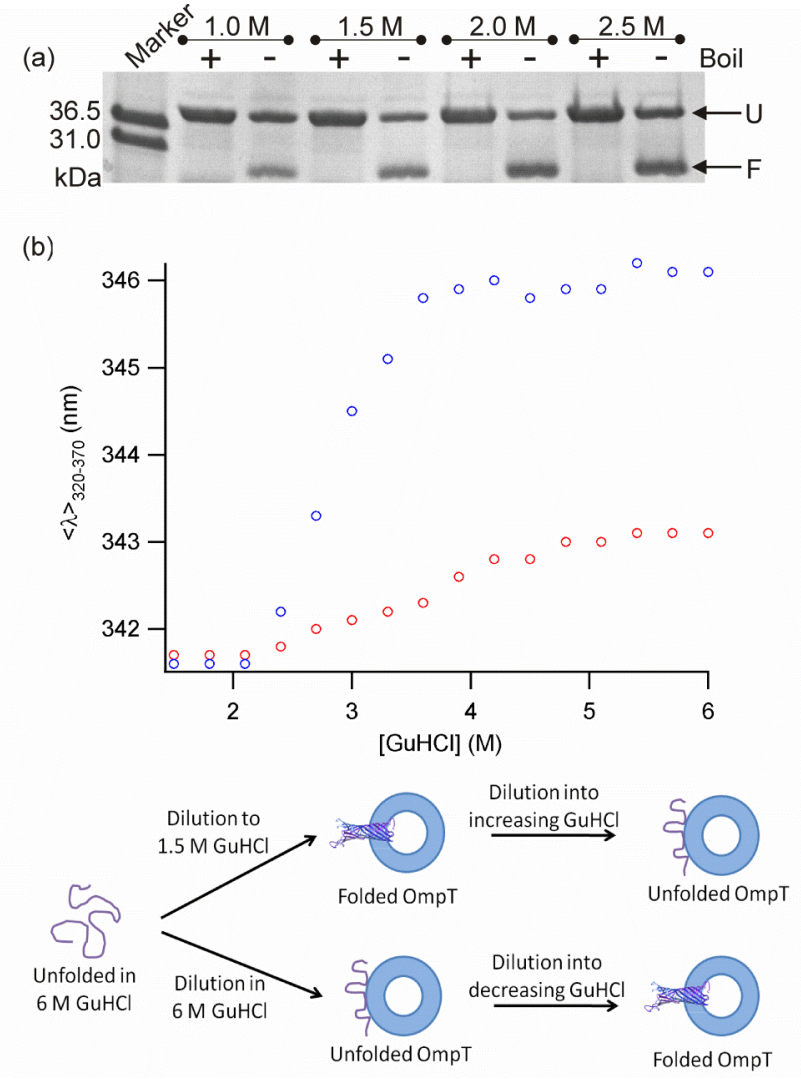

Figure 4.4 Testing OmpT folding and reversibility at acidic pH. (a) cold SDS-PAGE folding screen showing the folded (F) and unfolded (U) conformations and (b) urea denaturation in the folding (blue) and unfolding (red) directions. A dilution scheme is provided in the lower panel. SDS-PAGE experiments contained 4 µM OmpT and denaturation samples contained 0.5 µM OmpT. SDS-PAGE samples are labelled according to the final concentration of GuHCl. All samples contained 100 nm *di*$C_{16:1}$PC LUVs at an LPR of 3200:1 and 50 mM sodium citrate buffer, pH 3.8, 37 °C.

4.2.2 A Short Unfolding Method Reveals Double Exponential Unfolding Kinetics of OmpT

The results obtained thus far suggest that the hysteresis observed in previous equilibrium denaturation experiments could not be overcome by employing methods

to reduce aggregation in the urea denaturation samples. In the work of Huysmans *et al*[167], folding of HT PagP could be made reversible without hysteresis. These experiments were thought to measure the equilibrium between the folded, membrane-inserted form of HT PagP and an unfolded, lipid-adhered form. Assuming this to be correct, hysteresis in OmpT folding may occur if some of the OmpT does not remain adhered to the liposome in the unfolded state, giving rise to two different unfolded populations. Such a situation would be expected to give rise to parallel refolding pathways, and indeed, kinetic analysis of the HT PagP folding reaction revealed that at high LPR, where folding is reversible without hysteresis, folding kinetics display a burst phase followed by a single exponential phase[240]. Lowering the LPR causes hysteresis of folding and unfolding to occur and correspondingly, a second, slower exponential phase is observed[240]. This was proposed to arise from a second population of HT PagP which is not initially lipid-adhered[240]. If a similar scenario exists upon OmpT refolding, the population which is not lipid-adhered may require a long time to adhere and refold, or may never refold due to a loss of folding competency. In the latter case, if loss of folding competency were due to aggregation then the extent of aggregation and hence degree of hysteresis should be dependent upon protein concentration. Either situation could be expected to result in hysteresis of the folding and unfolding denaturation curves.

While it is not experimentally feasible to allow very long incubation times in a liposome-based folding system due to the lifespan of liposomes in aqueous solution being limited to approximately 5 days[158], it may be possible to circumvent the formation of two unfolded OmpT populations. If unfolding is allowed to proceed to an unfolded, lipid-adhered state then refolding is immediately initiated before dissociation from the liposome can take place, it may be possible to set up a "pseudo-equilibrium" which could be used to measure OmpT stability. To employ this method, the unfolding time of OmpT must first be determined kinetically. OmpT was refolded in 3 M urea, 50 mM sodium phosphate, pH 8.0 in 100 nm *di*$C_{16:1}$PC LUVs for 16 h, before a ten-fold dilution into 9 M urea was used to initiate unfolding. The change in tryptophan fluorescence emission at 335 nm was measured to follow the progress of the reaction, revealing that OmpT unfolding occurs rapidly and is completed after around 500 s (Figure 4.5a). A downward slope is observed at

the plateau of the reaction, which could be due to photobleaching of the sample or to an additional slow unfolding phase.

Equilibrium denaturation was next carried out using the kinetic data to set the unfolding time of the samples accordingly. For the unfolding direction, 1.2 μM OmpT was allowed to fold for 16 h in 3 M urea, 50 mM sodium phosphate, pH 8.0, 25 °C at an LPR of 3200:1 before a three-fold dilution into aliquots of increasing urea concentration (3.2–9.2 M). These samples were allowed to equilibrate for 500 s before the tryptophan emission spectrum of each sample was measured. Samples in the refolding direction were made by folding 4.8 μM OmpT for 16 h under identical buffer conditions before a four-fold dilution into 9.2 M urea to unfold the protein. After 500 s equilibration, a three-fold dilution was then carried out to refold samples at urea concentrations of 3.2–9.2 M, which were allowed to equilibrate for a minimum of 8 h before the tryptophan emission spectrum of each sample was measured. The results reveal that even very brief unfolding times give rise to a large degree of hysteresis between the folding and unfolding denaturation curves (Figure 4.5b).

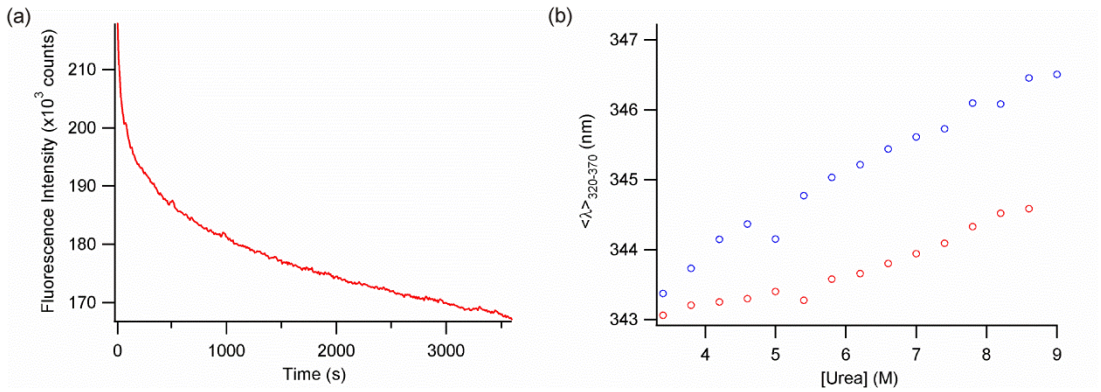

Figure 4.5 OmpT equilibrium denaturation using a short unfolding method. (a) OmpT unfolding kinetics in 9 M urea and (b) OmpT urea denaturation in the unfolding (red) and refolding (blue) directions. All experiments contained 0.4 μM OmpT and were carried out in 100 nm *di*$C_{16:1}$PC LUVs at an LPR of 3200:1, 50 mM sodium phosphate, pH 8.0, 25 °C.

Next, it was decided to analyse the unfolding kinetics of the OmpT unfolding reaction in more detail. The unfolding trace was first fitted (see Section 2.5.7) to a

single exponential function (Figure 4.6a, top, Table 4.3), with a linear term included to account for the slope in the baseline. Visual inspection of the fit and analysis of the fitting residuals (Figure 4.6a, bottom) show that the data are not well described by this function, and so a fitting to a double exponential function including a linear term was carried out instead (Figure 4.6b, top, Table 4.3). This fit was found to satisfactorily describe the data as judged by the fitting residuals (Figure 4.6b, bottom) and resulted in a large, ten-fold difference in the two unfolding rate constants obtained (Table 4.3).

Fit	k_1 ($\times 10^{-2}$ s^{-1})	Fitting Error ($\times 10^{-2}$ s^{-1})	k_2 ($\times 10^{-3}$ s^{-1})	Fitting Error ($\times 10^{-3}$ s^{-1})
Single exponential	–	–	3.85	0.12
Double exponential	2.63	0.08	1.88	0.04

Table 4.3 OmpT rate constants obtained from exponential fits of unfolding data in *di*$C_{16:1}$PC LUVs, 50 mM sodium phosphate, pH 8.0. All exponential fits included a linear term to account for the downward slope in the plateau of the data.

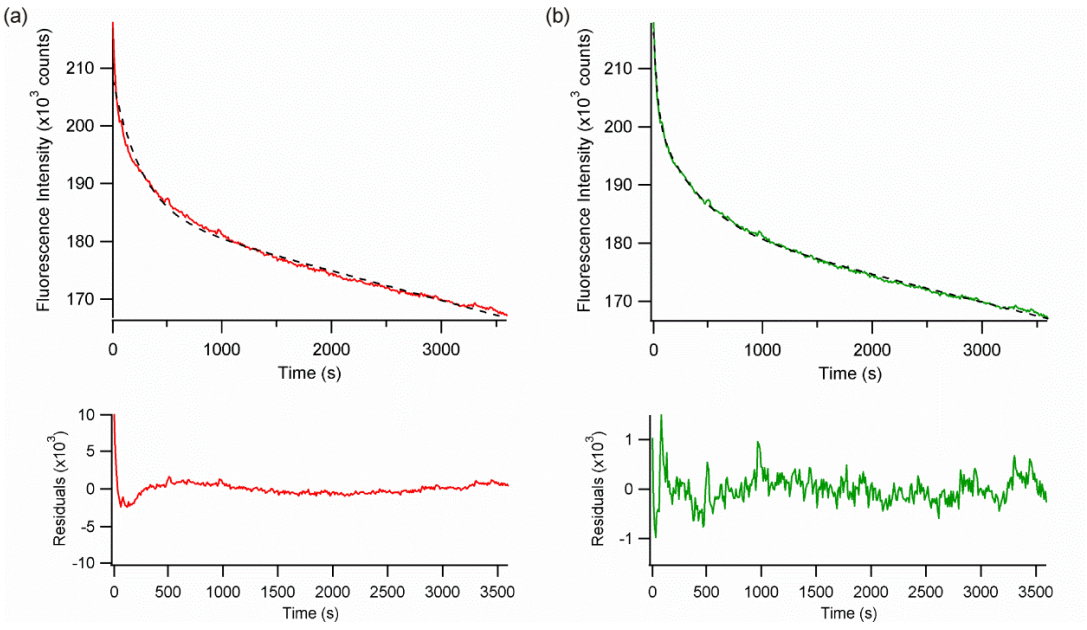

Figure 4.6 OmpT unfolding kinetics are best described by a double exponential function. (a) OmpT unfolding (red line) fitted to a single exponential function with a linear term (black, dashed line) and (b) OmpT unfolding (green line) fitted to a double exponential function with a linear term (black, dashed line). The fitting residuals are shown in the lower panels of the figure. Experiments contained 0.4 μM OmpT and were carried out in 100 nm *di*$C_{16:1}$PC LUVs at an LPR of 3200:1, 9 M urea, 50 mM sodium phosphate, pH 8.0, 25 °C.

The reproducibility of the unfolding rate constants was tested by measuring replicate unfolding samples using a four-cell changer and comparing the rate constants obtained from a minimum of three different liposome batches. The results demonstrate that OmpT consistently unfolds through two exponential phases at pH 8.0 (Figure 4.7) and reproducible rate constants are obtained (summarised in Table 4.4, Figure 4.16). In order to determine if OmpT unfolding under acidic conditions is also described by two unfolding phases, kinetic measurements of OmpT unfolding were next investigated at pH 3.8. OmpT was refolded in 1.5 M GuHCl, 50 mM sodium citrate, pH 3.8 in 100 nm *di*$C_{16:1}$PC LUVs for 16 h, before a ten-fold dilution into 6 M GuHCl was used to initiate unfolding and the tryptophan fluorescence emission at 335 nm was recorded. The resulting transients were fitted globally to both single (Figure 4.8a) and double (Figure 4.8b) exponential functions and the suitability of the fit determined using the fitting residuals (Figure 4.8, bottom). These experiments revealed that a double exponential function is required to satisfactorily describe the data (summarised in Table 4.4, Figure 4.16). This result is in contrast with the single exponential unfolding kinetics previously reported by Huysmans *et al* for HT PagP[167; 240] under conditions where folding is reversible without hysteresis (see Section 4.2.3) and led to the hypothesis that the second phase in the unfolding kinetics of OmpT may correspond to the OmpT detachment from the lipid surface.

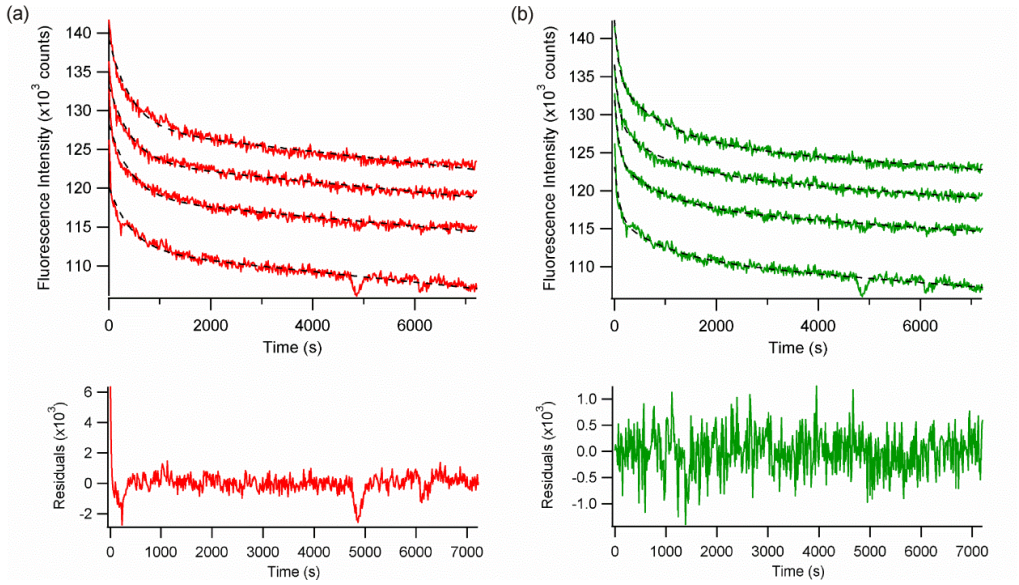

Figure 4.7 OmpT unfolding kinetics are reproducible. (a) OmpT unfolding (red lines) fitted to a single exponential function with a linear term (black, dashed lines) and (b) OmpT unfolding (green lines) fitted to a double exponential function with a linear term (black, dashed lines). Typical fitting residuals are shown in the lower panels of the figure. Experiments contained 0.4 μM OmpT and were carried out in 100 nm *di*$C_{16:1}$PC LUVs at an LPR of 3200:1, 9 M urea, 50 mM sodium phosphate, pH 8.0, 25 °C. Four replicate unfolding transients are shown in the top panels.

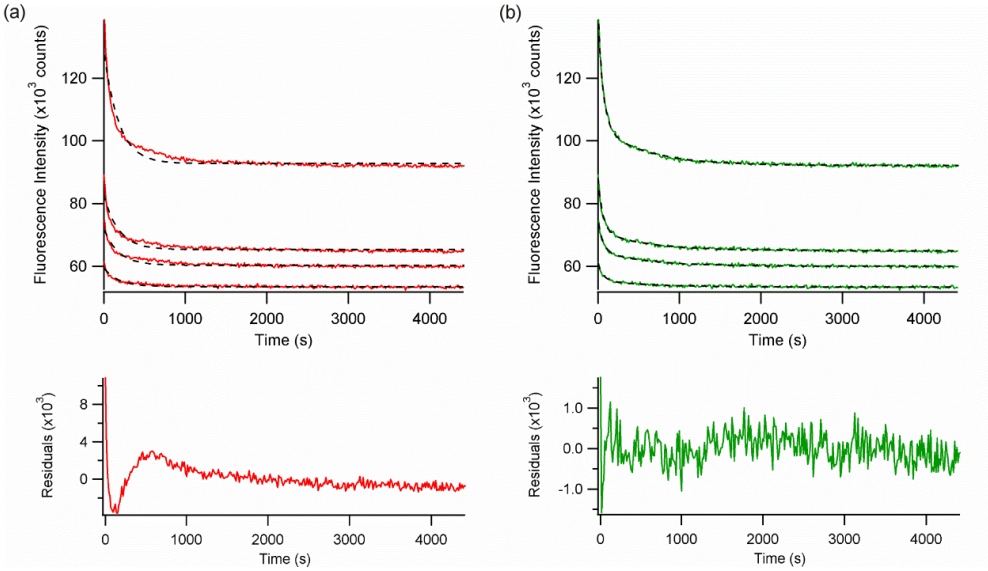

Figure 4.8 OmpT unfolding kinetics under acidic conditions are best described by a double exponential function. (a) OmpT unfolding (red lines) fitted to a single exponential function (black, dashed lines) and (b) OmpT unfolding (green lines) fitted to a double exponential function (black, dashed lines). Typical fitting residuals are shown in the lower panels of the figure. Experiments contained 0.4 μM OmpT and were carried out in 100 nm *di*$C_{16:1}$PC LUVs at an LPR of 3200:1, 6 M GuHCl, 50 mM sodium citrate, pH 3.8, 25 °C. Four replicate unfolding transients are shown in the top panels.

4.2.3 HIS-TAGGED PAGP FOLDS REVERSIBLY AND UNFOLDS BY SINGLE EXPONENTIAL KINETICS

The validity of the hypothesis that unfolding kinetics described by a double exponential could be indicative of lipid-detachment and hence hysteresis of folding and unfolding was tested further using HT PagP as a model system of reversible folding. The free energy of unfolding and the unfolding kinetics of HT PagP have been characterised previously[167; 240]. It was decided to test if these results could be replicated and ensure the resulting thermodynamic and kinetic parameters agreed with the published values. HT PagP was over-expressed in *E. coli* BL21 (DE3) cells as insoluble inclusion bodies and purified by nickel affinity chromatography according to the method of Huysmans *et al*[238], which is described in detail in Sections 2.4.1, 2.4.2 and 2.4.4). The resulting HT PagP was greater than 95 % in purity and, although a faint second band that may be attributed to the presence of a small amount of dimerised protein is observed by SDS-PAGE, was sufficiently pure for biophysical analysis. A typical yield of 50 mg pure protein per litre of culture was obtained.

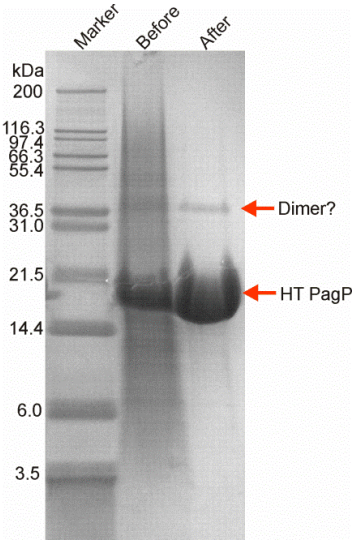

Figure 4.9 Purification of HT PagP using nickel affinity chromatography. Samples of the solubilised inclusion bodies before purification and the resulting protein after purification are shown. The second band in the pure sample is most likely due to the presence of a small amount of dimer. The size in kiloDaltons (kDa) of the protein markers is indicated.A detailed description of the purification strategy can be found in Sections 2.4.1, 2.4.2 and 2.4.4.

Folding reversibility was tested following the method of Huysmans *et al*[167], in which 1.2 μM HT PagP in 7 M urea was allowed to fold for 16 h in 50 mM sodium phosphate, pH 8.0, 25 °C in 100 nm *di*$C_{12:0}$PC LUVs at an LPR of 3200:1 before a three-fold dilution into aliquots of increasing urea concentration (7–10 M). These unfolding samples were allowed to equilibrate for a minimum of 8 h before the tryptophan emission spectrum of each sample was measured. Samples in the refolding direction were made by folding 3.6 μM HT PagP for 16 h under identical buffer conditions before a three-fold dilution into 10 M urea to unfold the protein. After 8 h equilibration, a three-fold dilution was then carried out to refold samples at urea concentrations of 7–10 M, which were allowed to equilibrate for a minimum of 8 h before the tryptophan emission spectrum of each sample was measured. The resulting denaturation curves could be overlaid (Figure 4.10) demonstrating that folding of HT PagP under these conditions is reversible without hysteresis. Fitting the resulting data to a two-state model as described in Section 2.5.6 yielded fit parameters ($\Delta G^{\circ H_2O}_{UN}$ = -60.3 ± 6.24 kJ mol^{-1} and M_{UN} = 6.92 ± 0.74 kJ mol^{-1} M^{-1}) that are in good agreement with those previously reported ($\Delta G^{\circ H_2O}_{UN}$ = -60.2 ± 0.3 kJ mol^{-1} and M_{UN} = 6.86 ± 0.20 kJ mol^{-1} M^{-1})[240].

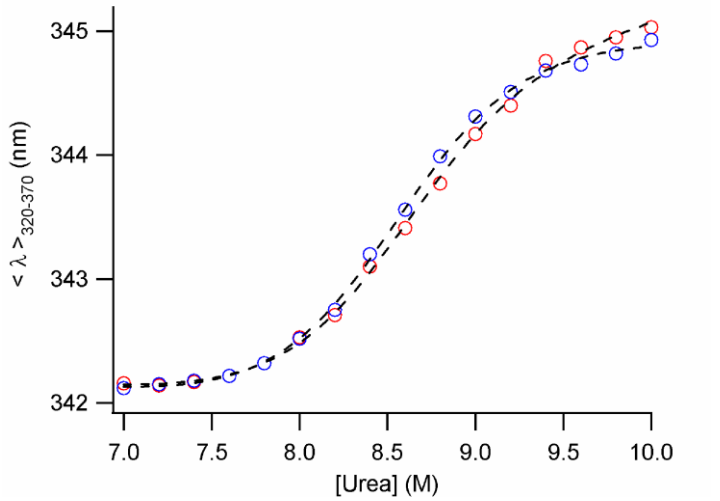

Figure 4.10 Folding of HT PagP is reversible without hysteresis. HT PagP urea denaturation in the unfolding (red) and refolding (blue) directions. All experiments contained 0.4 μM HT PagP and were carried out in 100 nm *di*$C_{12:0}$PC LUVs at an LPR of 3200:1, 50 mM sodium phosphate, pH 8.0, 25 °C. Black dashed lines indicate fits to a two-state folding model using a quantum yield correction factor of 1.65 (see Section 2.5.6).

The unfolding kinetics of HT PagP were next measured by allowing the protein to refold in 7 M urea, 50 mM sodium phosphate, pH 8.0 in 100 nm *di*$C_{12:0}$PC LUVs for 16 h, before a ten-fold dilution into 10 M urea was used to initiate unfolding. The change in tryptophan fluorescence emission at 335 nm was measured to follow the progress of the reaction and the resulting traces were globally fit to exponential functions as described in Section 2.5.7. The data were globally fit to both a single exponential function (Figure 4.11a, top) and a double exponential function (Figure 4.11b, top) revealing that the fitting residuals are not significantly improved in the double exponential fit (Figure 4.11, bottom panels). Additionally, although a relatively large difference in the two folding rate constants was observed ($k_1 = 8.59\times10^{-3} \pm 8.95\times10^{-4}\ s^{-1}$ and $k_2 = 9.84\times10^{-4} \pm 6.79\times10^{-6}\ s^{-1}$), the amplitude of k_1 is less than 10 % of that of k_2 (8253 counts and 90361 counts, respectively), suggesting that a single exponential fit is sufficient to describe the data.

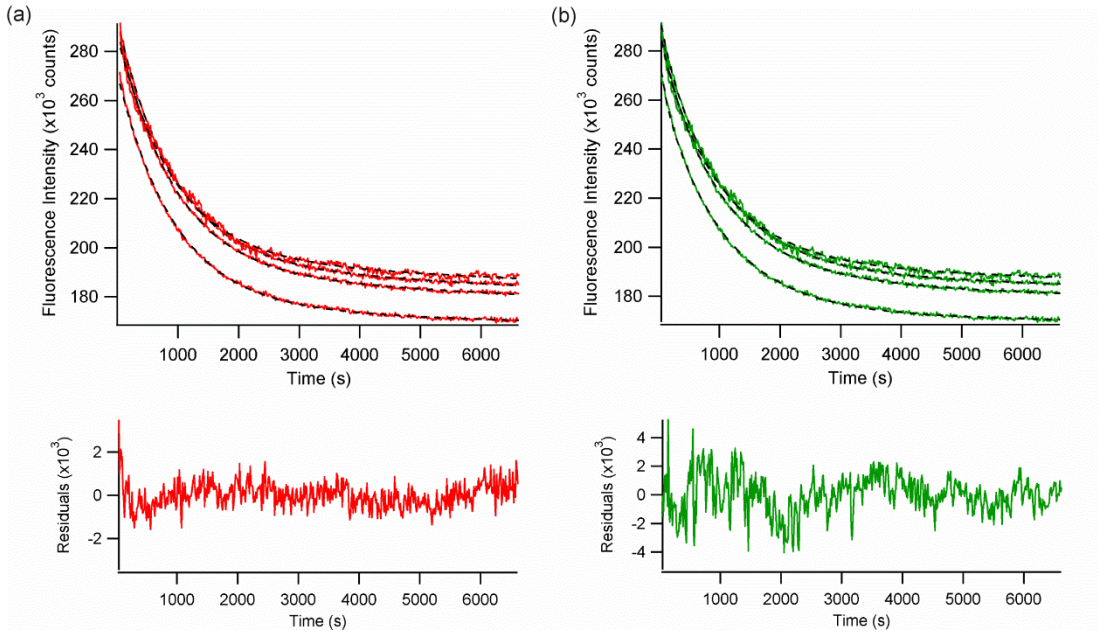

Figure 4.11 HT PagP unfolding kinetics are best described by a single exponential function. (a) HT PagP unfolding (red lines) fitted to a single exponential function (black, dashed lines) and (b) HT PagP unfolding (green lines) fitted to a double exponential function (black, dashed lines). Typical fitting residuals are shown in the lower panels of the figure. Experiments contained 0.4 μM HT PagP and were carried out in 100 nm *di*$C_{12:0}$PC LUVs at an LPR of 3200:1, 10 M urea, 50 mM sodium phosphate, pH 8.0, 25 °C. Four replicate unfolding transients are shown in the top panels.

Unfolding kinetics were measured in replicates of four in a total of four different batches of liposomes, and the obtained rate constant averaged. The measured rate constant ($k = 1.08\times10^{-3} \pm 8.25\times10^{-5}$ s^{-1}, summarised in Table 4.4, Figure 4.16) is comparable to the previously reported value[240] of 0.14 min^{-1}, which corresponds to approximately 2.2×10^{-3} s^{-1}. These results confirm that under reversible folding conditions, this protein unfolds *via* a single exponential process and is consistent with the hypothesis that lipid-adhesion in the unfolded state is linked to folding reversibility.

4.2.4 The Relationship Between Folding Reversibility and Unfolding Kinetics in Untagged PagP

Analysis of OmpT and HT PagP suggest a possible link between unfolding kinetics and folding reversibility. This may arise, however, due to the different proteins used as models for irreversible and reversible folding systems. The validity of the hypothesis that unfolding kinetics and folding reversibility are linked was further examined using the untagged construct of PagP first reported by Burgess *et al*[162]. PagP was over-expressed in *E. coli* BL21 (DE3) cells as insoluble inclusion bodies, which were isolated according to the method of Burgess *et al*[162] (described in Sections 2.4.1 and 2.4.2). Further purification of PagP was achieved by gel filtration using a Superdex 75 HiLoad 26/60 column (GE Healthcare) in 6 M GuHCl, 25 mM Tris-HCl, pH 8.0 as described in Section 2.4.5. The resulting protein was greater than 95 % pure (Figure 4.12) and was typically obtained in yields of 70 mg pure protein per litre of culture.

Previous work by Dr Alice I. Bartlett (University of Leeds) on the reversibility of PagP folding in 100 nm *di*$C_{12:0}$PC LUVs revealed that under conditions in which HT PagP folding is reversible without hysteresis (50 mM sodium phosphate, pH 8.0, 25 °C in 100 nm *di*$C_{12:0}$PC LUVs at an LPR of 3200:1), the folding of PagP shows hysteresis (Figure 4.13a). In contrast, under similar conditions to those reported by Moon *et al*[165] for the reversible folding of OmpLa (50 mM sodium citrate buffer,

pH 3.8, 37 °C in 100 nm *di*$C_{12:0}$PC LUVs at an LPR of 3200:1), PagP folding is reversible without hysteresis (Figure 4.13b). The PagP reversibility experiments were performed according to the protocols described by Huysmans *et al*[167] and Moon *et al*[165], however, the 3-(N,N-dimethylmyristyl-ammonio)propanesulfonate was excluded from the dilution steps in the latter protocol as this did not affect whether hysteresis was observed in PagP folding. In spite of this omission, the parameters obtained following fitting to a two-state folding model as described in Section 2.5.6 ($\Delta G^{\circ H_2O}_{UN}$ = -89.5 ± 5.55 kJ mol^{-1} and M_{UN} = 20.9 ± 1.3 kJ mol^{-1} M^{-1}) were comparable to those that have since been reported by Moon *et al*[174] (Table 4.1).

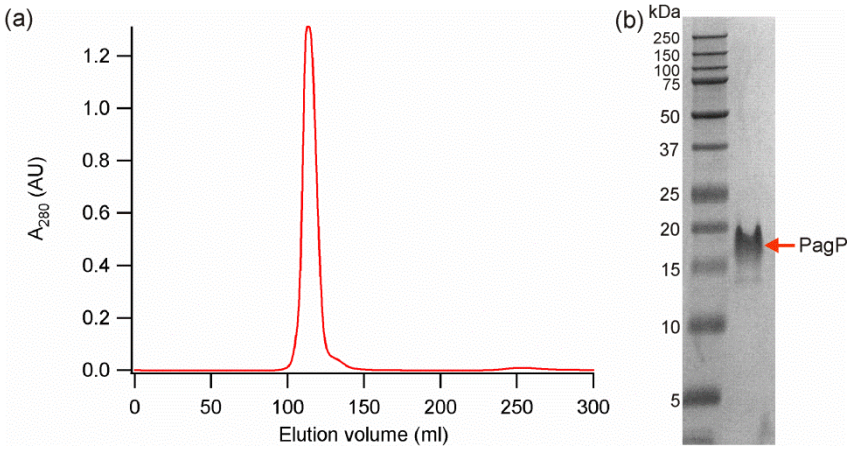

Figure 4.12 Purification of PagP. (a) Typical elution profile from the size exclusion column. (b) SDS-PAGE gel showing a PagP sample after purification. The size in kiloDaltons (kDa) of the protein markers is indicated. A detailed description of the purification strategy can be found in Sections 2.4.1, 2.4.2 and 2.4.5.

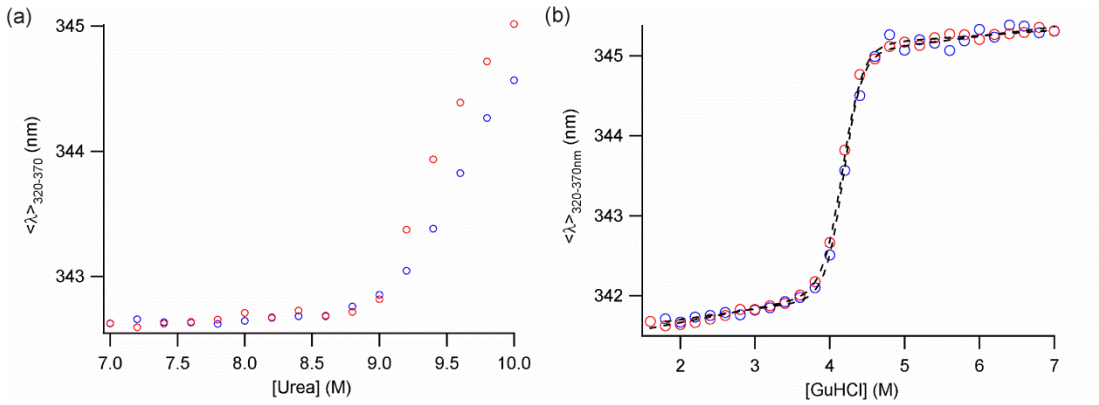

Figure 4.13 Folding reversibility of PagP is dependent on buffer conditions. PagP urea denaturation in the unfolding (red) and refolding (blue) directions in (a) 50 mM sodium phosphate, pH 8.0, 25 °C and (b) 50 mM sodium citrate, pH 3.8, 37 °C. All experiments contained 0.4 μM PagP and were carried out in 100 nm *di*$C_{12:0}$PC LUVs at an LPR of 3200:1. Black dashed lines indicate fits to a two-state folding model using a quantum yield correction factor of 2.39 (see Section 2.5.6). The fit parameters obtained were $\Delta G^{\circ H_2O}_{UN}$ = -89.5 ± 5.55 kJ mol^{-1} and M_{UN} = 20.9 ± 1.3 kJ mol^{-1} M^{-1}. These data were acquired by Dr Alice I. Bartlett (University of Leeds).

The ability to modulate hysteresis of PagP folding by changing buffer conditions allows the unfolding kinetics of PagP to be compared under conditions which favour either folding with or without hysteresis. The unfolding kinetics of PagP were measured by allowing the protein to refold in 100 nm *di*$C_{12:0}$PC LUVs for 16 h in either 7 M urea, 50 mM sodium phosphate, pH 8.0 or 3 M GuHCl, 50 mM sodium citrate, pH 3.8. Ten-fold dilutions into either 10 M urea or 6 M GuHCl were then used to initiate unfolding. Unfolding kinetic data were measured and fitted to exponential functions as before (Section 4.2.3). At pH 8.0, where PagP folding shows hysteresis, the unfolding kinetics are found to be poorly described by a single exponential function (Figure 4.14a) and a double exponential function (Figure 4.14b) is required to give a satisfactory fit, as judged by the fitting residuals (Figure 4.14, bottom). Conversely, at pH 3.8 a single exponential function provides a good fit to the experimental data (Figure 4.15a) and fitting the same data to a double exponential function (Figure 4.15b) yields no improvement in the fitting residuals (Figure 4.15, bottom). Additionally, the two rate constants obtained from the double exponential fit are almost identical (k_1 = 4.4×10^{-3} ± 1.4×10^{-4} s^{-1} and k_2 = 8.1×10^{-3} ± 5.4×10^{-4} s^{-1}), indicating that this is truly a single exponential process.

Replicate unfolding kinetics measured in different liposome batches showed good reproducibility under both conditions (summarised in Figure 4.16).

Overall, these data suggest that the previously identified relationship between hysteresis in urea denaturation experiments and unfolding kinetics is not a protein-dependent effect but is observed for the three different OMPs tested thus far. The second phase observed in the unfolding kinetics of the OMPs under conditions which favour hysteresis in urea denaturation experiments is hypothesised to correspond to the unfolded protein becoming detached from the liposome. This second unfolding phase is not observed for folding systems which fold reversibly without hysteresis as these exist in a two-state equilibrium between the folded, membrane-inserted state and the unfolded, lipid-adsorbed state. These data suggest the ability to predict the presence of hysteresis in folding and unfolding experiments based on the unfolding kinetics of an OMP, however, they do not provide evidence that the second phase is indeed due to lipid detachment.

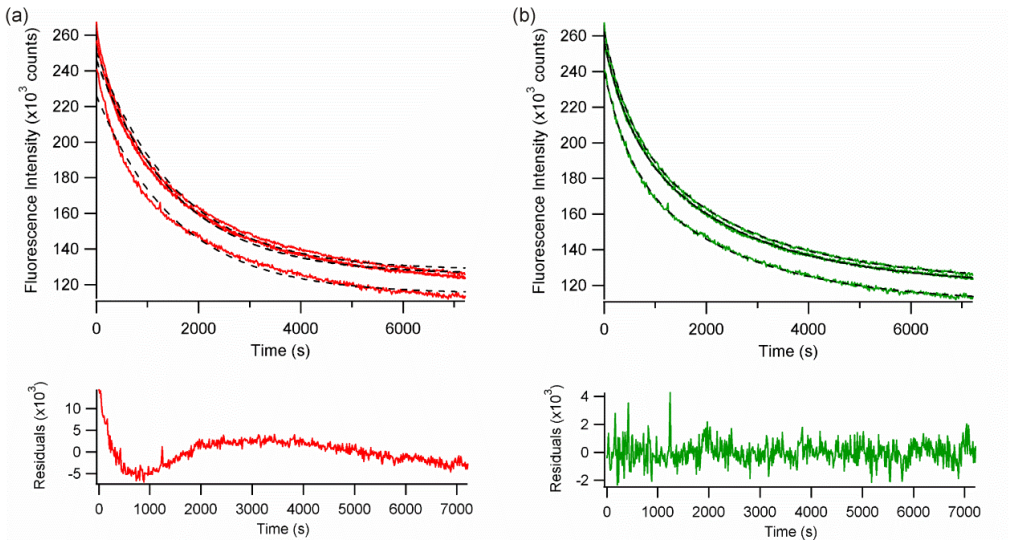

Figure 4.14 PagP unfolding kinetics are best described by a double exponential function at pH 8.0. (a) PagP unfolding (red lines) fitted to a single exponential function (black, dashed lines) and (b) PagP unfolding (green lines) fitted to a double exponential function (black, dashed lines). Typical fitting residuals are shown in the lower panels of the figure. Experiments contained 0.4 μM PagP and were carried out in 100 nm *di*$C_{12:0}$PC LUVs at an LPR of 3200:1, 10 M urea, 50 mM sodium phosphate, pH 8.0, 25 °C. Four replicate unfolding transients are shown in the top panels.

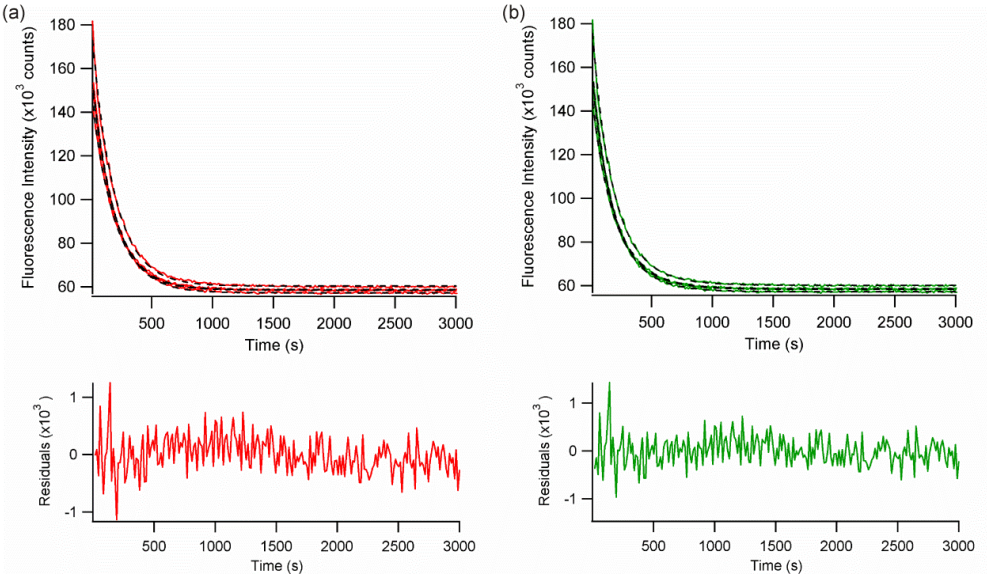

Figure 4.15 PagP unfolding kinetics are best described by a single exponential function at pH 3.8. (a) PagP unfolding (red lines) fitted to a single exponential function (black, dashed lines) and (b) PagP unfolding (green lines) fitted to a double exponential function (black, dashed lines). Typical fitting residuals are shown in the lower panels of the figure. Experiments contained 0.4 μM PagP and were carried out in 100 nm *di*$C_{12:0}$PC LUVs at an LPR of 3200:1, 6 M GuHCl, 50 mM sodium citrate, pH 3.8, 37 °C. Four replicate unfolding transients are shown in the top panels.

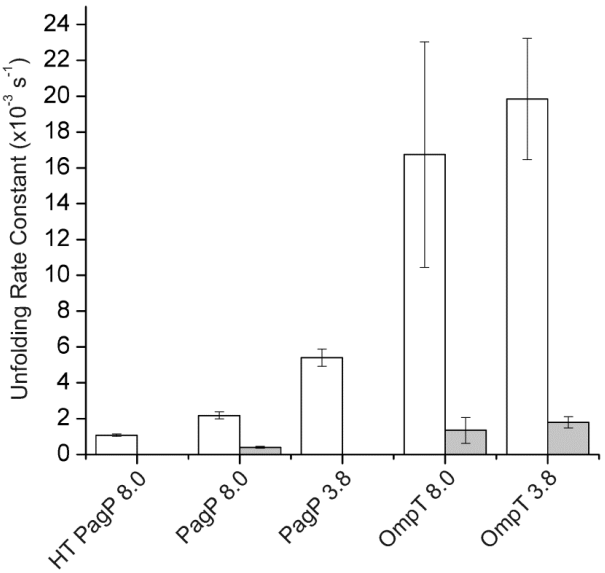

Figure 4.16 Summary of unfolding rate constants of HT PagP, PagP and OmpT. The rate constants obtained from exponential fits of the unfolding transients, k_1 and k_2, are shown by white and grey bars, respectively. Where no value of k_2 is shown, data were best described by a single exponential process. Error bars depict the standard deviations of the rate constants. All samples contained 0.4 μM OMP with 100 nm LUVs (*di*$C_{12:0}$PC for HT PagP and PagP, *di*$C_{16:1}$PC for OmpT) at an LPR of 3200:1 and 50 mM buffer (sodium phosphate, pH 8.0 or sodium citrate, pH 3.8). Rate constants and their associated errors are also provided in Table 4.4.

Protein	Lipid	pH	Hysteresis	Exponential Fit	k_1 ($\times 10^{-3}$ s^{-1})	Mean k_1 ($\times 10^{-3}$ s^{-1})	Std Dev k_1 ($\times 10^{-3}$ s^{-1})	k_2 ($\times 10^{-4}$ s^{-1})	Mean k_2 ($\times 10^{-4}$ s^{-1})	Std Dev k_2 ($\times 10^{-4}$ s^{-1})
HT PagP	*di*$C_{12:0}$PC	8.0	No	Single	1.05	1.08	0.07	–	–	–
					1.14					
					1.15					
					0.98					
PagP	*di*$C_{12:0}$PC	8.0	Yes	Double	2.17	2.18	0.20	4.29	3.95	0.48
					2.07			4.36		
					1.97			3.16		
					2.50			3.99		
		3.8	No	Single	5.41	5.40	0.48	–	–	–
					4.65					
					5.98					
					5.56					
OmpT	*di*$C_{16:1}$PC	8.0	Yes	Double	27.1	16.7	6.3	25.5	13.5	7.2
					10.2			6.40		
					14.2			10.0		
					15.5			12.0		
		3.8	Yes	Double	17.6	19.8	3.4	22.2	17.9	3.2
					24.6			14.4		
					17.3			17.2		

Table 4.4 Summary of unfolding kinetic data. Std Dev is an abbreviation of standard deviation. Each rate constant was obtained by fitting four replicate unfolding transients to exponential functions as described in Section 2.5.7.

4.2.5 Use of Dansyl-Labelled Lipid to Detect Differences in the Unfolding Kinetics of PagP Constructs

Under conditions where hysteresis is observed in urea denaturation experiments, the second phase observed in the unfolding kinetics of OMPs is hypothesised to arise from the unfolded protein becoming detached from the lipid. This hypothesis was investigated using FRET studies. For this, it was decided to compare HT PagP and PagP unfolding so that models both with and without observed hysteresis could be examined without changing buffer conditions between experiments. Both constructs of PagP have twelve tryptophan residues throughout the protein structure (Figure 1.24a), whose fluorescence emission is used to report on folding and unfolding of the proteins throughout this thesis. These residues, however, can also be utilised as FRET donors. 5-(Dimethylamino)naphthalene-1-sulfonyl (dansyl) labelled *di*$C_{18:1}$PS was chosen as a suitable FRET acceptor as the excitation maximum of dansyl occurs at approximately 335 nm[283], the emission wavelength used in the unfolding kinetic studies of PagP due to the large fluorescence emission of tryptophan at this wavelength. Since FRET is strongly distance-dependent[284], if the OMP remains lipid bound in the unfolded state, it would be expected to display higher FRET efficiency between tryptophan residues and dansyl than if the OMP is not lipid-adhered. Higher FRET efficiency between the donor and acceptor groups would cause a higher fluorescence emission from the dansyl group at the end of the reaction, and so the unfolded state of the two proteins can be compared.

Before FRET measurements were made, the unfolding kinetics of both HT PagP and PagP were measured to ensure the presence of the dansyl labelled *di*$C_{18:1}$PS does not affect the unfolding kinetics. The unfolding kinetics of HT PagP and PagP were measured by allowing the protein to refold in 7 M urea, 50 mM sodium phosphate, pH 8.0 in 100 nm *di*$C_{12:0}$PC LUVs doped with 1 % *di*$C_{18:1}$PS-dansyl for 16 h, before a ten-fold dilution into 10 M urea was used to initiate unfolding. The change in tryptophan fluorescence emission at 335 nm was measured to follow the progress of the reaction as before and the resulting traces were globally fit to both single and double exponential functions as described in Section 2.5.7. The low fluorescence amplitude of the resulting traces is a result of the FRET between tryptophan and dansyl, causing a reduction in the tryptophan fluorescence emission intensity, and makes the data more

difficult to fit with confidence. In spite of this, the resulting rate constants were found to be in reasonable agreement with those measured previously in 100 nm *di*$C_{12:0}$PC LUVs (Table 4.5), with HT PagP unfolding *via* single exponential kinetics (Figure 4.17) and PagP unfolding *via* double exponential kinetics (Figure 4.18). All unfolding rate constants obtained in 100 nm *di*$C_{12:0}$PC LUVs doped with 1 % *di*$C_{18:1}$PS-dansyl are slower than the corresponding rate constants measured in 100 nm *di*$C_{12:0}$PC LUVs, and this may perhaps be attributed to FRET effects altering the rate of change of fluorescence intensity at 335 nm.

Construct	Lipid	k_1 ($\times 10^{-3}$ s^{-1})	Std Dev k_1 ($\times 10^{-3}$ s^{-1})	k_2 ($\times 10^{-4}$ s^{-1})	Std Dev k_2 ($\times 10^{-4}$ s^{-1})
HT PagP	*di*$C_{12:0}$PC	1.08	0.07	–	–
HT PagP	*di*$C_{12:0}$PC + dansyl	0.72	0.08	–	–
PagP	*di*$C_{12:0}$PC	2.18	0.20	3.95	0.48
PagP	*di*$C_{12:0}$PC + dansyl	1.67	0.11	2.30	0.20

Table 4.5 Comparison of unfolding rate constants of PagP and HT PagP in 100 nm LUVs of *di*$C_{12:0}$PC in the presence or absence of 1 % *di*$C_{18:1}$PS-dansyl. Std Dev is an abbreviation of standard deviation. Each rate constant is obtained by averaging the calculated rate constants from a minimum of three different liposomes batches.

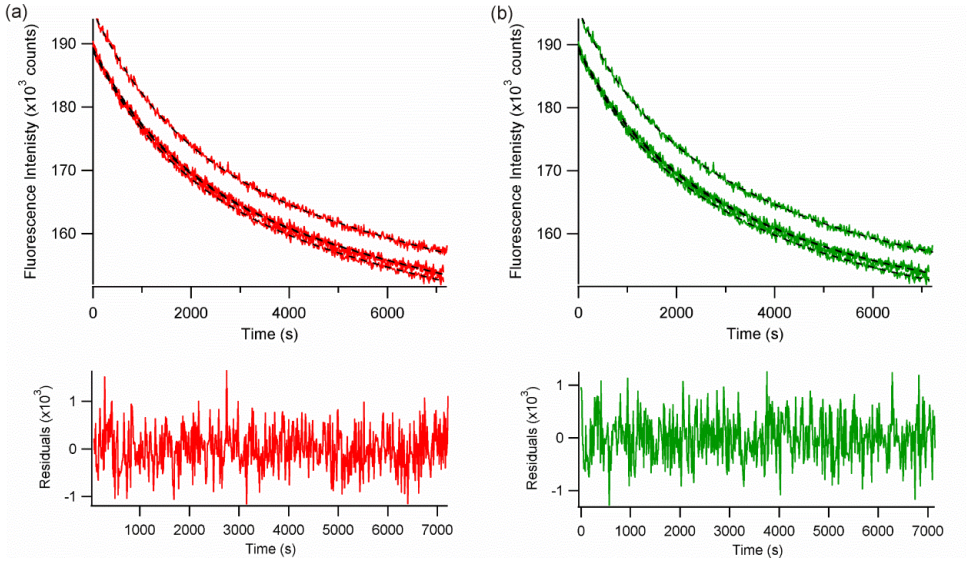

Figure 4.17 Unfolding kinetics of HT PagP in 100 nm *di*$C_{12:0}$PC LUVs are not affected by the presence of 1 % *di*$C_{18:1}$PS-dansyl. (a) HT PagP unfolding (red lines) fitted to a single exponential function (black, dashed lines) and (b) HT PagP unfolding (green lines) fitted to a double exponential function (black, dashed lines). Typical fitting residuals are shown in the lower panels of the figure. Experiments contained 0.4 μM HT PagP and were carried out in 100 nm *di*$C_{12:0}$PC LUVs doped with 1 % *di*$C_{18:1}$PS-dansyl at an LPR of 3200:1, 10 M urea, 50 mM sodium phosphate, pH 8.0, 25 °C. Three replicate unfolding transients are shown in the top panels.

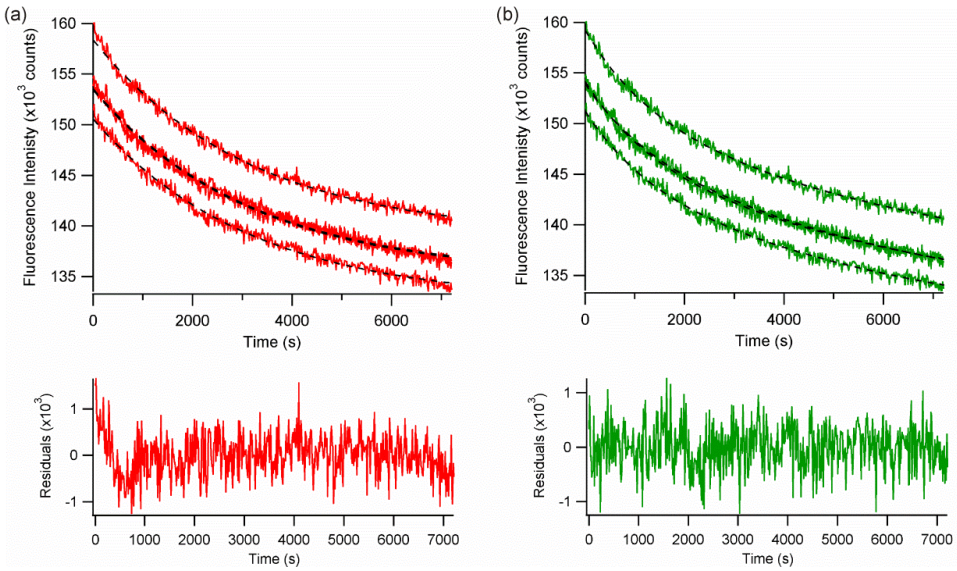

Figure 4.18 Unfolding kinetics of PagP in 100 nm *di*$C_{12:0}$PC LUVs are not affected by the presence of 1 % *di*$C_{18:1}$PS-dansyl. (a) PagP unfolding (red lines) fitted to a single exponential function (black, dashed lines) and (b) PagP unfolding (green lines) fitted to a double exponential function (black, dashed lines). Typical fitting residuals are shown in the lower panels of the figure. All experiments contained 0.4 μM PagP and were carried out in 100 nm *di*$C_{12:0}$PC LUVs doped with 1 % *di*$C_{18:1}$PS-dansyl at an LPR of 3200:1, 10 M urea, 50 mM sodium phosphate, pH 8.0, 25 °C. Three replicate unfolding transients are shown in the top panels.

To monitor changes in FRET between tryptophan residues in the PagP constructs and lipid-attached dansyl, unfolding reactions were initiated as described previously; however, following excitation of tryptophan residues at 280 nm the fluorescence emission spectrum from 300–500 nm was recorded at regular intervals. By following the entire spectrum over these wavelengths and time, changes in both tryptophan emission and dansyl emission can be measured. Protein unfolding was monitored by following the change in $\langle\lambda\rangle_{320\text{-}370}$ over time as this follows changes in spectral shape and should not be affected by low intensity changes caused by tryptophan quenching. Dansyl emission was monitored by following the change in the area under the intensity versus wavelength curve ($A_{460\text{-}540}$) between 460 and 540 nm. The relative emission at any time-point was calculated by normalising the $A_{460\text{-}540}$ at the relevant time-point between the maximum $A_{460\text{-}540}$ (that when folded protein is present) and the minimum $A_{460\text{-}540}$ (that of the dansyl-doped liposomes alone).

For both PagP and HT PagP, the unfolding rate constants measured by the change in $\langle\lambda\rangle_{320\text{-}370}$ (Table 4.6) were similar to those measured by following changes in the intensity of tryptophan emission at 335 nm (Table 4.5). It should be noted that the former method contains fewer data points and hence yields rate constants with lower accuracy than the latter method, particularly for the fast rate (k_1) of PagP unfolding. Additionally, unfolding rate constants obtained by monitoring changes in relative dansyl emission are also similar to those measured by the change in $\langle\lambda\rangle_{320\text{-}370}$ (Table 4.6). This indicates that changes in dansyl emission arise due to changes in the FRET between tryptophan residues and dansyl as the PagP constructs unfold. Unfolding of HT PagP was highly reproducible when monitored using either $\langle\lambda\rangle_{320\text{-}370}$ or $A_{460\text{-}540}$, and FRET at the end of the unfolding reaction was evident through the elevated dansyl emission intensity compared with dansyl-doped liposomes in the absence of protein (Figure 4.19). Unfolding of PagP by both methods showed more variability between samples, but was overall reproducible. At the end of the unfolding reaction, the dansyl emission was equal to that of dansyl-doped liposomes in the absence of protein, suggesting that no FRET is occurring (Figure 4.20).

Construct	Method	k_1 ($\times 10^{-3}$ s^{-1})	Fitting Error k_1 ($\times 10^{-3}$ s^{-1})	k_2 ($\times 10^{-4}$ s^{-1})	Fitting Error k_2 ($\times 10^{-4}$ s^{-1})
HT PagP	$\langle\lambda\rangle_{320\text{-}370}$	0.90	0.02	–	–
HT PagP	$A_{460\text{-}540}$	1.90	0.12	–	–
PagP	$\langle\lambda\rangle_{320\text{-}370}$	7.80	0.69	3.05	0.24
PagP	$A_{460\text{-}540}$	8.33	0.26	4.33	0.30

Table 4.6 Comparison of unfolding rate constants of PagP and HT PagP in 100 nm LUVs of *di*$C_{12:0}$PC in the 1 % *di*$C_{18:1}$PS-Dansyl by two spectral methods. $\langle\lambda\rangle_{320\text{-}370}$ refers to the intensity averaged wavelength of tryptophan emission, the calculation of which is described in Section 2.5.6. $A_{460\text{-}540}$ refers to the area under the dansyl emission curve, the calculation of which is described in Section 2.5.8.

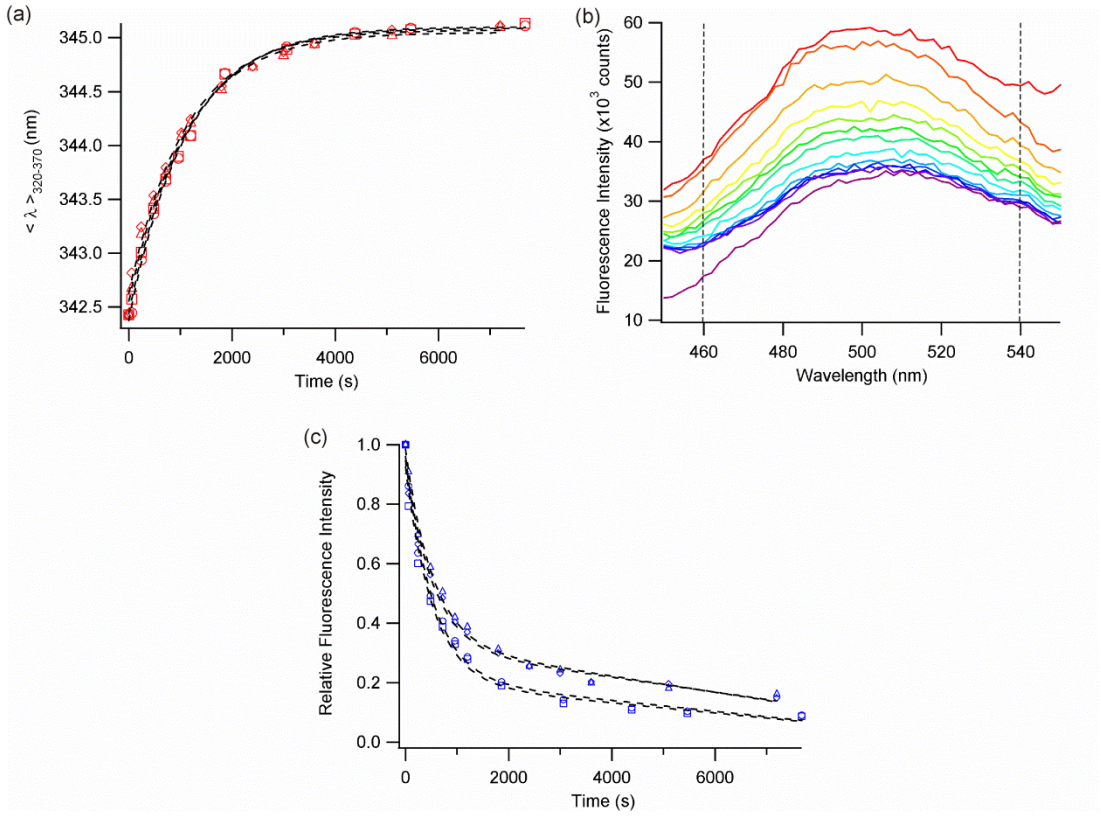

Figure 4.19 Unfolding kinetics of HT PagP in 100 nm *di*$C_{12:0}$PC LUVs doped with 1 % *di*$C_{18:1}$PS-dansyl. (a) Unfolding of HT PagP monitored using $\langle\lambda\rangle_{320\text{-}370}$, (b) emission of dansyl following excitation tryptophan residues in HT PagP at 280 nm (red line indicates folded HT PagP present, violet line indicates *di*$C_{12:0}$PC LUVs doped with 1 % *di*$C_{18:1}$PS-dansyl in the absence of protein, other solid lines show changes in dansyl emission over time with orange indicating the start of the unfolding reaction and blue indicating the end of the unfolding reaction) and (c) unfolding of HT PagP monitored using $A_{460\text{-}540}$. In (a) and (c), four replicate experiments are depicted (circle, square, triangle and diamond markers), while fits to single exponential functions are shown as black dashed lines. In (b) black dashed lines indicate the range of wavelengths used to calculate efficiency of energy transfer. All experiments contained 0.4 μM HT PagP and were carried out in 100 nm *di*$C_{12:0}$PC LUVs doped with 1 % *di*$C_{18:1}$PS-dansyl at an LPR of 3200:1, 10 M urea, 50 mM sodium phosphate, pH 8.0, 25 °C.

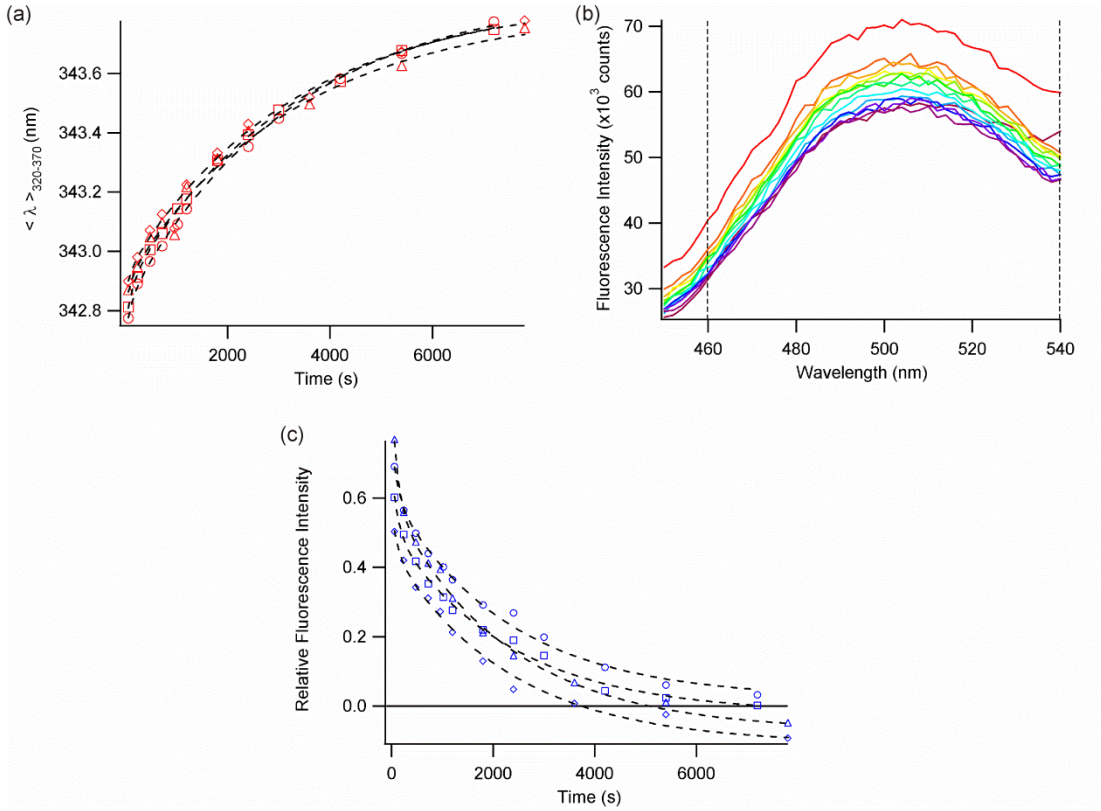

Figure 4.20 Unfolding kinetics of PagP in 100 nm *di*$C_{12:0}$PC LUVs doped with 1 % *di*$C_{18:1}$PS-dansyl. (a) Unfolding of PagP monitored using $\langle\lambda\rangle_{320\text{-}370}$, (b) emission of dansyl following excitation tryptophan residues in HT PagP at 280 nm (red line indicates folded PagP present, violet line indicates *di*$C_{12:0}$PC LUVs doped with 1 % *di*$C_{18:1}$PS-dansyl in the absence of protein, other solid lines show changes in dansyl emission over time with orange indicating the start of the unfolding reaction and blue indicating the end of the unfolding reaction) and (c) unfolding of PagP monitored using $A_{460\text{-}540}$. In (a) and (c), four replicate experiments are depicted using differently shaped markers (circle, square, triangle and diamond), while fits to double exponential functions are shown as black dashed lines. In (b) black dashed lines indicate the range of wavelengths used to calculate efficiency of energy transfer. All experiments contained 0.4 μM PagP and were carried out in 100 nm *di*$C_{12:0}$PC LUVs doped with 1 % *di*$C_{18:1}$PS-dansyl at an LPR of 3200:1, 10 M urea, 50 mM sodium phosphate, pH 8.0, 25 °C.

A total of eight replicate unfolding reactions were carried out for both PagP and HT PagP and the relative dansyl fluorescence at the end of the reaction was compared to give a measure of the residual FRET between tryptophan residues and the dansyl groups (Figure 4.21). The amount of residual FRET at the end of each reaction was highly reproducible over all eight replicates for both HT PagP and PagP. Although the absolute difference between the blue and violet spectra in Figure 4.19 is small, it was consistent over all replicates. The average residual FRET is shown to be approximately

20 % for HT PagP and 0 % for PagP compared with dansyl-doped liposomes in the absence of protein. This difference was found to be highly significant using a t-test ($p < 0.01$). These data suggest that the unfolded state of HT PagP is sufficiently close to the liposome surface for measurable FRET to occur between tryptophan residues and dansyl, but this is not the case for PagP.

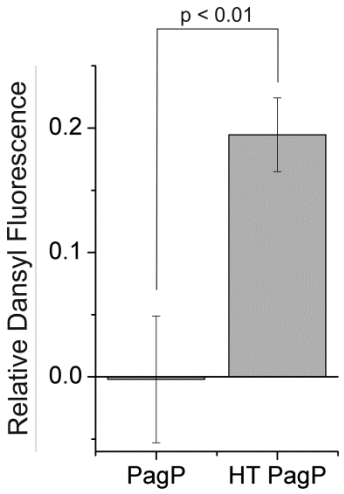

Figure 4.21 Relative dansyl fluorescence remaining at the end point of unfolding reactions of PagP and HT PagP in 100 nm *di*$C_{12:0}$PC LUVs doped with 1 % *di*$C_{18:1}$PS-dansyl. Data are the average of eight replicate unfolding samples, with error bars depicting the standard error. All experiments contained 0.4 μM PagP or HT PagP and were carried out in 100 nm *di*$C_{12:0}$PC LUVs doped with 1 % *di*$C_{18:1}$PS-dansyl at an LPR of 3200:1, 10 M urea, 50 mM sodium phosphate, pH 8.0, 25 °C.

4.3 DISCUSSION

Measurement of the free energy of unfolding of OMPs *in vitro* has proved to be difficult due to the necessity of identifying experimental conditions under which folding is completely reversible without hysteresis. Hysteresis between the folding and unfolding denaturation curves is commonly reported[150; 151; 165; 174] and, as a result, there are few studies which have been successful in determining free energies of unfolding of OMPs (Table 4.1). Two main approaches for overcoming hysteresis have been employed in these studies. The first is to allow the samples to equilibrate for very long times prior to measurements, which has been successful in achieving OmpA reversibility in both detergent micelles and amphipols[150; 151]. Measurement of samples repeatedly over up to

52 days results in the shift of the refolding denaturation curve until it overlays with the unfolding denaturation curve[151], and this effect is typical in systems with extremely high kinetic barriers to unfolding. Very high kinetic stability has been observed in dissociation of subunits of the *E. coli* Type 1 pilus[281]. The extrapolated half-life of this process is 3×10^9 years and this results in stable pili with subunits that never dissociate under physiological conditions[281]. Additionally, unfolding of the prokaryotic α-lytic protease has been shown to be inhibited by high kinetic barriers as the native state of this protein is actually thermodynamically unstable compared with the unfolded state[202]. Upon denaturation, α-lytic protease would require millennia to refold unaided, however, folding is catalysed by the "pro-region", an N-terminal domain which is degraded after folding occurs *in vivo*[203]. The high kinetic stability of α-lytic protease has been attributed to a conserved β-hairpin in the C-terminal domain[204]. Protein engineering experiments to alter the conformation of this β-hairpin to a more favourable one were successful in increasing the kinetic stability of α-lytic protease but gave rise to impaired proteolytic function[204]. The experiments on both pilis subunits and α-lytic protease demonstrate the advantages of large kinetic barriers to unfolding *in vivo*, but also highlight the limitations on protein stability imposed by functional constraints[202; 203; 204; 281]. Such extremely high kinetic barriers may act to inhibit OMP unfolding and ensure no disruption to the essential functions carried out by OMPs *in vivo*[46; 64; 282]. For studies of OMP folding in liposomes, it is not possible to incubate samples for more than a few days[158] to allow full equilibration of kinetically stable samples and so conditions must be found under which folding reversibility without hysteresis is achievable on shorter timescales. While this has been achieved for two different PagP constructs, OmpW, OmpLa and OmpA[149; 165; 167; 174], problems with aggregation led to hysteresis under many conditions and continues to hamper work in this area.

Following the work described in Chapter 3, folding of OmpT in *di*$C_{16:1}$PC liposomes was carried out in different conditions with the aim of achieving reversible folding without hysteresis. Changing buffer conditions, protein concentration, denaturant and equilibration time with the aim of preventing aggregation all failed to improve the large degree of observed hysteresis. With so many examples of this phenomenon apparent in the literature, it is perhaps surprising that no previous studies have attempted to ascertain which factors control OMP folding reversibility *in vitro*. Following the problems experienced in achieving folding reversibility without hysteresis, the focus of

this study shifted to examine possible causes of hysteresis in unfolding and folding equilibration.

During the course of the reversibility studies of OmpT, a short unfolding time was employed and this revealed that OmpT unfolds *via* double exponential kinetics. This is notable as studies on the folding kinetics of HT PagP reported single exponential unfolding kinetics of this protein under conditions where folding is reversible without hysteresis[167; 240]. Additionally, these studies suggested that the measured reversible equilibrium was between the folded, membrane-inserted state of HT PagP and the unfolded, lipid-adsorbed state[167]. Together these observations led to the hypothesis that if OmpT becomes detached from the liposome in the unfolded state, it could then refold *via* parallel pathways with different kinetics. Folding of OmpT from a lipid-detached state could give rise to aggregation of unfolded OmpT, or a loss of folding competency, leading to the observed hysteresis in the folding and unfolding equilibrium.

This hypothesis was first tested by examining the folding equilibria for the presence of hysteresis and the unfolding kinetics of HT PagP, PagP and OmpT under different conditions to establish whether single exponential unfolding kinetics are indeed linked to folding reversibility. The results show that for the OMP constructs tested, double exponential unfolding kinetics are observed when the folding equilibrium exhibits hysteresis. To examine whether the second exponential phase observed when the folding equilibrium shows hysteresis arises due to the unfolded protein becoming detached from the membrane surface, liposomes were doped with dansyl-labelled lipid as this moiety acts as a FRET acceptor to tryptophan[283]. Unfolding kinetics of PagP and HT PagP were then measured in these dansyl-doped liposomes under identical conditions where HT PagP folds reversibly without hysteresis and PagP folds and unfolds with hysteresis. The results obtained suggest that FRET between dansyl and tryptophan still occurs in the unfolded state of HT PagP but not in the unfolded state of PagP. The Förster distance (R_0) of the dansyl-tryptophan FRET pair has been estimated to be 21 Å[283]. Assuming that there should be observable FRET between 0.5–2 R_0[284], it can be concluded that for no FRET to be detectable in the unfolded state of PagP, all tryptophan residues in the protein must be further than 42 Å from the liposome in the unfolded state. This leads to the conclusion that it is unlikely that PagP remains adhered to the liposome in the unfolded state. Conversely, residual FRET in the unfolded state of HT PagP suggests that the unfolded protein is further from the dansyl

on the liposome surfaces than the folded protein, but close enough for FRET to occur. This could arise from the HT PagP remaining lipid-bound in the unfolded state.

The large decrease in FRET as HT PagP unfolds could be explained by the fact that dansyl groups will be present on both the inner and outer surfaces of the liposome. In the presence of folded HT PagP, tryptophan residues in both aromatic girdles of this OMP (Figure 1.24a) will allow FRET to occur with dansyl groups on both surfaces of the bilayers. In contrast, upon unfolding HT PagP will cease to be membrane inserted and the lipid-adhered unfolded state will only facilitate FRET with the dansyl groups on the surface of the liposome nearest the unfolded protein. Additionally, the FRET efficiency with the dansyl groups on the outer surface of the bilayer could be expected to be lower in the unfolded than the folded state due to the effects of different orientations of the donor and acceptor with respect to each other[284]. Together these data provide evidence that hysteresis in the folding equilibria of PagP and HT PagP is dependent on lipid adsorption of the unfolded state. These experiments, however, provide an indirect measurement of the adsorption state of the unfolded protein. One possible experimental method to directly measure if the unfolded states of HT PagP and PagP are lipid-attached or free in solution would be to dope the liposomes with biotinylated lipid and use streptavidin beads to separate the lipid-adhered and lipid-inserted protein from protein free in solution. Alternatively, sucrose density gradient centrifugation could be used to separate lipid-bound OMP species from those free in solution. In both of the suggested experiments, fractions could then be analysed using cold SDS-PAGE and thus the folded state of the protein could also be assessed. Additionally, the subset of OMPs examined could be expanded to include all of the proteins whose folding equilibria have been previously examined[174].

While the work presented in this chapter suggests that lipid adhesion in the unfolded state could be important for folding reversibility without hysteresis of different PagP constructs, this cannot be stated conclusively. Additionally, it is not yet known whether this effect is generic for all OMPs under all conditions, or if other factors could be at play. The recent study by Moon *et al* compared fluorescence emission spectra of unfolded PagP and OmpW, under conditions where folding of both proteins was reversible without hysteresis, to that of *N*-acetyl-L-tryptophanamide under identical buffer conditions[174]. This led to the conclusion that both proteins were likely to be completely solvated in aqueous solution and not bound to the membrane surface[174].

These data suggest that it is also possible that folding reversibility without hysteresis is not determined by lipid adhesion in the unfolded state, but by the ability of the unfolded protein to re-adsorb to the liposome upon dilution of denaturant. This would imply that folding reversibility without hysteresis depends on the delicate balance of lipid adsorption and aggregation of the denatured state in low concentrations of urea. If this is true in the case of all OMPs, it is perhaps surprising that HT PagP folding is reversible under conditions where very high urea concentrations are required to ensure solubility[238], while PagP folding is irreversible even though this construct remains soluble at much lower concentrations of urea[162]. It is apparent that much more work needs to be done on the unfolded states of OMPs under conditions where the folding equilibrium has been characterised in order to fully understand the factors which govern reversible folding.

In vivo, large numbers of chaperones have been identified in the *E. coli* periplasm which are thought to be involved in OMP folding and membrane insertion[76; 83; 107]. The presence of these proteins in the OMP folding pathway is essential in preventing the misfolding and aggregation of these insoluble proteins *in vivo*, and their absence in studies *in vitro* must be compensated for by use of carefully designed experimental conditions. The understanding of exactly which requirements must be fulfilled to achieve reversible folding of an OMP *in vitro* will be invaluable in aiding future studies of OMP folding, stability and function.

5 DISSECTING THE EFFECTS OF PERIPLASMIC CHAPERONES ON THE FOLDING OF OUTER MEMBRANE PROTEINS

5.1 INTRODUCTION

The folding pathway of a C-terminally His-tagged construct of PagP (the mature PagP sequence followed by leucine, glutamic acid and $(histidine)_6$, referred to as HT PagP herein) in the absence of chaperones has been described in detail in Section 1.8.2. The work in this chapter aims to establish conditions under which the folding of an OMP can be studied in the presence of periplasmic chaperones and shed light on the mechanism of action of these chaperones.

Previous analysis on HT PagP revealed that high concentrations of urea are required in order to maintain this construct in a soluble, folding competent state[167; 238; 240] (see Section 1.8.2). This is in stark contrast with the OMP folding screen reported by Burgess *et al*[162], which showed the folding of nine different OMPs, including an untagged construct of PagP, into a range of synthetic lipid vesicles at urea concentrations as low as 1 M using the cold SDS-PAGE method. Although the different OMPs studied folded with different efficiencies under each condition, in general, folding at high pH into vesicles comprised of lipids with shorter acyl chains favoured the highest yields (Figure 5.1)[162]. These results suggest that the development of an OMP folding assay in the presence of soluble periplasmic chaperones enabling the examination of the effects of these proteins on folding rate and yield is possible.

Although many proteins thought to act as periplasmic chaperones have been identified (see Section 1.5.2), the mechanisms by which they interact with unfolded OMPs to promote correct folding and membrane insertion remain poorly understood[76; 83]. In this chapter, the effect of two chaperones, SurA and Skp (discussed in detail in Sections 1.5.2.1 and 1.5.2.2, respectively), on the folding kinetics of OMPs are investigated. Following analysis of the folding kinetics of PagP, OmpT and OmpP in low concentrations of urea, only PagP was taken forward for study in the presence of chaperones based on highly reproducible folding kinetics that are well described by an exponential function.

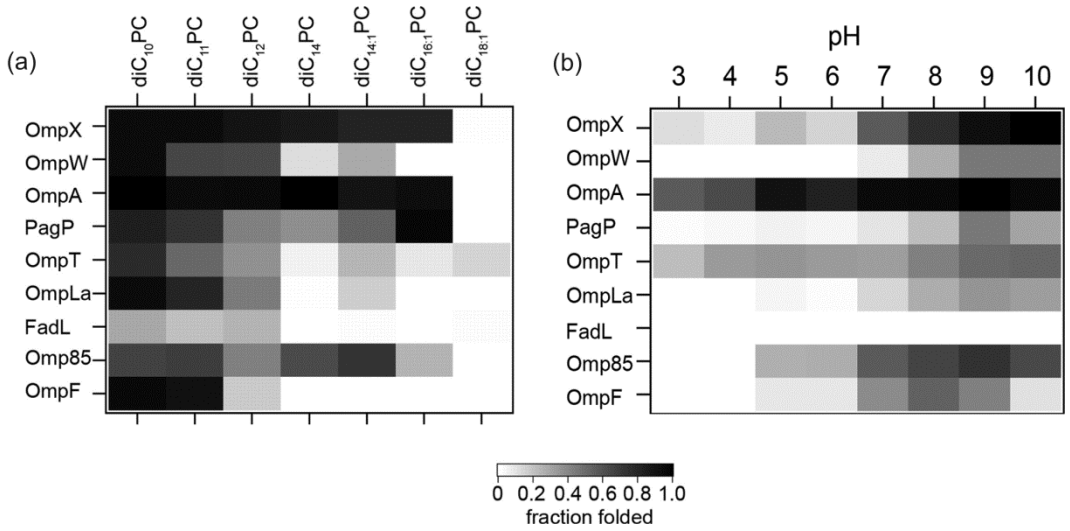

Figure 5.1 Effect of lipid acyl chain length and pH on the efficiency of OMP folding. (a) Fraction of folded OMP observed in the presence of phosphocholine (PC) lipids of different chain lengths. OMPs were incubated with 100 nm large unilamellar vesicles (LUVs) at a lipid to protein ratio (LPR) of 800:1 for 15 h at 38.3 °C in 10 mM sodium borate, pH 10.0 containing 1 M urea and 2 mM EDTA. (b) Effect of pH on the observed fraction folded of each OMP. OMPs were incubated with 100 nm *di*$C_{12:0}$PC LUVs at an LPR of 800:1 for 15 h in 10 mM buffer at the appropriate pH containing 1 M urea and 2 mM EDTA. SDS-PAGE loading buffer was added to quench the folding reactions and the samples were analysed by cold SDS-PAGE (*i.e.* not boiled prior to loading onto the gel). The fraction folded for each reaction was determined by quantifying the relative intensities of the folded and unfolded bands by densitometry. This figure has been adapted from Burgess *et al*, 2008[162].

Folding kinetics of PagP into both zwitterionic *di*$C_{12:0}$PC liposomes and negatively charged 80:20 *di*$C_{12:0}$PC:*di*$C_{12:0}$PG liposomes (for lipid structures see Figure 5.2) were investigated using a combination of spectroscopic and SDS-PAGE assays. The results presented show that the membrane composition and ionic strength of the buffer strongly influence the effect of Skp on PagP folding, suggesting that electrostatic interactions play an important role in the mechanism of action of Skp. In contrast, SurA did not affect the observed folding rates of PagP, consistent with the view that these chaperones act by distinct mechanisms in partially redundant chaperone pathways. The ability of Skp to prevent OMP aggregation was also investigated, revealing that even in conditions where aggregation is strongly favoured, Skp can rescue the folding and membrane insertion of HT PagP. These results demonstrate the use of minimalist *in vitro* assays for investigating the complex folding environment of the periplasm and

demonstrate the key role of Skp in holding unfolded and aggregation-prone OMPs prior to their delivery, directly or indirectly, to the membrane where folding can be initiated.

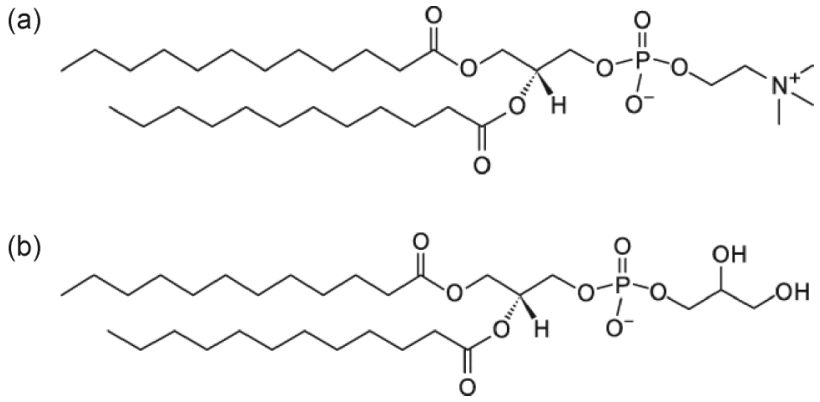

Figure 5.2 Structures of C12 lipids used in this chapter. (a) 1,2-dilauroyl-*sn*-glycero-3-phosphocholine (*di*$C_{12:0}$PC) and (b) 1,2-dilauroyl-*sn*-glycero-3-phospho-(1'-*rac*-glycerol) (*di*$C_{12:0}$PG).

5.2 RESULTS

5.2.1 DEVELOPMENT OF KINETIC FOLDING ASSAYS IN LOW CONCENTRATIONS OF UREA

To facilitate biophysical studies of folding in the presence of periplasmic chaperones, it is necessary to establish conditions under which the OMP is in a soluble, folding-competent state while simultaneously maintaining the chaperone in a folded and functional state. Satisfying this criterion allows direct comparison between the folding reactions in the presence and absence of chaperones, giving insights into the mechanisms of chaperone action. To characterise the folding of OmpT, OmpP and PagP (expressed and purified as shown in Sections 3.2.2.2, 3.2.2.3 and 4.2.4), an unfolded stock of OMP in 10 M urea was diluted directly into 100 nm LUVs at an LPR of 3200:1 and folding was monitored using tryptophan fluorescence emission. The enrichment of tryptophan residues in the membrane intrafacial region of membrane proteins[271; 272] makes tryptophan fluorescence a sensitive probe of folding and membrane insertion.

The folding parameters tested were chosen based on the results of the screens described in Section 3.2.3 and also the work described by Huysmans *et al*[167; 238; 240] and Burgess *et al*[162]. Two lipid compositions were tested: zwitterionic *di*$C_{12:0}$PC and negatively charged 80:20 *di*$C_{12:0}$PC:*di*$C_{12:0}$PG. The former was chosen based on the observation that lipids with shorter acyl chains favour higher folding yields and the previous success of folding studies of HT PagP in this lipid[162; 167; 238; 240], while the latter was chosen to give a negative surface charge similar to that observed in the OM of Gram-negative bacteria *in vivo*[66]. Data were fitted to exponential functions as described in Section 2.6.1.

5.2.1.1 OmpT

OmpT was over-expressed and purified as described in Sections 3.2.2.2 and 3.2.2.3. OmpT folding was initiated by dilution from a concentrated stock of protein (100 μM) in 10 M urea to 0.4 μM OmpT into 100 nm LUVs and the fluorescence emission at 335 nm (following excitation at 280 nm) monitored for up to 2 h. The lipid composition, LPR, temperature and pH were varied. The final concentration of urea in the folding samples ranged from 1–2 M.

Folding of OmpT into 100 nm *di*$C_{12:0}$PC LUVs at an LPR of 3200:1 was carried out in a final concentration of 2 M urea, 50 mM sodium phosphate, pH 8.0 at 25 °C (Figure 5.3). Kinetic folding traces were not well described by either a single exponential function (Figure 5.3a) or a double exponential function (Figure 5.3b). Increasing the temperature to 37 °C resulted in a faster folding reaction, as judged by the time taken to reach plateau, however, the kinetic folding traces were once again not well described by either a single (Figure 5.3c) or double (Figure 5.3d) exponential function. The origin of the complex folding kinetics appears to be from a short "lag" phase at the beginning of the transient, which was hypothesised to arise from the initial association of the unfolded protein with the liposome surface. This association has been shown for HT PagP to occur in the dead-time of manual-mixing experiments and manifests as a burst-phase in fluorescence folding transients[167; 240]. If this occurs on a slower time-scale for OmpT folding into *di*$C_{12:0}$PC liposomes, it may give rise to the observed lag-phase.

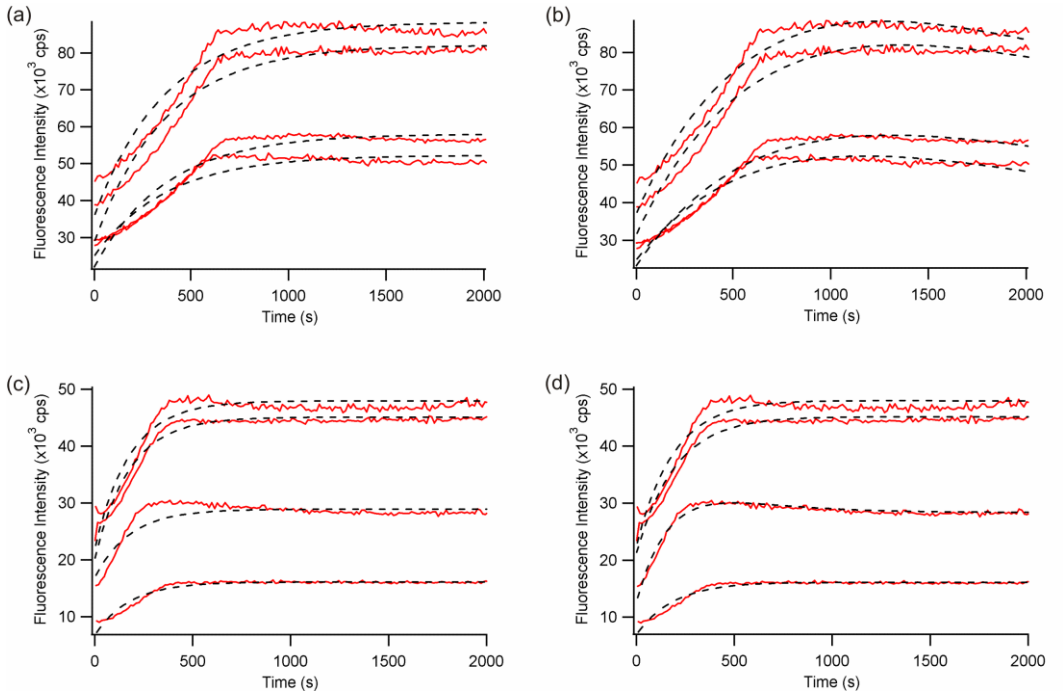

Figure 5.3 Kinetics of OmpT folding into 100 nm *di*$C_{12:0}$PC LUVs at pH 8.0. Fluorescence kinetic traces are shown as red lines, with the exponential fit as black dashed lines. Folding was measured at 25 °C and fitted to (a) a single exponential function and (b) a double exponential function. Folding was also measured at 37 °C and fitted to (c) a single exponential function and (d) a double exponential function. All experiments contained 0.4 μM OmpT at an LPR of 3200:1 in 2 M urea, 50 mM sodium phosphate, pH 8.0. Four replicate transients are shown in each panel.

To investigate if the lag-phase does in fact originate from association of the unfolded OmpT to the lipid, the folding reaction was repeated in 100 nm *di*$C_{12:0}$PC LUVs in 50 mM sodium phosphate, pH 8.0 containing 2 M urea and monitored for up to 2 h at 25 °C. The LPR was varied from 400:1 to 6400:1 under the assumption that lipid adsorption is a bimolecular reaction dependent on both the protein and lipid concentrations, as suggested by the results reported for HT PagP[240]. The data show that decreasing LPR overall lowers the folding rate, as judged by the time taken to reach plateau, and increases the length of the lag-phase, but in all cases the lag-phase is still present (Figure 5.4).

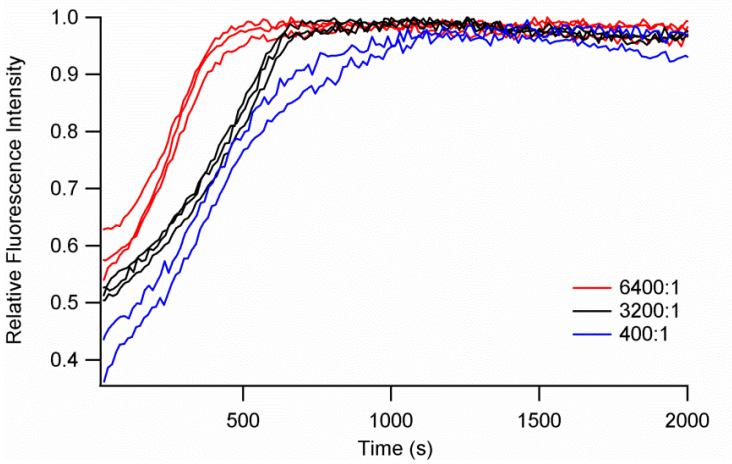

Figure 5.4 Effect of changing the LPR on kinetics of OmpT folding into 100 nm *di*$C_{12:0}$PC LUVs. The LPR was varied between 6400:1 (red lines), 3200:1 (black lines) and 400:1 (blue lines). Kinetic traces have been normalised to the final fluorescence signal to highlight differences. All experiments were carried out at 25 °C and contained 0.4 μM OmpT in 2 M urea, 50 mM sodium phosphate, pH 8.0. 2–3 replicate transients are shown for each LPR.

Changing the buffer conditions from 50 mM sodium phosphate, pH 8.0 to 50 mM glycine, pH 9.5, but keeping all other conditions identical, also failed to eliminate the lag-phase in 100 nm *di*$C_{12:0}$PC LUVs (Figure 5.5a), suggesting that this phenomenon may arise due to the properties of the lipid surface itself rather than the buffer conditions or lipid concentration. When the folding experiment in 50 mM glycine, pH 9.5 was repeated with 100 nm 80:20 *di*$C_{12:0}$PC:*di*$C_{12:0}$PG LUVs, which changes the charge at the surface of the lipid from zwitterionic to negatively charged, the lag-phase was no longer observed in the time-course of the experiment and the refolding transients were well described by a single exponential function with a rate constant of $6.61\times10^{-3} \pm 8.61\times10^{-4}$ s^{-1} (Figure 5.5b, Table 5.1). Together these data suggest that the lag-phase does indeed correspond to adsorption of the unfolded OmpT to the liposome surface, and the presence of this effect make the data difficult to analyse quantitatively.

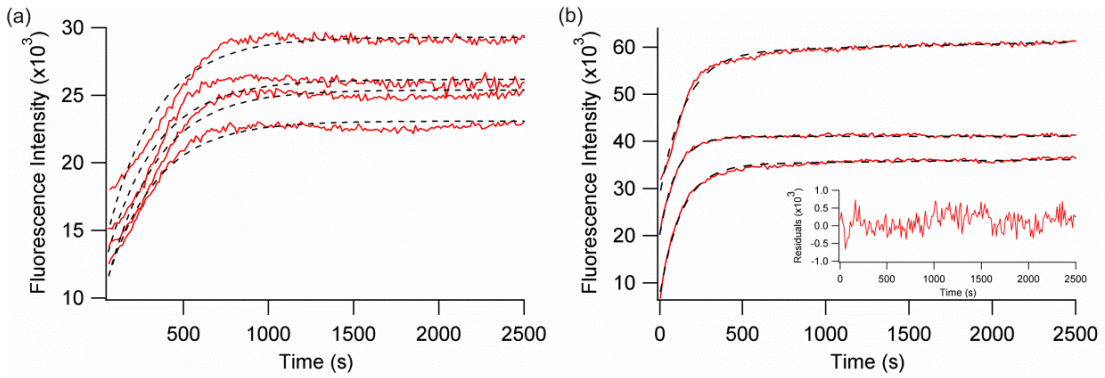

Figure 5.5 Effect of changing lipid surface charge on kinetics of OmpT folding into 100 nm LUVs. Fluorescence folding transients of OmpT into (a) *di*$C_{12:0}$PC and (b) 80:20 *di*$C_{12:0}$PC:*di*$C_{12:0}$PG. Fits to single exponential functions are shown as black dashed lines, with typical fitting residuals for the 80:20 *di*$C_{12:0}$PC:*di*$C_{12:0}$PG data shown in the inset of panel (b). All experiments were carried out at 37 °C and contained 0.4 μM OmpT in 2 M urea, 50 mM glycine, pH 9.5 with an LPR of 3200:1. 3–4 replicate transients are shown in each panel.

The final strategy employed to find simple, reproducible folding kinetics in 100 nm *di*$C_{12:0}$PC LUVs was to reduce the urea concentration. Folding reactions were carried out in 50 mM sodium phosphate, pH 8.0 at 25 °C and an LPR of 3200:1 with urea concentrations of 1.5 M and 1 M (Figure 5.6). At a urea concentration of 1.5 M, the lag phase is no longer visible in the folding transient (Figure 5.6a), however, the data are not satisfactorily described by a single exponential function as determined by the fitting residuals (Figure 5.6b). Conversely, in 1 M urea the folding reaction is well described by a single exponential function (Figure 5.6c), as determined by the fit residuals (Figure 5.6d), yielding a rate constant of $1.51\times10^{-2} \pm 9.34\times10^{-5}$ s^{-1} (Table 5.1).

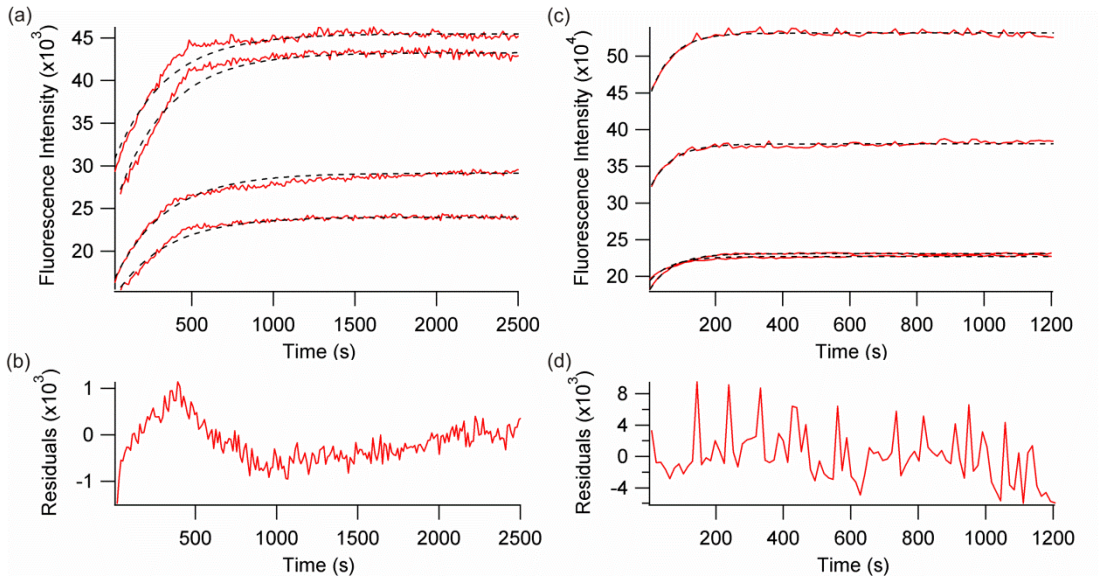

Figure 5.6 Effect of changing final urea concentration on kinetics of OmpT folding into 100 nm *di*$C_{12:0}$PC LUVs. Fluorescence folding transients in (a) 1.5 M urea and (c) 1 M urea, with fits to a single exponential function shown as black dashed lines. Typical fitting residuals (b) for the 1.5 M urea data and (d) the 1 M urea data are shown. All experiments contained 0.4 μM OmpT at an LPR of 3200:1 in 50 mM sodium phosphate, pH 8.0 and were measured at 25 °C. Four replicate transients are shown in the top panels.

Conditions	Ave. Rate Constant ($\times 10^{-3}$) s^{-1}	Standard Error ($\times 10^{-3}$) s^{-1}
100 nm *di*$C_{12:0}$PC LUVs, LPR 3200:1, 50 mM sodium phosphate, pH 8.0, 1 M urea, 25 °C	15.1	0.1
100 nm 80:20 *di*$C_{12:0}$PC:*di*$C_{12:0}$PG LUVs, LPR 3200:1, 50 mM glycine, pH 9.5, 2 M urea, 37 °C	6.61	0.86

Table 5.1 Summary of OmpT folding rate constants. Rate constants were obtained by measuring tryptophan fluorescence emission at 335 nm over time, and fitting the observed transients to a single exponential function. Global fits were obtained over four replicates from a single batch of liposomes, and the average of the global fits from folding reactions into three batches of liposomes calculated. The standard error of the mean was calculated by taking the number of liposome replicates to be 3.

Taken together, these results suggest OmpT folding kinetics are not well suited to study with chaperones as the folding kinetics obtained are mostly complex. While reproducible and exponential kinetics are observed under two conditions (100 nm *di*$C_{12:0}$PC LUVs at an LPR of 3200:1 in 1 M urea, 50 mM sodium phosphate, pH 8.0,

25 °C; and 100 nm 80:20 *di*$C_{12:0}$PC:*di*$C_{12:0}$PG LUVs at an LPR of 3200:1 in 2 M urea, 50 mM glycine, pH 9.5, 37 °C), these two sets of conditions are not directly comparable excluding the possibility of investigating the effect of lipid surface charge on chaperone activity. Additionally, the rate constant obtained in 100 nm *di*$C_{12:0}$PC LUVs at an LPR of 3200:1 in 1 M urea, 50 mM sodium phosphate, pH 8.0, 25 °C shows OmpT folding occurs very rapidly under these conditions and if chaperone activity were to further increase the rate of folding it would no longer be possible to measure the folding transient using this manual mixing technique. For these reasons, it was decided not to use OmpT as a model protein to investigate the mechanism of action of periplasmic chaperones.

5.2.1.2 OmpP

OmpP was over-expressed and purified as described in Sections 3.2.2.2 and 3.2.2.3. OmpP folding kinetics were first measured by dilution from a 100 μM stock unfolded in 10 M urea into 100 nm *di*$C_{12:0}$PC LUVs at an LPR of 3200:1 in 2 M urea, 50 mM sodium phosphate, pH 8.0, 37 °C. Tryptophan fluorescence emission was measured at 335 nm (following excitation at 280 nm) and monitored for up to 2 h. The resulting kinetic transients were well described by a single exponential function yielding an average rate constant of $1.82\times10^{-2} \pm 3.92\times10^{-3}$ s^{-1} (Figure 5.7a, Table 5.2). Folding into 100 nm 80:20 *di*$C_{12:0}$PC:*di*$C_{12:0}$PG LUVs under the same conditions occurred too quickly to be measured by this technique (data not shown), however, increasing the final urea concentration to 3 M was sufficient to slow the folding rate and gave an average rate constant of $2.93\times10^{-3} \pm 8.91\times10^{-4}$ s^{-1} (Figure 5.7b, Table 5.2).

OmpP folding kinetics into both 100 nm *di*$C_{12:0}$PC LUVs (Figure 5.8a) and 100 nm 80:20 *di*$C_{12:0}$PC:*di*$C_{12:0}$PG LUVs (Figure 5.8b) were also measured in 2 M urea, 50 mM glycine buffer, pH 9.5, which revealed single exponential transients in both lipids under these buffer conditions. Average rate constants obtained under these conditions were $8.52\times10^{-3} \pm 7.73\times10^{-4}$ s^{-1} and $1.02\times10^{-2} \pm 2.70\times10^{-3}$ s^{-1} for *di*$C_{12:0}$PC and 80:20 *di*$C_{12:0}$PC:*di*$C_{12:0}$PG LUVs, respectively (Table 5.2).

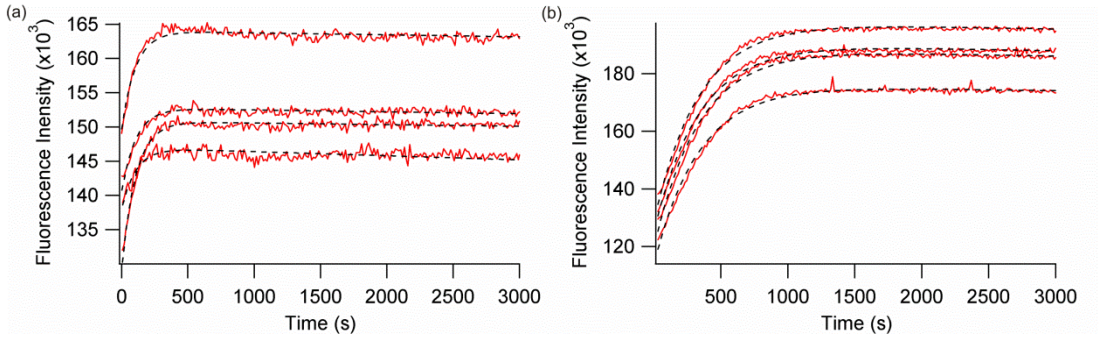

Figure 5.7 OmpP folding kinetics into 100 nm LUVs at pH 8.0. (a) OmpP folding into *di*$C_{12:0}$PC in 2 M urea and (b) OmpP folding into 80:20 *di*$C_{12:0}$PC:*di*$C_{12:0}$PG in 3 M urea measured using tryptophan fluorescence emission. Fits to single exponential functions are shown as black dashed lines. All experiments contained 0.4 μM OmpP at an LPR of 3200:1 in 50 mM sodium phosphate, pH 8.0 and were measured at 37 °C. Four replicate transients are shown in each panel.

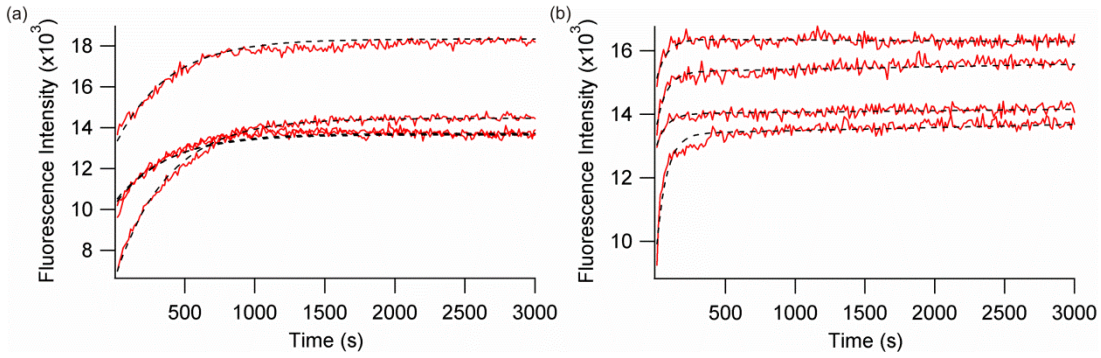

Figure 5.8 OmpP folding kinetics into 100 nm LUVs at pH 9.5. OmpP folding into (a) *di*$C_{12:0}$PC and (b) 80:20 *di*$C_{12:0}$PC:*di*$C_{12:0}$PG liposomes measured using tryptophan fluorescence emission. Fits to single exponential functions are shown as black dashed lines. All experiments contained 0.4 μM OmpP at an LPR of 3200:1 in 2 M urea, 50 mM sodium phosphate, pH 8.0 and were measured at 37 °C. Four replicate transients are shown in each panel.

Using OmpP as a model protein, conditions have been established under which reproducible folding kinetics can be measured in low concentrations of urea in two different lipid environments. These conditions allow the possibility of studying folding in the presence of periplasmic chaperones. The folding rate constant in 80:20 *di*$C_{12:0}$PC:*di*$C_{12:0}$PG LUVs, however, may be too fast to assess any increase in folding rate conferred unless a stopped-flow mixing assay could be developed successfully and the folding rate accelerated sufficiently for measurement by this technique.

Conditions	Average Rate Constant ($\times 10^{-2}$) s^{-1}	Standard Error ($\times 10^{-2}$) s^{-1}
100 nm *di*$C_{12:0}$PC LUVs, LPR 3200:1, 50 mM sodium phosphate, pH 8.0, 2 M urea, 37 °C	1.82	0.39
100 nm 80:20 *di*$C_{12:0}$PC:*di*$C_{12:0}$PG LUVs, LPR 3200:1, 50 mM sodium phosphate, pH 8.0, 3 M urea, 37 °C	0.29	0.09
100 nm *di*$C_{12:0}$PC LUVs, LPR 3200:1, 50 mM glycine, pH 9.5, 2 M urea, 37 °C	0.85	0.07
100 nm 80:20 *di*$C_{12:0}$PC:*di*$C_{12:0}$PG LUVs, LPR 3200:1, 50 mM glycine, pH 9.5, 2 M urea, 37 °C	1.02	0.27

Table 5.2 Summary of OmpP folding rate constants. Rate constants were obtained by measuring tryptophan fluorescence emission at 335 nm over time and fitting the observed transients to a single exponential function. Global fits were obtained over four replicates from a single batch of liposomes, and the average of the global fits from folding reactions into three batches of liposomes calculated. The standard error of the mean was calculated by taking the number of liposome replicates to be 3.

For the introduction of soluble chaperones into the assay, it was decided to add an addition dilution step such that the unfolded OmpP stock in 10 M urea was first diluted to 0.24 M urea for 5 min in the absence of lipid, either with or without chaperone. This low urea incubation was used to promote binding to the chaperone as the presence of 2–3 M urea required for efficient liposome insertion may disrupt interaction with the unfolded OMP. Following incubation, the mixture was then added to liposomes at the appropriate final concentration of urea. This protocol was first carried out in the absence of chaperones to determine whether low urea incubation would change the observed folding kinetics. While direct dilution into *di*$C_{12:0}$PC LUVs from an unfolded stock in 10 M urea yielded exponential folding kinetics, pre-incubation at 0.24 M urea before addition to liposomes did not result in a folding transient (Figure 5.9). These data suggest that OmpP does not remain folding-competent at 0.24 M urea and this additional step to promote binding would have to be omitted from the assay, which may disrupt chaperone binding.

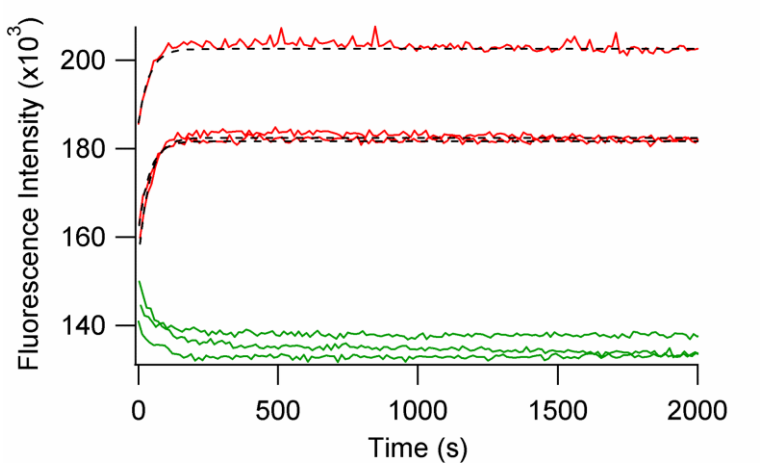

Figure 5.9 Effect of incubation at 0.24 M urea on OmpP folding into 100 nm *di*$C_{12:0}$PC LUVs. Unfolded OmpP solubilised in 10 M urea was added to 100 nm *di*$C_{12:0}$PC LUVs at a final concentration of 0.4 μM OmpP, 2 M urea either by direct dilution (red lines) or via a pre-incubation step of 2.4 μM OmpP in 0.24 M urea for 5 min in the absence of lipid (green lines). Fits to a single exponential function are shown as black dashed lines. All reactions were carried out at an LPR of 3200:1 in 50 mM sodium phosphate, pH 8.0, 37 °C. The replicate transients of each experiment are shown.

5.2.1.3 PagP

The work described in Section 5.2.1.3 was carried out by Dr. Alice I. Bartlett, University of Leeds. It is included here as the prelude to the work carried out in this thesis. This work, including the work in the remainder of this chapter, has been published in: McMorran L. M.; Bartlett A. I.; Huysmans G. H. M.; Radford S. E. & Brockwell D. J. (2013) Dissecting the effects of periplasmic chaperones on the *in vitro* folding of the outer membrane protein PagP. *J. Mol. Biol.* **425**, 3178-3191

PagP was over-expressed and purified as described in Section 4.2.4. To characterise the folding of untagged PagP, an unfolded stock in 10 M urea was diluted into 100 nm *di*$C_{12:0}$PC LUVs at an LPR of 3200:1 in 50 mM glycine, pH 9.5 and folding was monitored using tryptophan fluorescence emission at 335 nm (following excitation at 280 nm). The observed transients were well described by single exponential functions over the time-course of the experiment when the final urea concentration was varied between 2 M and 4 M (Figure 5.10a). Interestingly, the observed folding rate constant increased approximately five-fold between 2 M and 4 M urea (Table 5.3), suggesting that increasing the concentration of urea is able to change the balance between folding

and aggregation such that folding is favoured. PagP has been reported to be particularly aggregation-prone when compared with other OMPs[164; 238] and when PagP folding is monitored using cold SDS-PAGE, higher urea concentrations result in fewer observable aggregates (Figure 5.11). Under these conditions, however, the reported folding rate constants in 2 M urea were independent of protein concentration between 0.4 μM and 0.04 μM PagP (Table 5.4, Figure 5.12), suggesting that the explanation for this phenomenon may be more complex.

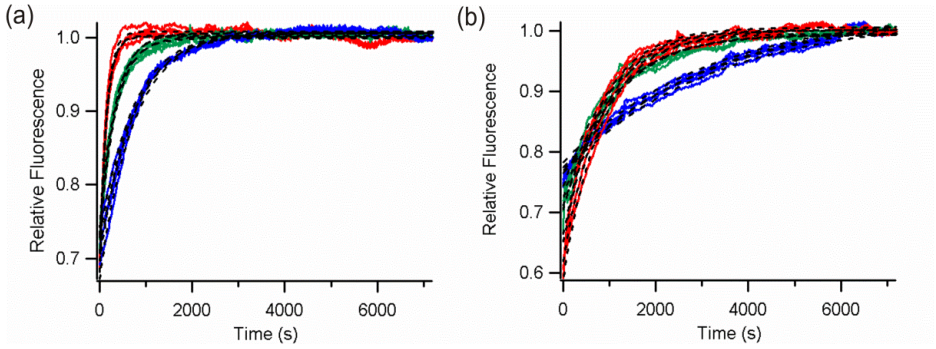

Figure 5.10 Effect of urea concentration and lipid composition on PagP folding kinetics. Kinetic traces have been normalised to the final fluorescence signal to highlight the changes in k_{obs} in 2 M (blue lines), 3 M (green lines) and 4 M (red lines) urea for PagP folding into (a) *di*$C_{12:0}$PC LUVs and (b) 80:20 *di*$C_{12:0}$PC:*di*$C_{12:0}$PG LUVs. Fits to single exponential functions are shown as black dashed lines. All kinetic samples contained 0.4 μM PagP, LPR 3200:1, 50 mM glycine, pH 9.5 and were measured at 37 °C. These data were acquired by Dr Alice I. Bartlett (University of Leeds).

Lipid	[Urea] (M)	Rate Constant ($\times 10^{-3}$ s^{-1})	Error in Global Fit ($\times 10^{-3}$ s^{-1})
di$C_{12:0}$PC	2.0	1.37	0.01
	3.0	2.90	0.01
	4.0	7.02	0.06
80:20 *di*$C_{12:0}$PC: *di*$C_{12:0}$PG	2.0	0.32	0.00
	3.0	0.92	0.01
	4.0	1.20	0.01

Table 5.3 Urea dependence of PagP folding into *di*$C_{12:0}$PC and 80:20 *di*$C_{12:0}$PC:*di*$C_{12:0}$PG liposomes. Rate constants were obtained by measuring tryptophan fluorescence emission at 335 nm over time and globally fitting the observed transients to a single exponential function. All kinetic samples contained 0.4 μM PagP, LPR 3200:1, 50 mM glycine, pH 9.5 and were measured at 37 °C. These data were acquired by Dr Alice I. Bartlett (University of Leeds).

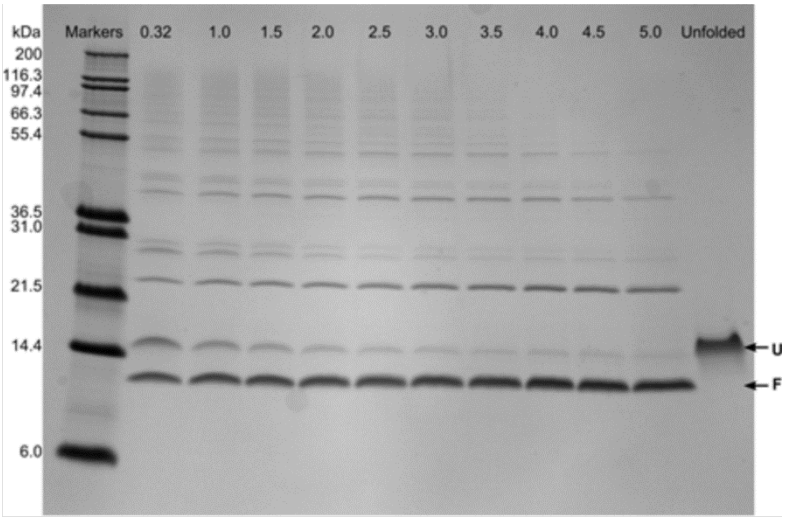

Figure 5.11 Urea-dependence of PagP folding monitored using cold SDS-PAGE. All samples contained 4 μM PagP and 100 nm *di*$C_{12:0}$PC LUVs at an LPR of 3200:1 in 10 mM glycine buffer, pH 9.5, 2 mM EDTA and were refolded at 37 ºC for 15 hours prior to 1:1 dilution with 2 × SDS-PAGE loading buffer. Lanes on the SDS-PAGE gel are numbered to indicate the final concentration of urea (M) in each sample. The unfolded and folded forms of PagP are denoted by U and F, respectively. These data were acquired by Dr Alice I. Bartlett (University of Leeds).

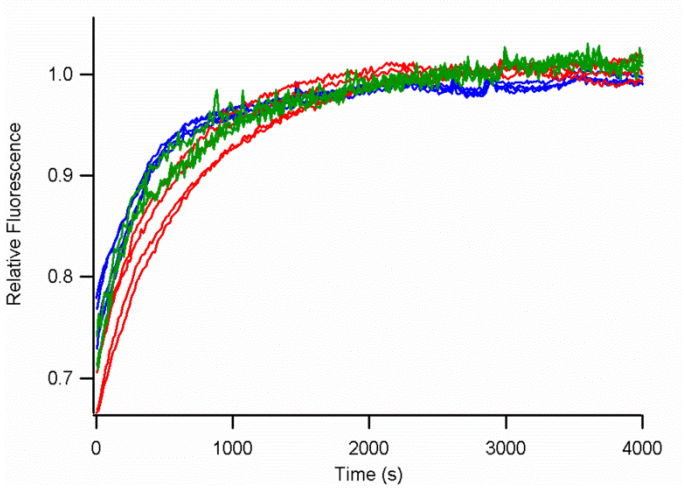

Figure 5.12 Protein concentration dependence of PagP folding into 100 nm *di*$C_{12:0}$PC LUVs. Kinetic traces are normalised to the final fluorescence signal. Samples contained 0.4 µM PagP (red lines), 0.1 µM (blue lines) or 0.04 µM PagP (green lines). All samples contained 50 mM glycine, pH 9.5, 2 M urea and were measured at 37 °C at an LPR of 3200:1. Note that an additional slow phase visible for 0.04 µM PagP was not observed reproducibly in replicate experiments and is not considered further. These data were acquired by Dr Alice I. Bartlett (University of Leeds).

PagP Concentration (µM)	Average Rate Constant ($\times 10^{-3}$ s^{-1})	Standard Error ($\times 10^{-3}$ s^{-1})
0.4	1.61	0.37
0.1	2.35	0.40
0.04	1.56	0.60

Table 5.4 Protein concentration dependence of PagP folding into 100 nm *di*$C_{12:0}$PC LUVs. Rate constants were obtained by measuring tryptophan fluorescence emission at 335 nm over time, and fitting the observed transients to a single exponential function. The mean of four transients was calculated and standard error (n = 4) was calculated. All kinetic samples contained LPR 3200:1, 50 mM glycine, pH 9.5 and were measured at 37 °C. These data were acquired by Dr Alice I. Bartlett (University of Leeds).

When these experiments were repeated using 100 nm 80:20 *di*$C_{12:0}$PC:*di*$C_{12:0}$PG LUVs, the obtained folding transients were similarly well described by a single exponential function over the time-course of the experiments (Figure 5.10b). The resulting rate constants were lower than those measured in *di*$C_{12:0}$PC LUVs, however, the same overall trend was observed with the rate constant increasing approximately four-fold

between 2 M and 4 M urea (Table 5.3). The lower folding rate constants in 80:20 *di*$C_{12:0}$PC:*di*$C_{12:0}$PG LUVs may reflect a less favourable folding reaction as these liposomes carry an overall negative surface charge and PagP, with a theoretical pI of 5.5 calculated using the ExPASy ProtParam Tool[251], will also be negatively charged under the conditions of the assay.

Folding of PagP into both *di*$C_{12:0}$PC and 80:20 *di*$C_{12:0}$PC:*di*$C_{12:0}$PG LUVs was successful, with reproducible exponential folding kinetics being observed in urea concentrations as low as 2 M. Folding in 50 mM glycine, pH 9.5, 37 °C was carried out either by direct dilution into liposomes from an unfolded stock in 10 M urea or by first incubating unfolded PagP at 0.24 M urea for 5 min before the addition to lipid and the resulting rate constants were compared. Direct dilution to 0.4 μM PagP in 100 nm *di*$C_{12:0}$PC LUVs at a final concentration of 2 M urea yielded a rate constant of $1.34\times10^{-3} \pm 7.7\times10^{-5}$ s^{-1}. When 2.4 μM PagP was incubated in 0.24 M urea prior to six-fold dilution into 100 nm *di*$C_{12:0}$PC LUVs, a similar rate constant of $1.52\times10^{-3} \pm 9.8\times10^{-6}$ s^{-1} was observed, suggesting that PagP remains folding-competent in these conditions. Taken together, these results indicate PagP as an excellent model protein for the study of the effects of periplasmic chaperones on the folding of an OMP.

5.2.2 PagP Populates a Collapsed, Folding-Competent State in Very Low Concentrations of Urea

The conformation adopted by PagP in 0.24 M urea, pH 9.5, 37 °C in the absence of lipid was further characterised using spectroscopic methods as it represents an early species in the folding pathway that could be recognised by chaperones before liposome insertion occurs. Analytical gel filtration of PagP in 0.24 M urea, 50 mM glycine, pH 9.5 showed that PagP elutes as a single peak with an elution volume of approximately 8 ml (Figure 5.13), suggesting that PagP populates a single non-aggregated species under these conditions.

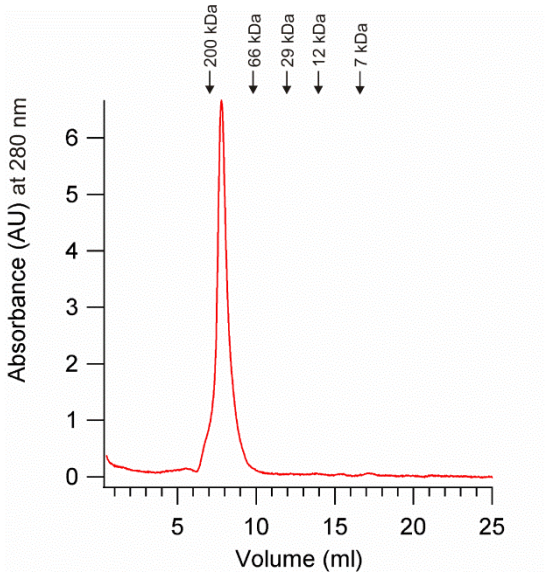

Figure 5.13 PagP populates a single, non-aggregated conformation in 0.24 M urea. 2 μM PagP in 50 mM glycine, pH 9.5 containing 0.24 M urea was incubated for 5 min at room temperature before injection on to a Superdex 75 10/300 GL column. Elution volumes of calibrant proteins are indicated with arrows.

The far-UV CD spectrum of PagP in 0.24 M urea is distinct from both the spectrum of PagP folded in *di*$C_{12:0}$PC LUVs and the spectrum of PagP unfolded in 10 M urea (Figure 5.14a). The spectrum of PagP folded in *di*$C_{12:0}$PC LUVs (Figure 5.14a) shows a negative peak at 218 nm, indicative of β-sheet secondary structure, and a Cotton band at 232 nm, which has been shown to arise from interactions between Tyr-26 and Trp-66 and is characteristic of natively folded PagP[235; 238]. The Cotton band is not observed in the spectrum of the folding-competent PagP species in 0.24 M urea (Figure 5.14a), however, the negative ellipticity around 215 nm suggests the presence of β-sheet structure. By contrast, the fluorescence emission spectra of the folding-competent state in 0.24 M urea and the native state in *di*$C_{12:0}$PC LUVs have identical intensity-averaged wavelength (< λ >) values of 342.9 nm but differ in signal intensity (Figure 5.14b). Both spectra show a higher fluorescence yield and were blue-shifted when compared with that of unfolded PagP in 10 M urea (Figure 5.14), suggesting the folding-competent state is collapsed in 0.24 M urea such that the tryptophan residues are partially protected from exposure to the solvent.

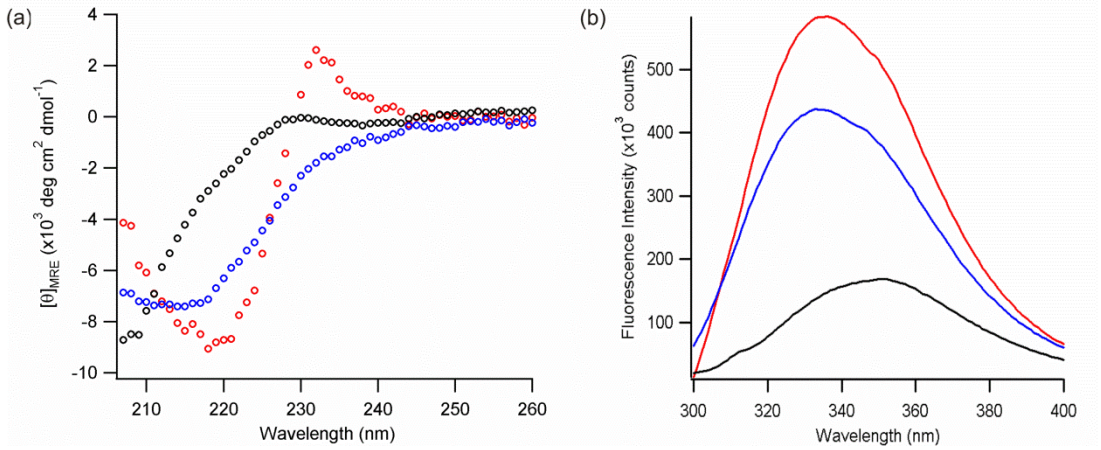

Figure 5.14 PagP in 0.24 M urea populates a collapsed state with residual structure as assessed by far UV-CD and tryptophan fluorescence emission spectroscopy. (a) The far-UV CD spectra of 10 µM PagP refolded in 50 mM glycine, pH 9.5 in *di*$C_{12:0}$PC liposomes (red circles); 10 µM PagP in 0.24 M urea, 50 mM glycine, pH 9.5 (blue circles) and 10 µM PagP unfolded in 10 M urea, 50 mM glycine, pH 9.5 (black circles). (b) The fluorescence emission spectra of 0.4 µM PagP folded in *di*$C_{12:0}$PC liposomes, 50 mM glycine, pH 9.5 (red line); 0.4 µM PagP in 0.24 M urea, 50 mM glycine, pH 9.5 (blue line) and 0.4 µM PagP unfolded in 10 M urea, 50 mM glycine, pH 9.5 (black line). The $\langle\lambda_{320\text{-}370nm}\rangle$ calculated for the unfolded spectrum in 10 M urea is 346.3 nm. The $\langle\lambda_{320\text{-}370nm}\rangle$ values for the folded spectrum and spectrum in 0.24 M urea are 342.9 nm. All spectra were acquired at 37 °C.

The fluorescent molecule 8-anilinonaphthalene-1-sulfonic acid (ANS), binds to the exposed hydrophobic areas on the surface of partially folded proteins, causing a characteristic large increase in the fluorescence emission of the ANS and also a blue-shift in the peak emission wavelength (λ_{max})[285; 286]. Fluorescence emission spectra of ANS were acquired in both the presence and absence of 1 µM PagP in 0.24 M urea to probe the conformation of the protein. In 10 M urea the emission spectrum of ANS had a λ_{max} of 509 nm and was unchanged by the addition of PagP (Figure 5.15), suggesting that the protein is in a fully unfolded conformation under these conditions. In contrast, the emission spectrum of ANS in 0.24 M urea shows a shift in λ_{max} to 476 nm in the presence of 1 µM PagP, with an associated three-fold increase in intensity (Figure 5.15). These data show that in 0.24 M urea, PagP adopts a soluble, molten globule[287] like conformation with residual secondary structure, as observed by far-UV CD, but with loosely defined tertiary structure and exposed hydrophobic amino acid side chains.

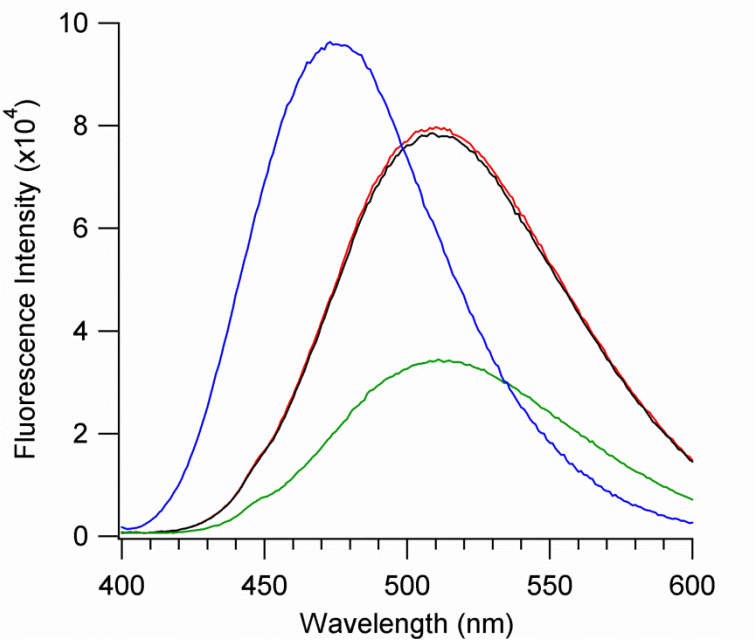

Figure 5.15 ANS binding to PagP reveals a molten globule like state in 0.24 M urea. The fluorescence spectrum of 250 μM ANS in 10 M urea in the absence (red) and presence (black) of 1 μM PagP and in 0.24 M urea in the absence (green) and presence (blue) of 1 μM PagP. All samples contained 50 mM glycine, pH 9.5 and were measured at 37 °C.

5.2.3 SurA Does Not Affect PagP Folding in Vitro

SurA is a periplasmic chaperone from *E. coli* which contains two PPIase domains[86]. Deletion of these domains showed that SurA can still function *in vivo* and exhibits a chaperone activity which is not dependent on PPIase activity[87]. The observation that SurA can be cross-linked to BamA *in vivo*[94], the ability of SurA to assist the BAM complex in the folding of OmpT *in vitro*[137] and proteomics experiments on SurA depletion strains of *E. coli*[85; 98] have all led to the hypothesis that SurA is the primary chaperone involved in transport of unfolded OMPs across the periplasm, making it an obvious choice for study.

SurA was over-expressed in BL21 (DE3) cells and purified according to the method of Hagan *et al*[137] as described in Section 2.4.6. Briefly, purification from whole cell lysate was achieved by nickel affinity chromatography in 20 mM Tris-HCl, 5 mM imidazole, pH 8.0. SurA was eluted using a stepwise gradient against 20 mM Tris-HCl, 500 mM imidazole, pH 8.0 (Figure 5.16a). The N-terminal His-tag was removed by thrombin

cleavage and the cleaved protein purified using a second nickel affinity chromatography purification step. The resulting protein was greater than 95 % pure (Figure 5.16b) and was typically obtained in yields of 6 mg pure protein per litre of culture.

Prior to inclusion in PagP folding assays, the stability of SurA was assessed using equilibrium denaturation experiments. Tryptophan fluorescence emission was chosen as a suitable probe since SurA contains four tryptophan residues which are located throughout the structure in the P1, P2 and C-terminal domains (Figure 1.11). The resulting denaturation curves indicate that SurA remains natively folded in buffer containing 2 M or 3 M urea (Figure 5.17) suggesting it could be added directly into folding assays at these concentrations of denaturant. Equilibrium denaturation experiments of SurA were conducted by Dr. Alice I. Bartlett, University of Leeds.

Folding assays were carried out by diluting unfolded PagP in 10 M urea into a three-fold molar excess of SurA in 50 mM glycine, pH 9.5 containing a final concentration of 0.24 M urea and 2.4 μM PagP and incubating for 5 min. This mixture was then diluted six-fold into either 100 nm *di*$C_{12:0}$PC or 100 nm 80:20 *di*$C_{12:0}$PC:*di*$C_{12:0}$PG LUVs at a final urea concentration of 2–3 M and the folding reaction followed using tryptophan fluorescence as a probe.

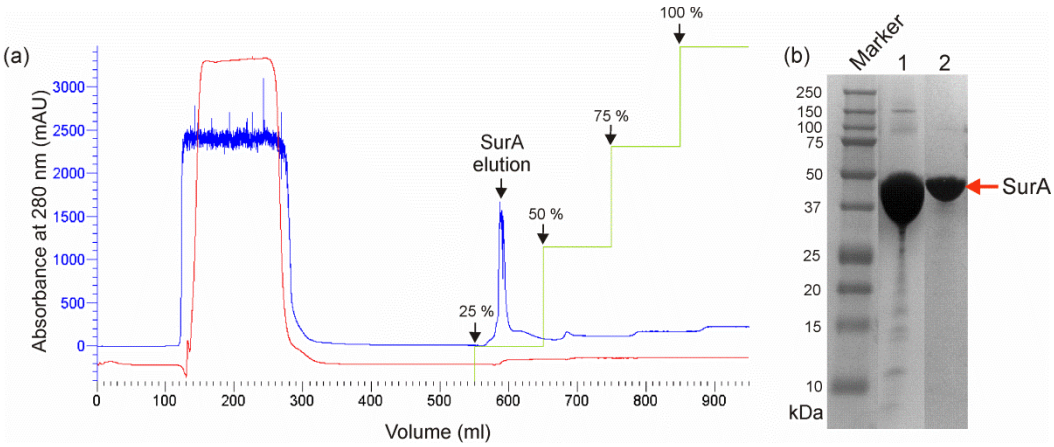

Figure 5.16 Purification of SurA. (a) Typical elution profile of SurA during the first nickel affinity purification showing the A_{280nm} (blue line), solution conductivity (red line) and percentage of imidazole-containing buffer (green line). The step gradient of imidazole-containing buffer and SurA elution peak are indicated with arrows. (b) SDS-PAGE of SurA-containing fraction (1) after first nickel affinity chromatography purification and (2) after second nickel affinity chromatography purification following thrombin cleavage of the His-tag. The size in kiloDaltons (kDa) of the protein markers is indicated.

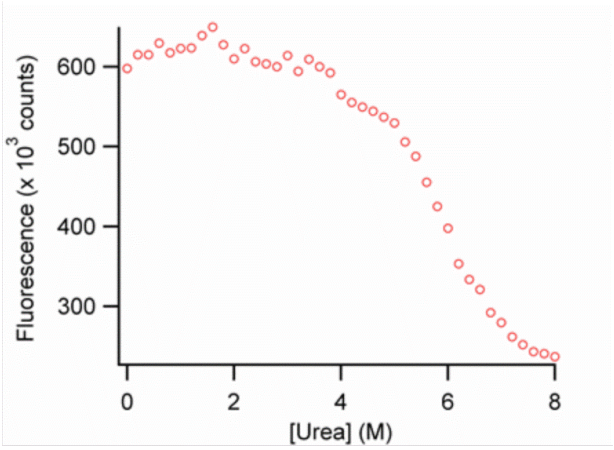

Figure 5.17 Equilibrium unfolding of SurA monitored using tryptophan fluorescence emission. SurA dissolved in 50 mM glycine buffer, pH 9.5 was diluted into separate aliquots of 50 mM glycine buffer containing different concentrations of urea (0–8 M in 0.2 M increments). The final concentration of protein was 2.5 μM. The aliquots were equilibrated overnight (16 h) at 37 °C before measurement. The fluorescence emission intensity was measured at 335 nm (following excitation at 280 nm) for 60 s and the average signal calculated. These data were acquired by Dr Alice I. Bartlett (University of Leeds).

Inclusion of SurA in the PagP folding assay had no effect on the observed folding rate constants in either 100 nm *di*$C_{12:0}$PC (k_{obs} = 1.34×10^{-3} ± 4.5×10^{-5} s^{-1} for PagP alone and

k_{obs} = 1.41×10^{-3} ± 1.3×10^{-4} s^{-1} with SurA in buffer containing 2 M urea) (Figure 5.18a, top panel, Figure 5.31, Table 5.5) or 100 nm 80:20 *di*$C_{12:0}$PC:*di*$C_{12:0}$PG LUVs (k_{obs} = 7.9×10^{-4} ± 1.4×10^{-4} s^{-1} for PagP alone and k_{obs} = 6.3×10^{-4} ± 1.7×10^{-4} s^{-1} with SurA) (Figure 5.18b, top panel, Figure 5.31, Table 5.5). Cold SDS-PAGE analysis at the endpoint of the reaction revealed no change in folding yield when SurA was added to the assay in either lipid composition (Figure 5.18a, b; bottom panels). These results indicate that under these conditions, the presence of SurA has no effect on either the rate or yield of PagP folding into liposomes, regardless of the net charge carried on the lipid surface. Kinetic folding assays and endpoint cold SDS-PAGE assays were carried out by Dr. Alice I. Bartlett, University of Leeds.

The lack of any influence of SurA on the folding kinetics or yield of PagP raised the question of whether these two proteins interact *in vitro*. This was first investigated by immobilising HT PagP (10 μM) on nickel Sepharose resin, then washing over SurA in 50 mM glycine, pH 9.5. Any protein remaining bound to the resin was then eluted using a mixture of 5 M urea and 500 mM imidazole. Analysis of the fractions by SDS-PAGE showed that SurA remained in the unbound fraction, suggesting it forms no interaction with HT PagP (Figure 5.19). Additionally, analytical gel filtration of PagP in 50 mM glycine, pH 9.5, 0.24 M urea showed no change in peak size or elution volume upon addition of a six-fold molar excess of SurA suggesting there is no interaction between these two proteins (Figure 5.20). This was confirmed by SDS-PAGE analysis of the eluted fractions (Figure 5.20).

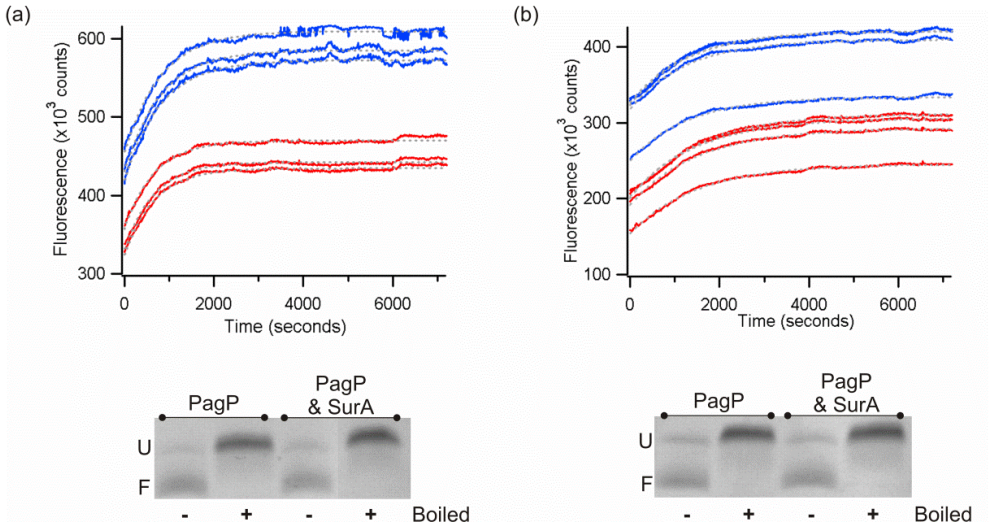

Figure 5.18 Kinetic traces for PagP folding into (a) *di*$C_{12:0}$PC and (b) 80:20 *di*$C_{12:0}$PC:*di*$C_{12:0}$PG liposomes in the presence (blue lines) or absence (red lines) of a three-fold molar excess of SurA. Dashed grey lines represent the fits of the data to a single exponential function. Three replicate samples of each assay are shown. The differences in the final fluorescence signal arise from very small PagP concentration differences between samples, which appear due to the large molar extinction coefficient (82390 M^{-1} cm^{-1}) of PagP. Note, however, that each sample refolded successfully with a similar folding yield. The lower panels show cold SDS-PAGE analysis of PagP samples allowed to fold overnight in the presence of a six-fold molar excess of SurA. The folded and unfolded forms of PagP are denoted by F and U, respectively. Kinetic samples contained 0.4 μM PagP and 2–3 M urea. SDS-PAGE gel samples contained 4 μM PagP and 1 M urea. All samples had an LPR of 3200:1, contained 50 mM glycine, pH 9.5 and were measured at 37 °C. These data were acquired by Dr Alice I. Bartlett (University of Leeds).

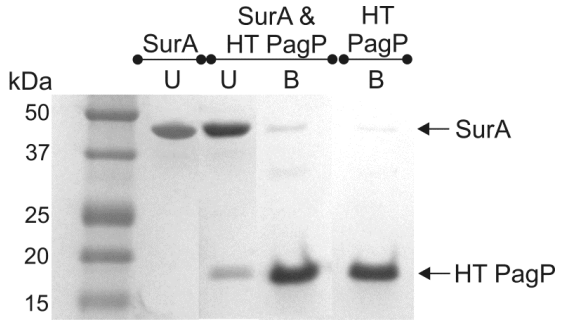

Figure 5.19 SurA does not bind to HT PagP under the conditions used for the PagP folding kinetic assays. 10 μM HT PagP was immobilised on nickel Sepharose resin before incubation with 10 μM SurA in 50 mM glycine buffer, pH 9.5. Bound and unbound fractions were analysed by SDS-PAGE. SurA remains in the unbound (U) fraction, while HT PagP is bound (B) to the resin until eluted with 500 mM imidazole. Control experiments containing HT PagP or SurA only were conducted under identical conditions for comparison.

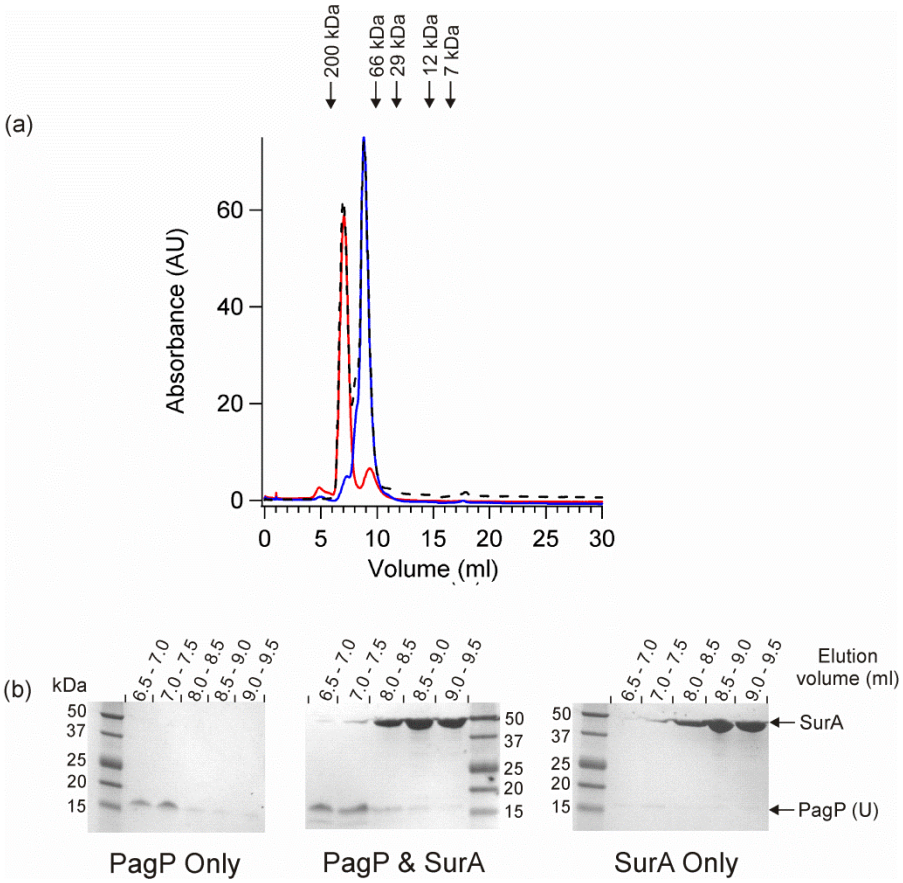

Figure 5.20 PagP and SurA do not interact under the conditions used in the kinetic folding assays. (a) 10 µM PagP was added to 60 µM SurA in 50 mM glycine, pH 9.5 containing 0.24 M urea and incubated for 5 min at room temperature before injection on to a Superdex 75 10/300 GL column (black dashed line). Samples containing either 10 µM PagP (red line) or 60 µM SurA (blue line) in 50 mM glycine, pH 9.5 containing 0.24 M urea were also analysed. Elution volumes of calibrant proteins are indicated with arrows. (b) Fractions were collected and analysed by SDS-PAGE.

5.2.4 Skp Has a Dramatic Effect on PagP Folding

Skp is a trimeric chaperone from *E. coli*, which has been suggested to have a role in rescuing OMPs which deviate from the main folding pathway in the periplasm[94; 98]. Skp has been shown by NMR to bind to both OmpX[92] and OmpA[102] *in vitro*, forming stable complexes which maintain the transmembrane domains of the OMPs in unfolded conformations. Additionally, fluorescence spectroscopy revealed that Skp could bind to bacterial OMPs with dissociation constants in the nanomolar range but binding to the human mitochondrial OMP, VDAC1, was not observed[101; 174]. Using cold SDS-PAGE analysis, it has been suggested that Skp impedes the folding of OmpA into zwitterionic bilayers while accelerating folding into negatively charged bilayers[161]. Exactly how

these observations relate to the role of Skp *in vivo* remains poorly understood, making Skp an attractive target of study.

Skp was over-expressed in BL21 (DE3) cells and purified as described in Section 2.4.7. Briefly, following overnight expression of Skp, polymyxin B sulphate was used to extract the periplasmic proteins from the harvested cells. The periplasmic extract was then purified first by using anion exchange in 20 mM Tris-HCl, 100 mM NaCl, pH 8.0 and then by further purifying the flow-through by cation exchange in the same buffer. Skp was eluted from the cation exchange column using a continuous gradient of increasing NaCl in 20 mM Tris-HCl, pH 8.0 (Figure 5.21a) and the purest fractions, as judged by SDS-PAGE, were retained. The resulting protein was greater than 95 % pure (Figure 5.21b) and was typically obtained in yields of 3 mg pure protein per litre of culture.

Prior to inclusion in the PagP folding assay, equilibrium denaturation experiments were carried out to determine the stability of Skp. Far-UV CD was chosen as an appropriate probe to follow unfolding since Skp contains no tryptophan residues and thus is only weakly excited by UV light at a wavelength of 280 nm. In 50 mM glycine, pH 9.5, Skp is seen to be unfolded by the addition of urea with the unfolding midpoint ($[urea]_{50\%}$) occurring at approximately 1.4 M urea (Figure 5.22). In spite of the low stability of Skp, kinetic assays were carried out at a final urea concentration of 2 M under the assumption that first allowing Skp to bind to PagP in 0.24 M urea may stabilise the Skp trimer sufficiently such that any effect on PagP folding could still be observed. Binding of Skp to OMPs has been observed to occur with a stoichiometry of one Skp trimer per OMP monomer[101; 159], so a two-fold molar excess of Skp trimers were added to the PagP kinetic assay to ensure complete PagP binding.

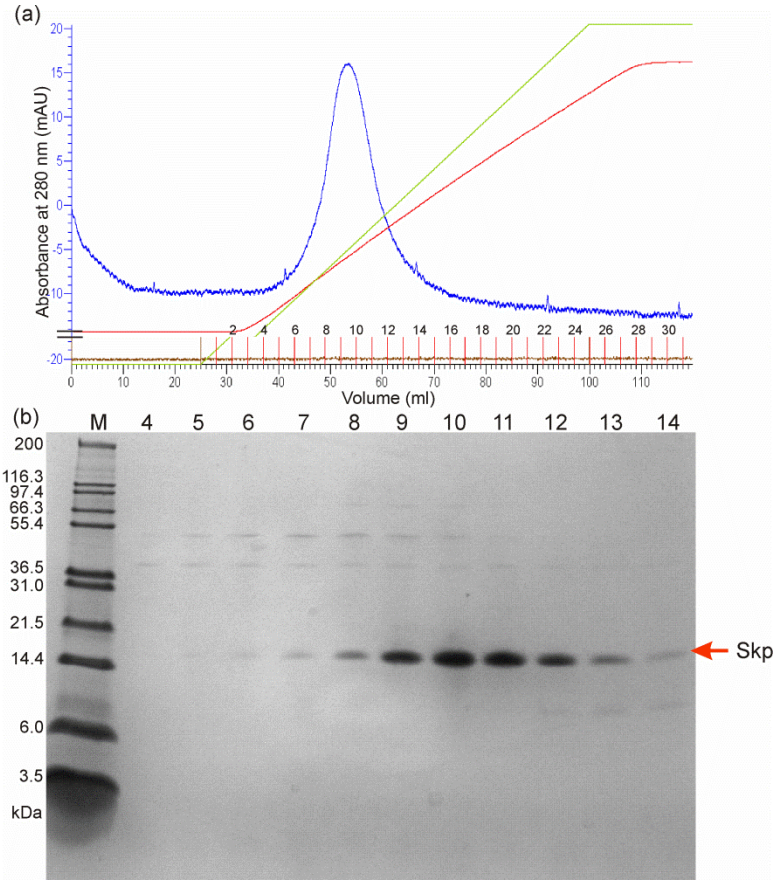

Figure 5.21 Purification of Skp. (a) Typical elution profile of Skp from a cation exchange column showing the A_{280nm} (blue line), solution conductivity (red line) and gradient of salt-containing buffer (green line). Eluted fraction numbers are indicated on the graph. Fractions 9–12 were typically retained. (b) SDS-PAGE of Skp-containing fractions after cation exchange. Lanes on the gel are labelled using the same fraction numbers as in (a). The size in kiloDaltons (kDa) of the protein markers (M) is indicated.

Inclusion of Skp trimers in the PagP folding assay showed a dramatic impact on the kinetics of PagP folding into 100 nm *di*$C_{12:0}$PC or 100 nm 80:20 *di*$C_{12:0}$PC:*di*$C_{12:0}$PG LUVs in a manner which is dependent on liposome composition (Figure 5.23). The folding of PagP into zwitterionic *di*$C_{12:0}$PC LUVs in the presence of Skp is inhibited such that no fluorescence change is observed over the 2 h time-course (Figure 5.23a, top and middle, Figure 5.31, Table 5.5). This finding was verified using cold SDS-PAGE to follow the folding of PagP in the presence and absence of Skp (Figure 5.24). When the folding reaction was allowed to proceed for 16 h, the final folded yield of PagP was unchanged in the presence of Skp as determined by cold SDS-PAGE (Figure 5.23a, bottom). By contrast, in 80:20 *di*$C_{12:0}$PC:*di*$C_{12:0}$PG LUVs, which are negatively

charged, Skp increased the observed rate constant of PagP folding by approximately nine-fold (k_{obs} PagP folding alone = $2.8\times10^{-4} \pm 1.8\times10^{-5}$ s^{-1}, k_{obs} PagP folding in the presence of Skp = $2.4\times10^{-3} \pm 4.3\times10^{-4}$ s^{-1}) (Figure 5.23b, top and middle, Figure 5.31, Table 5.5). Cold SDS-PAGE analysis of the folding reaction after 16 h revealed that in spite of the order of magnitude increase in PagP folding rate constant, the folding yield remains unchanged in the presence of Skp (Figure 5.23b, bottom). Skp is a very basic protein with a pI around 9.5 and a large macrodipole moment on the chaperone surface[105]. Together with the data described above, this suggests that electrostatics play an important role in the modulation of Skp-mediated membrane delivery as has been previously suggested[105]. Kinetic folding assays and endpoint cold SDS-PAGE assays (but not the cold SDS-PAGE time-course experiments) were carried out by Dr. Alice I. Bartlett, University of Leeds.

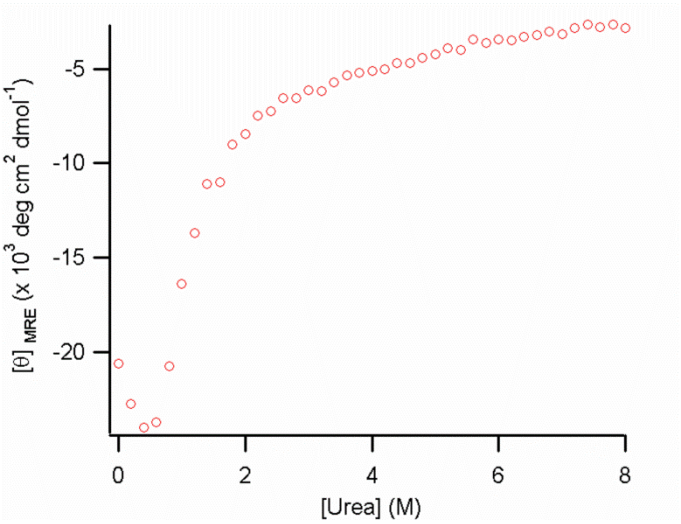

Figure 5.22 Equilibrium unfolding of Skp monitored using far-UV CD. Skp dissolved in 50 mM glycine buffer, pH 9.5 was diluted into separate aliquots of 50 mM glycine buffer containing different concentrations of urea (0–8 M in 0.2 M increments). The final concentration of Skp was 15 μM. Aliquots were equilibrated for 16 h at 37 °C then the ellipticity at 222 nm was recorded for 60 s and the average signal calculated.

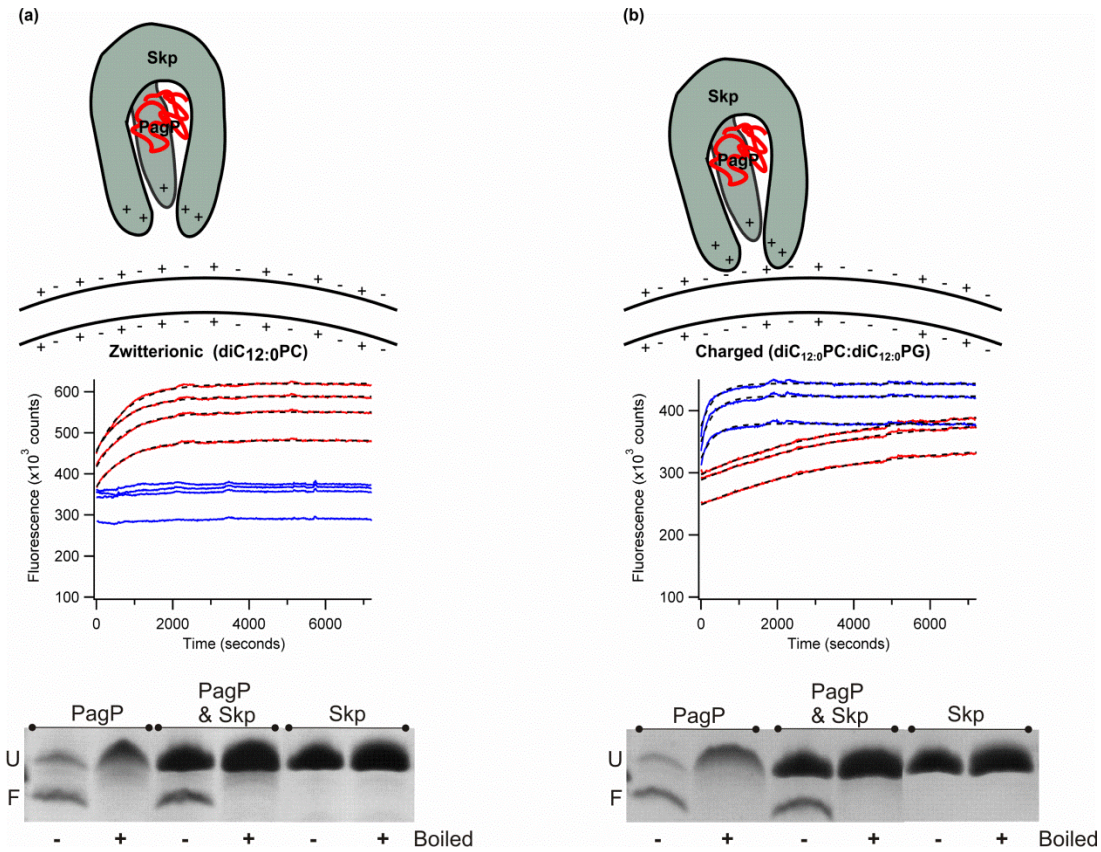

Figure 5.23 Kinetic traces of PagP folding into (a) *di*$C_{12:0}$PC and (b) 80:20 *di*$C_{12:0}$PC:*di*$C_{12:0}$PG liposomes in the presence (blue lines) or absence (red lines) of a two-fold molar excess of Skp trimers. Dashed black lines represent the fits of the data to a single exponential function. The differences in the final fluorescence signal arise from small PagP concentration differences between samples, which appear due to the large molar extinction coefficient (82390 M^{-1} cm^{-1}) of PagP. Note, however, that each sample refolded successfully with a similar folding yield. The lower panels show cold SDS-PAGE analysis of PagP samples allowed to fold overnight in the presence of a two-fold molar excess of Skp trimers. The folded and unfolded forms of PagP are denoted by F and U, respectively. Kinetic samples contained 0.4 µM PagP and 2 M urea in 50 mM glycine, pH 9.5. SDS-PAGE gel samples contained 4 µM PagP and 1 M urea in 50 mM glycine, pH 9.5. All samples had an LPR of 3200:1 and were measured at 37 °C. These data were acquired by Dr Alice I. Bartlett (University of Leeds).

Skp has recently been shown to form a tightly bound complex with PagP, having a measured dissociation constant of 11.8 nM at pH 8.0[174]. To confirm Skp binding under the experimental conditions described here, HT PagP (10 µM) was immobilised on nickel Sepharose resin as described in Section 2.6.5 then incubated with an equal concentration of Skp trimers in 50 mM glycine, pH 9.5 (Figure 5.25). In the presence of HT PagP, Skp co-elutes from the resin with HT PagP, as observed in SDS-PAGE experiments, when 500 mM imidazole is added suggesting that, in contrast with SurA,

Skp readily binds to HT PagP (Figure 5.25a). This result was confirmed by Western blotting with an anti-Histag antibody, as the apparent molecular weights of Skp and HT PagP are very similar and hence are difficult to resolve on a Coomassie Blue stained gel (Figure 5.25b).

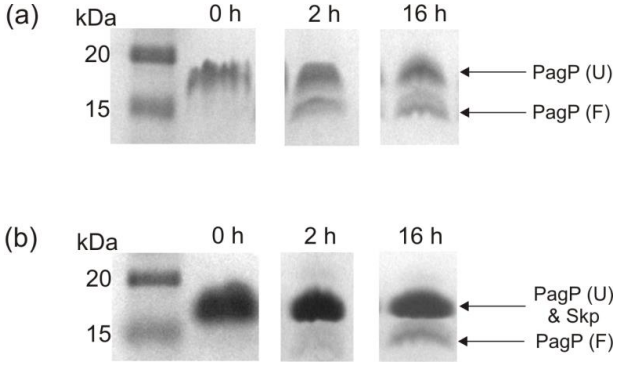

Figure 5.24 Skp retards the folding of PagP into *di*$C_{12:0}$PC liposomes over a 2 h time-course. (a) PagP folding into *di*$C_{12:0}$PC liposomes in the absence of Skp. (b) PagP folding into *di*$C_{12:0}$PC liposomes in the presence of Skp. 12 μM PagP was incubated (with a two-fold excess of Skp trimers, if appropriate) for 5 min at room temperature in 50 mM glycine, pH 9.5 containing 0.24 M urea before a three-fold dilution into *di*$C_{12:0}$PC liposomes (LPR 3200:1) in 50 mM glycine, pH 9.5 containing 2 M urea at 37 °C. Samples were removed at 0, 2 and 16 hours and analysed by cold SDS-PAGE immediately.

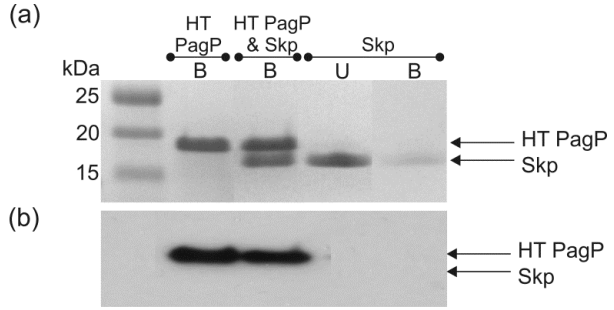

Figure 5.25 Skp binds to HT PagP under the conditions used for the PagP folding kinetic assays as analysed by (a) SDS-PAGE and (b) Western blot analysis with an anti-Histag antibody. 10 μM HT PagP was immobilised on nickel Sepharose resin before incubation with 10 μM Skp in 50 mM glycine buffer, pH 9.5. Control experiments containing HT PagP or Skp only were conducted under identical conditions for comparison. Skp alone does not bind to the resin and is present in the unbound (U) fraction, while in the presence of HT PagP, Skp co-elutes in the bound (B) fraction.

5.2.5 Skp Activity is Influenced by Electrostatic Interactions

To further investigate the extent to which the modulation of Skp-mediated delivery of PagP to zwitterionic and negatively charged liposomes occurs by electrostatic interactions, another highly basic protein was added to the folding assay in place of Skp. Hen egg-white lysozyme was added at a three-fold molar excess over PagP to assays examining folding into both 100 nm *di*$C_{12:0}$PC or 100 nm 80:20 *di*$C_{12:0}$PC:*di*$C_{12:0}$PG LUVs (Figure 5.26). In *di*$C_{12:0}$PC LUVs, lysozyme showed a similar inhibition of PagP folding over the 2 h fluorescence time-course to that of PagP incubated with an excess of Skp (Figure 5.26a, Figure 5.31, Table 5.5). This finding could not be verified using a cold SDS-PAGE time-course due to the folded conformation of untagged PagP and lysozyme running at the same apparent molecular weight (for example, see Figure 5.32). Inclusion of lysozyme in 80:20 *di*$C_{12:0}$PC:*di*$C_{12:0}$PG LUVs gave rise to a small, approximately two-fold increase in observed rate constant for PagP folding ($k_{obs} = 7.9\times10^{-4} \pm 1.4\times10^{-4}$ s^{-1} for PagP alone and $k_{obs} = 1.5\times10^{-3} \pm 1.8\times10^{-4}$ s^{-1} with lysozyme) (Figure 5.26b, Figure 5.31, Table 5.5). This small increase is in stark contrast with the order of magnitude increase observed in the presence of Skp. The data in 80:20 *di*$C_{12:0}$PC:*di*$C_{12:0}$PG LUVs were acquired by Dr. Alice I. Bartlett (University of Leeds).

Binding of lysozyme to PagP was investigated using both analytical gel filtration and nickel affinity chromatography (Figure 5.27). Analytical gel filtration in 50 mM glycine, pH 9.5 showed that the elution peak of PagP is shifted in the presence of lysozyme, indicating an interaction between these two proteins (Figure 5.27a). The atypical elution profile in the presence of lysozyme arises from the interaction of lysozyme with the Superdex resin, as evidenced by lack of an elution profile of lysozyme alone under the same buffer conditions. The manufacturer's literature also indicates that very basic proteins interact with the resin in the absence of salt[288]. As a result of the interaction of lysozyme with Superdex resin, low (2–6 μM) protein concentrations were used for this assay, excluding the possibility of SDS-PAGE analysis of the eluted fractions. Binding was additionally testing by immobilising HT PagP on nickel Sepharose beads followed by incubation with an equal concentration (10 μM) of lysozyme in 50 mM glycine, pH 9.5 (Figure 5.27b). When the eluent was

analysed by SDS-PAGE, lysozyme was seen to have co-eluted with HT PagP demonstrating the ability of these two proteins to interact.

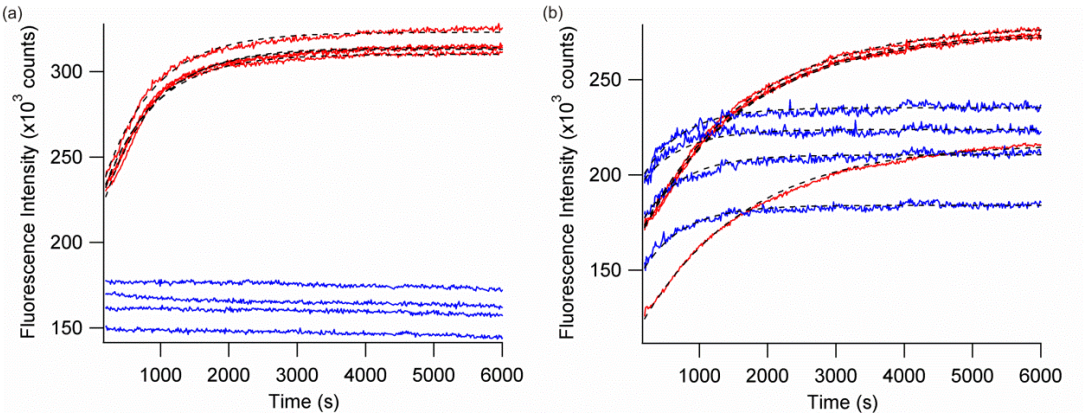

Figure 5.26 Kinetic traces of PagP folding into (a) *di*$C_{12:0}$PC and (b) 80:20 *di*$C_{12:0}$PC:*di*$C_{12:0}$PG liposomes in the presence (blue lines) or absence (red lines) of a three-fold molar excess of lysozyme. Dashed black lines represent the fits of the data to a single exponential function. The differences in the final fluorescence signal arise from small PagP concentration differences between samples, which appear due to the large molar extinction coefficient (82390 M^{-1} cm^{-1}) of PagP. All samples contained 0.4 µM PagP and 2–3 M urea in 50 mM glycine, pH 9.5, had an LPR of 3200:1 and were measured at 37 °C. The data shown in panel (b) were acquired by Dr Alice I. Bartlett (University of Leeds).

The ability of lysozyme, a basic protein which does not function as a chaperone, to partially mimic the activity of Skp provides further evidence that electrostatic interactions are important in the mechanism of action of Skp. To investigate whether the changes in rate constant observed when PagP folds in the presence of Skp are dependent on complementary electrostatic interactions between the protein and membrane, the kinetic refolding assays were repeated in the presence of NaCl. Inclusion of 200 mM NaCl in the folding buffer caused relief of Skp-mediated inhibition of PagP folding into 100 nm *di*$C_{12:0}$PC LUVs and resulted in similar PagP folding rate constants in the presence and absence of Skp (k_{obs} = 8.1×10^{-4} ± 9.6×10^{-5} s^{-1} and k_{obs} = 1.2×10^{-3} ± 1.4×10^{-4} s^{-1}, respectively) (Figure 5.28a, Figure 5.31, Table 5.5). In 100 nm 80:20 *di*$C_{12:0}$PC:*di*$C_{12:0}$PG LUVs, data could not be acquired in 200 mM NaCl as the resulting traces were too noisy to fit with confidence. The presence of salt has been reported previously to increase the aggregation propensity of OMPs *in vitro*[164] and this, coupled with the slow folding rate constants observed for PagP folding in

80:20 *di*$C_{12:0}$PC:*di*$C_{12:0}$PG liposomes (Figure 5.10, Table 5.3), could account for the difficulties encountered. Reduction of the final NaCl concentration to 100 mM, however, was sufficient to yield reproducible kinetic transients with $k_{obs} = 5.6\times10^{-4} \pm 6.2\times10^{-5}$ s^{-1} for PagP alone and $k_{obs} = 1.5\times10^{-3} \pm 9.4\times10^{-5}$ s^{-1} with Skp (Figure 5.28b, Figure 5.31, Table 5.5). These data show that the presence of NaCl reduces the rate enhancement of PagP folding by Skp from nine-fold in the absence of salt to three-fold in the presence of salt. These data, together with the relative effects of Skp and lysozyme, suggest that while electrostatic interactions play a role in the action of Skp, they cannot fully account for the observed changes. The data in *di*$C_{12:0}$PC LUVs were acquired by Dr. Alice I. Bartlett (University of Leeds).

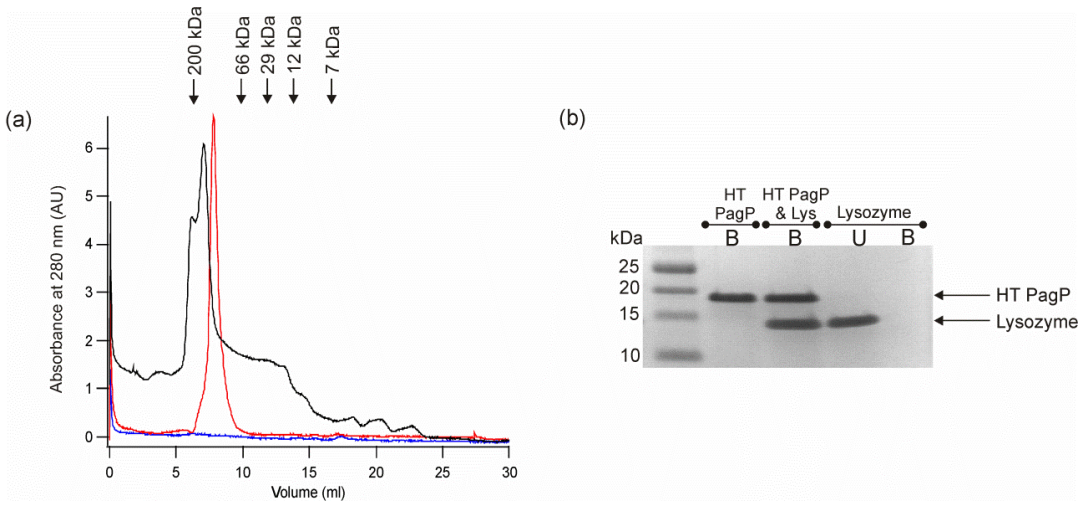

Figure 5.27 Lysozyme interacts with PagP under the conditions used in the kinetic folding assays as shown by (a) analytical gel filtration and (b) nickel affinity chromatography. For analytical gel filtration, 2 μM PagP was added to 6 μM lysozyme in 50 mM glycine, pH 9.5 containing 0.24 M urea and incubated for 5 min at room temperature before injection on to a Superdex 75 10/300 GL column (black line). Samples containing either 2 μM PagP (red line) or 6 μM lysozyme (blue line) in 50 mM glycine, pH 9.5 containing 0.24 M urea were also analysed. Note that the lysozyme interacts with the Superdex resin under these conditions, preventing its elution when injected alone and causing the atypical elution profile when injected with PagP. For the nickel affinity assay, 10 μM HT PagP was immobilised on nickel Sepharose resin before incubation with 10 μM lysozyme in 50 mM glycine buffer, pH 9.5. Control experiments containing HT PagP or lysozyme only were conducted under identical conditions for comparison. Lysozyme alone does not bind to the resin and is present in the unbound (U) fraction, while in the presence of HT PagP, lysozyme co-elutes in the bound (B) fraction.

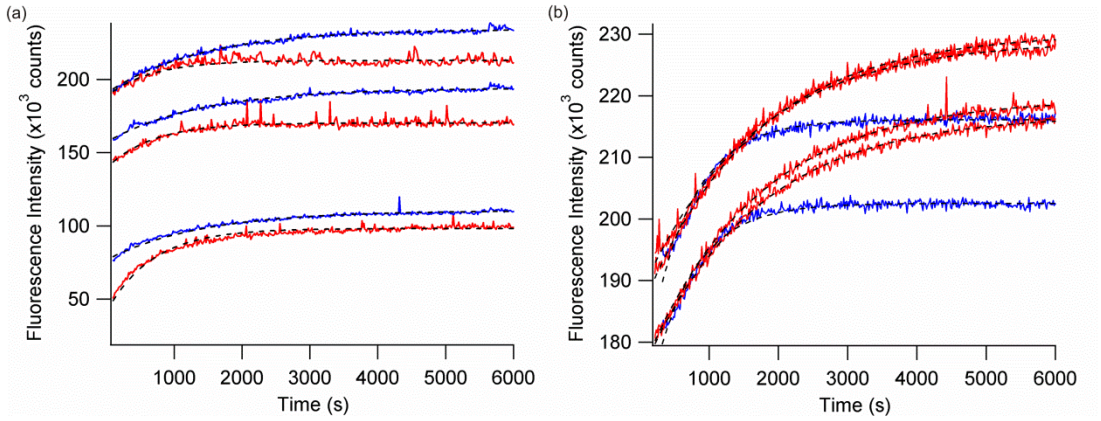

Figure 5.28 Effect of NaCl on Skp-mediated refolding of PagP. Kinetic traces of PagP folding into (a) *di*$C_{12:0}$PC liposomes in 200 mM NaCl and (b) 80:20 *di*$C_{12:0}$PC:*di*$C_{12:0}$PG liposomes in 100 mM NaCl in the presence (blue lines) or absence (red lines) of a two-fold molar excess of Skp trimers. Dashed black lines represent the fits of the data to a single exponential function. The differences in the final fluorescence signal arise from small PagP concentration differences between samples, which appear due to the large molar extinction coefficient (82390 M^{-1} cm^{-1}) of PagP. All samples contained 0.4 μM PagP and 2 M urea in 50 mM glycine, pH 9.5, had an LPR of 3200:1 and were measured at 37 °C. The data in panel (a) were acquired by Dr Alice I. Bartlett (University of Leeds).

One possible explanation for the ability of NaCl to influence Skp-mediated delivery of PagP to membranes is that electrostatic interactions are important for binding of Skp to its substrates and the presence of NaCl could therefore be disrupting binding rather than delivery to the membrane surface. To determine whether Skp is able to bind PagP in NaCl, the nickel affinity chromatography assay was repeated in the presence of 200 mM NaCl. As described previously, HT PagP (10 μM) was immobilised on nickel Sepharose resin and then incubated with an equal concentration of Skp trimers in 50 mM glycine, pH 9.5 containing 200 mM NaCl. The data show the co-elution of Skp with HT PagP (Figure 5.29), indicating that the addition of salt does not influence the binding of HT PagP to Skp. The results observed in the kinetic assays (Figure 5.28) suggest that NaCl disrupts electrostatic interactions between Skp and the target membrane, and hence retards the rate of PagP folding.

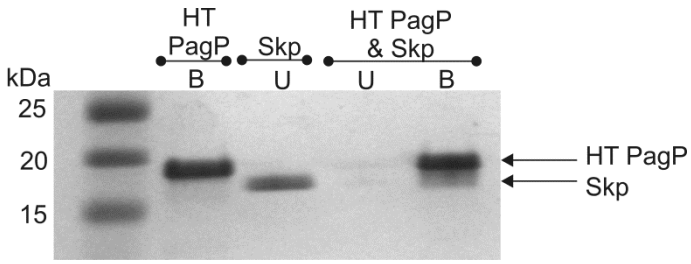

Figure 5.29 Skp binding to HT PagP is not disrupted by the presence of 200 mM NaCl. 10 μM HT PagP was immobilised on nickel Sepharose resin before incubation with 10 μM Skp in 50 mM glycine buffer, pH 9.5 containing 200 mM NaCl. Control experiments containing HT PagP or Skp only were conducted under identical conditions for comparison. Skp alone does not bind to the resin and is present in the unbound (U) fraction, while in the presence of HT PagP, Skp co-elutes in the bound (B) fraction.

5.2.6 Interaction With LPS-Containing Bilayers Has a Limited Effect on Skp-Mediated PagP Folding

Based on the results of cross-linking experiments, Skp has previously been suggested to interact with unfolded OMPs soon after their translocation into the periplasm[289]. However, under normal growth conditions, it has been hypothesised that Skp is not part of the major pathway of OMP assembly[107]. The crystal structure of Skp revealed a putative lipopolysaccharide (LPS) binding site on the surface of the protein[106] and this coupled with the observation that LPS could modulate the Skp-mediated refolding of OmpA[159; 161; 290] has led to the hypothesis that Skp may act to deliver OMPs to the OM under conditions of stress[107].

This hypothesis was tested by measuring the PagP folding rate constant in the presence and absence of a two-fold molar excess of Skp trimers following incorporation of 20 % (*w/w*) LPS into both 100 nm *di*$C_{12:0}$PC and 100 nm 80:20 *di*$C_{12:0}$PC:*di*$C_{12:0}$PG LUVs. Analysis of the resulting kinetic transients revealed that the folding rate constant of PagP into LPS-containing *di*$C_{12:0}$PC LUVs was doubled in the presence of Skp ($k_{obs} = 8.9\times10^{-4} \pm 6.7\times10^{-5}$ s^{-1} for PagP alone and $k_{obs} = 1.9\times10^{-3} \pm 3.3\times10^{-4}$ s^{-1} with Skp) (Figure 5.30a, Figure 5.31, Table 5.5). The folding rate constant in LPS-containing 80:20 *di*$C_{12:0}$PC:*di*$C_{12:0}$PG LUVs was increased by a smaller amount (approximately 50 %) in the presence of Skp ($k_{obs} = 6.8\times10^{-4} \pm 7.1\times10^{-6}$ s^{-1} for PagP alone and $k_{obs} = 1.0\times10^{-3} \pm 1.5\times10^{-4}$ s^{-1} with Skp) (Figure 5.30b, Figure 5.31, Table 5.5).

These data show that LPS, a negatively charged glycolipid, causes an increase in rate of Skp-mediated PagP folding into bilayers which are not already negatively charged, however, this increase is smaller than the increase observed when Skp is added to the folding reaction in negatively charged liposomes in the absence of LPS. Together, these results suggest that the presence of LPS may actually impede delivery of PagP to the membrane surface in the presence of Skp.

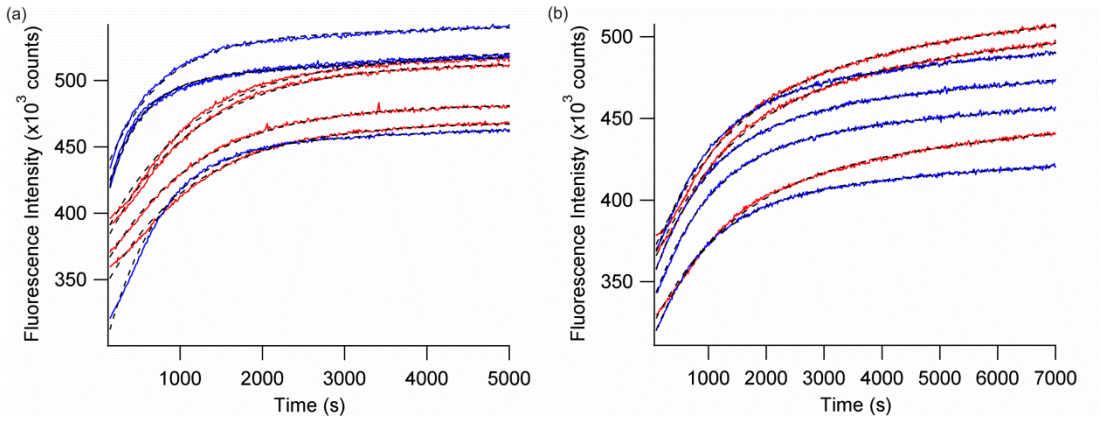

Figure 5.30 Effect of LPS on Skp-mediated refolding of PagP. Kinetic traces of PagP folding into (a) *di*$C_{12:0}$PC liposomes containing 20 % (*w/w*) LPS (b) 80:20 *di*$C_{12:0}$PC:*di*$C_{12:0}$PG liposomes containing 20 % (*w/w*) LPS in the presence (blue lines) or absence (red lines) of a two-fold molar excess of Skp trimers. Dashed black lines represent the fits of the data to a single exponential function. The differences in the final fluorescence signal arise from small PagP concentration differences between samples, which appear due to the large molar extinction coefficient (82390 M^{-1} cm^{-1}) of PagP. All samples contained 0.4 µM PagP and 2 M urea in 50 mM glycine, pH 9.5, had an LPR of 3200:1 and were measured at 37 °C.

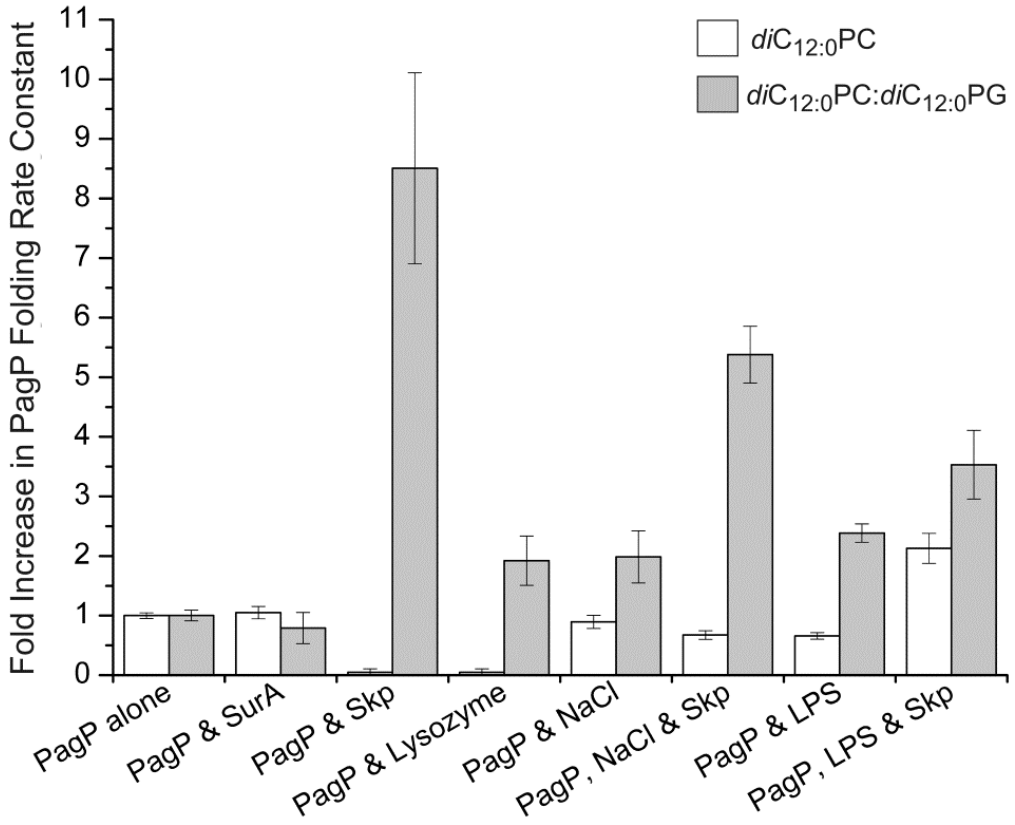

Figure 5.31 Relative rates of PagP folding under different conditions in *di*$C_{12:0}$PC (white) and 80:20 *di*$C_{12:0}$PC:*di*$C_{12:0}$PG (grey) liposomes. Each rate is normalised to the rate of PagP folding alone in the same lipid. Error bars depict the standard deviation of the average rate for each condition, propagated through the normalisation. All kinetic samples contained 0.4 μM PagP, LPR of 3200:1, 50 mM glycine, pH 9.5 and were measured at 37 °C. Unnormalised rate constants and their associated errors are provided in Table 5.5.

Lipid	Experiment	Average Rate Constant ($\times 10^{-3}$ s^{-1})	Standard Deviation ($\times 10^{-3}$ s^{-1})	Standard Error of the Mean ($\times 10^{-3}$ s^{-1})
di$C_{12:0}$PC	PagP alone (2 M urea)	1.34	0.08	0.05
	PagP & SurA (2 M urea)	1.41	0.23	0.13
	PagP & Skp (2 M urea)	0	N/A	N/A
	PagP & Lysozyme (2 M urea)	0	N/A	N/A
	PagP & NaCl (2 M urea)	1.20	0.25	0.14
	PagP, NaCl & Skp (2 M urea)	0.81	0.14	0.10
	PagP & LPS (2 M urea)	0.89	0.12	0.07
	PagP, LPS & Skp (2 M urea)	1.89	0.58	0.33
80:20 *di*$C_{12:0}$PC: *di*$C_{12:0}$PG	PagP alone (3 M urea)	0.79	0.20	0.14
	PagP alone (2 M urea)	0.28	0.03	0.02
	PagP & SurA (3 M urea)	0.63	0.25	0.17
	PagP & Skp (2 M urea)	2.41	0.60	0.43
	PagP & Lysozyme (3 M urea)	1.52	0.26	0.18
	PagP & NaCl (2 M urea)	0.56	0.11	0.06
	PagP, NaCl & Skp (2 M urea)	1.53	0.16	0.09
	PagP & LPS (2 M urea)	0.68	0.01	0.01
	PagP, LPS & Skp (2 M urea)	1.00	0.26	0.15

Table 5.5 Measured rate constants of PagP folding into liposomes *in vitro*. Rate constants were obtained by measuring changes in tryptophan fluorescence emission at 335 nm (following excitation at 280 nm) over time, and fitting the observed transients to a single exponential function. Global fits were obtained over four replicates from a single batch of liposomes, and the average of the global fits from folding reactions into three batches of liposomes calculated. The standard error of the mean was calculated by taking the number of liposome replicates to be 3.

5.2.7 Skp Displays Holdase Activity Against a Highly Aggregation-Prone PagP Construct

The data described herein and elsewhere[174] show that Skp binds to unfolded or partially folded PagP species. Although this interaction is inhibitory to folding and membrane insertion in *di*$C_{12:0}$PC LUVs, sequestration of PagP by Skp should decrease the local concentration of unfolded and partially folded PagP and hence decrease the aggregation propensity of these species. To probe this hypothesis, the ability of both Skp and lysozyme to reduce aggregation of HT PagP was tested. This construct was chosen as it

is particularly aggregation prone, having been shown to require high concentrations of urea (> 4 M) to remain soluble[231; 238]. HT PagP was diluted to 1 M urea in 50 mM glycine, pH 9.5 and incubated in these conditions for 5 min before dilution into 100 nm *di*$C_{12:0}$PC LUVs. Under these conditions, all of the HT PagP precipitated, leaving none detectable in the supernatant by either SDS-PAGE or Western blotting with an anti-Histag antibody (Figure 5.32). Conversely, when HT PagP was diluted to 1 M urea in the presence of a two-fold molar excess of Skp trimers, a significant proportion of the HT PagP was rescued from aggregation and retained in a folding-competent state as evidenced by the heat modifiability of the HT PagP band, a characteristic of folded OMPs[140] (Figure 5.32). Using densitometry, the amount of HT PagP which was aided in folding and membrane insertion by Skp was estimated to be 14 %. When the assay was repeated in the presence of a six-fold molar excess of lysozyme, none of the HT PagP was prevented from aggregation (Figure 5.32). These data suggest that Skp is able to sequester HT PagP, preventing its aggregation even under conditions which strongly disfavour folding, and then release the protein in a folding-competent state. By contrast, lysozyme was not able to prevent the aggregation of HT PagP under these conditions despite being shown to interact with PagP under conditions which favour folding (Figure 5.27). These data support the conclusions of other studies which have suggested that Skp acts as a holdase chaperone[102; 104; 108; 161], preventing the aggregation of unfolded OMPs even under conditions which strongly favour aggregation.

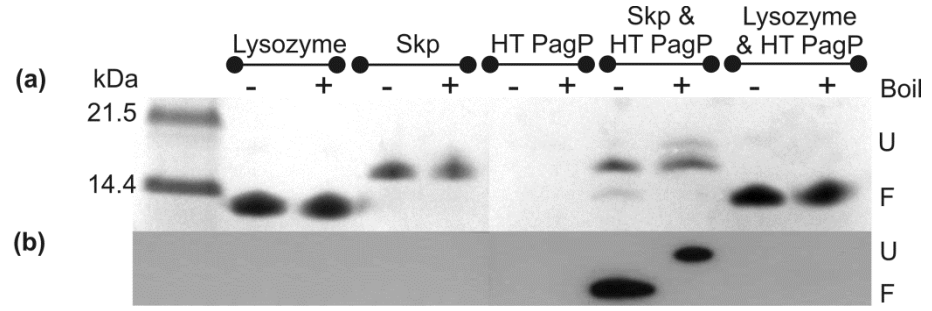

Figure 5.32 The holdase activity of Skp against His-tagged PagP (HT PagP), a highly aggregation-prone construct, demonstrated by (a) SDS-PAGE and (b) Western blot analysis using an anti-Histag antibody. 21.6 µM HT PagP in 1 M urea, 50 mM glycine, pH 9.5 was incubated in the presence of a two-fold molar excess of Skp trimers or a six-fold molar excess of hen egg white lysozyme before a six-fold dilution into *di*$C_{12:0}$PC liposomes (final HT PagP concentration = 3.6 µM; LPR 3200:1) at 37 °C. Any precipitate was removed by centrifugation and the supernatant analysed for the presence of HT PagP. The folded and unfolded forms of HT PagP are denoted by F and U, respectively. Control experiments containing Skp and lysozyme incubated alone were also conducted under identical conditions. Lysozyme runs at the same apparent molecular weight as folded HT PagP. The ability of Skp to divert HT PagP from aggregation was quantified by densitometry measurements.

5.3 DISCUSSION

Biophysical analysis of the folding pathway of PagP *in vitro* has been carried out previously, utilising a highly aggregation prone C-terminally His-tagged construct (HT PagP) which requires high concentrations of urea to remain in a soluble and folding-competent conformation[167; 238; 240]. The requirement for high concentrations of denaturant in the folding buffer excludes the possibility of investigating the effects of soluble, periplasmic chaperones on the folding mechanism of PagP. However, an untagged PagP construct reported by Burgess *et al* was recently shown to fold in a variety of lipid and buffer conditions in as little as 1 M urea[162]. In this study, a real-time spectroscopic assay has been developed to directly monitor the folding of the untagged PagP construct into liposomes in the presence and absence of periplasmic folding factors.

Inclusion of SurA in the PagP folding assay was shown to have no effect on the rate constant or yield of PagP folding into either zwitterionic *di*$C_{12:0}$PC or negatively charged 80:20 *di*$C_{12:0}$PC:*di*$C_{12:0}$PG liposomes. Furthermore, SurA was not observed to interact with PagP under the conditions tested, suggesting that SurA does not play a

direct role in assisting PagP folding *in vitro*. In seeming contrast with these observations, the work of Sklar *et al*[94] shows that depletion of SurA led to a loss of OM density and defects in OMP assembly. Additionally, differential proteomics experiments show that depletion of SurA causes a reduction in the abundance of some OMPs in *E. coli*[85], an effect exacerbated when Skp is also depleted[98]. Together, these results have led to the hypothesis that SurA is the primary chaperone involved in OMP transport and assembly in the periplasm[85; 94; 98]. In accordance with this view, SurA is the only soluble chaperone which has successfully been cross-linked to BamA *in vivo*[94; 291] and has been shown to increase the rate of BAM-mediated OmpT refolding *in vitro*[137] suggesting a role for SurA in delivery of unfolded OMPs to the BAM complex. However, while SurA has been shown to bind to OMP-derived peptides with affinities in the μM range[89] and a preference for Ar-X-Ar motifs[89; 90], few studies report binding of SurA to full length OMPs *in vitro*[92; 292]. Taken together with the data presented herein, these results suggest that SurA does not influence OMP folding *in vitro* in the absence of the BAM complex and as such its role may be in delivery of unfolded OMPs to the BAM complex to initiate folding and membrane insertion.

Inclusion of Skp in the PagP folding assay demonstrated that this chaperone results in striking changes in the rates of PagP folding, which are dependent on both bilayer charge and the ionic strength of the folding buffer. PagP folding into zwitterionic *di*$C_{12:0}$PC liposomes was strongly inhibited by Skp such that no folding was observed over a 2 h time-course. When the folding time-course was increased to 16 h, however, the yield of folded PagP at the end of the reaction was unchanged by the presence of Skp, as judged by cold SDS-PAGE. In stark contrast with these observations, folding of PagP into 80:20 *di*$C_{12:0}$PC:*di*$C_{12:0}$PG liposomes was accelerated almost ten-fold in the presence of Skp, suggesting that electrostatic interactions with the head-groups of the lipids in the bilayer may have an important role to play in Skp-mediated membrane delivery. Accordingly, the crystal structure of Skp reveals that the trimer has a "jellyfish" topology with long α-helical "tentacles" which define a central cavity and a notable positive charge distribution at the tips of the helices, resulting in an extremely large dipole moment (Figure 1.12)[105; 106] through which interaction with a negatively charged membrane would be favourable. The kinetic data presented herein are in agreement with data concerning the Skp-mediated folding of OmpA[161], which was monitored using cold SDS-PAGE methods and showed that Skp inhibits folding of

OmpA into zwitterionic 1,2-dioleoyl-sn-glycero-3-phosphocholine (DOPC) liposomes and reduces the folded yield over a 4 h time-course. In the same study, folding of OmpA into negatively charged liposomes at neutral or basic pH was shown to be facilitated by Skp[161].

The role of electrostatic interactions in Skp function was verified by the inclusion of NaCl in the folding assays, which partially reversed the effects of Skp in both lipid mixtures without disrupting substrate binding. Furthermore, the inclusion of negatively charged LPS was able to increase the rate of Skp-mediated refolding by a greater extent in zwitterionic liposomes compared with negatively charged liposomes. The ability of lysozyme to partially mimic the action of Skp provides further support for this idea, as it is also a highly basic protein. However, lysozyme was unable to achieve the dramatic acceleration of PagP folding observed in the presence of Skp, suggesting that other factors such as hydrophobic interactions and substrate orientation may play roles in Skp activity. Indeed, a recent study on the interaction of Skp with soluble proteins concluded that hydrophobic interactions were important for binding to these substrates[108]. While lysozyme was observed to bind to PagP and alter folding rate constants, it was unable to prevent the aggregation of HT PagP, in contrast with Skp. This difference may arise from the difference in structures of Skp and lysozyme. Skp has been shown to be able to bind to and sequester unfolded OMPs in its central cavity, a feature absent in lysozyme[293], by both fluorescence[101; 104; 292] and NMR spectroscopy[92; 102]. It could be assumed that interaction of Skp with the membrane surface would bring the client protein, bound within the cavity, in close to the membrane surface while excluding the solvent, hence promoting folding and membrane insertion.

The presence of a putative LPS binding site identified in the Skp structure[106] has contributed to the hypothesis that Skp may act in times of stress to recover OMPs which fall off the main folding pathway[107]. LPS is normally only present in the outer leaflet of the OM, and would not come into contact with Skp under normal growth conditions[64; 65; 107]. Previous analysis of the effect of LPS on the Skp-mediated refolding of OmpA[159; 161] reported that LPS increased the folding rate into liposomes irrespective of membrane charge, in contrast with the results presented herein. One possible reason for these seemingly contradictory results is that while the kinetic assays described in this study used LPS-containing bilayers, the work on OmpA allowed pre-binding of the

Skp:OMP complex to LPS before addition to lipid[159]. LPS has been shown to be able to spontaneously insert into bilayers from aqueous solution[294], suggesting that even a non-specific electrostatic interaction between Skp and LPS may bring the Skp:OMP complex in contact with the membrane during LPS insertion and trigger faster OMP folding. While it is not known which of these experimental designs closer mimics the possible LPS-Skp interaction under stress conditions *in vivo*, it is reasonable to assume that LPS biogenesis would also be affected[107] and result in LPS accumulation along the transport pathway[65]. Regardless of whether Skp encountered LPS contained within bilayers or free in the periplasm, it could be assumed that this Skp:LPS interaction may result in non-specific delivery of OMPs to either membrane and its importance *in vivo* remains to be elucidated.

The current prevailing model of OMP biogenesis suggests that SurA is the primary chaperone for OMP transport, with Skp acting on a separate minor pathway in cooperation with the dual protease-chaperone, DegP[107; 111]. This model has arisen from genetic[94; 95] and proteomic[85; 98] data showing that depletion of Skp alone has little effect on the composition of OMPs in the OM, while SurA depletion and simultaneous SurA and Skp depletion cause increasingly serious OMP assembly defects. While this model appears to fit the data, it is based on the assumption that only the three chaperones SurA, Skp and DegP play a notable role in OMP assembly. Very recently, it has been reported that depletion of both Skp and another periplasmic chaperone FkpA[110] causes impairment of the assembly of the essential OMP LptD[100]. Overexpression of SurA failed to compensate for this defect in the OMP assembly pathway, while overexpression of either Skp or FkpA failed to compensate for SurA depletion[100]. These new data are not easily explained by the current OMP folding model and add weight to the previously disfavoured hypothesis that Skp interacts with newly-translocated OMPs and passes them to other chaperones to initiate folding[295]. It has been previously observed using MD simulation data that Skp binding is initiated by the OMP N-terminal residues[104] and that the OMP PhoE interacts with Skp soon after translocation in a manner that is impeded by removal of N-terminal segments[103]. Overall this suggests that Skp and FkpA may serve the redundant function of interacting with unfolded OMPs as they emerge from the translocon, preventing their aggregation. Following completion of translocation, the unfolded OMP may then interact with other chaperones which are involved in promoting folding and membrane insertion, such as

SurA and the BAM complex. This model would provide an explanation for the difficulty in binding SurA to full-length OMPs *in vitro* as shown by the data presented in this study and the few studies which report binding elsewhere[92; 292], as well as the lack of SurA participation in OMP folding *in vitro* in the absence of other chaperones.

While it remains unclear which of the proposed models accurately describes the chaperone pathways encountered by unfolded OMPs in the *E. coli* periplasm, the data presented herein demonstrate the utility of simplified *in vitro* systems in dissecting this complex folding environment. This work paves the way for studies of PagP folding in the presence of other periplasmic chaperones and BAM complex components in future, allowing further insights to be gained about the interactions made by unfolded OMPs during their transit to the OM. Additionally, the ability of Skp to sequester an extremely aggregation prone OMP and promote its folding and membrane insertion *in vitro* is shown, suggesting a potential role for this chaperone in holding unfolded OMPs prior to their delivery, either directly or indirectly, to the membrane *in vivo*.

6 CONCLUDING REMARKS AND FUTURE DIRECTIONS

Since the realisation by Anfinsen that the primary amino acid sequence encodes the native, three-dimensional structure of a protein [4], the field of protein folding has seen phenomenal progress. Over the past fifty years, increases in computing power coupled with improvements in experimental techniques have led to elucidation of the folding pathways of small, soluble proteins at near-atomistic detail[17]. In stark contrast, however, elucidation of the folding mechanisms of membrane proteins has lagged significantly behind[27]. Development of suitable membrane mimetic systems that allow the folding of membrane proteins to be studied *in vitro* has resulted in more research being focussed in this area in recent years[27; 29; 46; 48; 296].

Integral membrane proteins with a β-barrel structure are found in mitochondria, chloroplasts and the OM of Gram-negative bacteria. Those of the bacterial OM carry out a diverse range of essential functions including (non)specific transport of small and large ligands[46; 64; 282], proteolytic and synthetic reactions[46; 282] and cellular recognition and adhesion[64; 282]. Accordingly, while the β-barrel transmembrane domain forms a common structural motif, the number, length and tilt angle of the β-strands, as well as the presence of non-membrane incorporated domains is highly variable[46; 48]. To date, detailed information on the folding pathways of only two OMPs, HT PagP[167; 238; 240] and OmpA[144; 145; 147; 149; 153; 161], has been elucidated, making it difficult to draw generic conclusions about the folding of this diverse family of proteins. In the field of soluble protein folding, the "fold approach" of studying the folding mechanisms of homologous proteins has been very successful in determining features of protein folding landscapes[175; 177]. This approach is now finding utility in the field of OMP folding, with the Fleming group studying the folding of nine different OMPs simultaneously, allowing the behaviour of these proteins to be compared and contrasted[162; 164; 165; 174]. While informative, these studies have not yet been able to achieve the level of detail acquired in comparative studies of soluble protein folding (Sections 1.1 and 1.7). The work described herein aims to expand current knowledge on OMP folding by striking a balance between the high levels of detail attained on the folding pathway of HT PagP and the utility of comparative folding studies by focussing on a small subset of OMPs.

In Chapter 3 the homologous *E. coli* OM proteases, OmpT and OmpP, were cloned from bacterial DNA and protocols for their over-expression and purification from inclusion bodies were developed. Refolding to the native state was achieved for both proteins into synthetics PC liposomes of both different hydrophobic thickness and different acyl chain saturation, demonstrating the structural malleability of both these OMPs and the surrounding lipid environment. The ability of OmpT and OmpP to successfully fold in different lipid environments without significantly perturbing the final protein structure suggests that local rearrangements of amino acid side chains, local membrane deformation or a combination of both must be occurring. These effects have previously been documented for an α-helical membrane protein using a combination of crystallography and MD simulations which studied the conformational changes of the sarco(endo)plasmic reticulum Ca^{2+}–ATPase[268]. Additionally, the effect of specific side-chains on the positioning of transmembrane helices has been well documented[156; 269; 270; 271; 272]. In contrast, the role of individual side chains in membrane positioning and OMP stability is much less characterised. Work to date on OmpA[155; 273] and OmpLa[297] has begun to examine the role of the energetic cost of burying amino acid side chains in the membrane interior and the energetic contributions of aromatic side chains in the intrafacial regions of the bilayer. So far this work has not uncovered any evidence for the repositioning of transmembrane regions due to side chain-lipid interactions in OMPs, and this may be attributable to the rigidity of the β-barrel structure. Assuming this to be the case, local membrane deformations are likely to be mostly responsible for the ability of OmpT and OmpP to remain natively folded and functional in lipid environments of varying hydrophobic thickness.

Examination of the relative stabilities of OmpT and OmpP reveal that in liposomes composed of lipids with saturated acyl chains, both proteins appear similarly stable. In contrast, when lipids with unsaturated acyl chains are present, leading to an increase in lateral pressure within the membrane, OmpP is much more destabilised than OmpT. These two OMPs have approximately 87 % sequence homology[244] and the homology model of OmpP based on the crystal structure of OmpT (Figure 3.2) reveals that residues which are not conserved in the two proteins are clustered in two areas: around the active site and the aromatic girdles. The differences in the active site are presumably necessary to alter the substrate specificity of the enzyme, while those in the aromatic girdles, in particular the three extra aromatic residues at the periplasmic

intrafacial region of OmpP, may act to increase the stability of the protein. Why, then, is OmpP less stable in liposomes comprised of lipids with unsaturated acyl chains if it has extra aromatic clusters in the intrafacial region? This presumably arises as the sequence required to confer the substrate specificity of OmpP is very destabilising and these extra residues are required to partly offset this destabilisation and ensure a stable fold is still attainable. This fold versus function relationship is well documented in soluble proteins[201; 277; 278]; however, this is the first reported example of evidence for its existence in OMPs.

The large destabilisation of OmpP in *di*$C_{16:1}$PC liposomes makes this OMP unsuitable for rigorous thermodynamic analysis as a full unfolding transition could not be measured. OmpT, however, was less destabilised and accordingly it was attempted to find conditions which promoted folding reversibility of this protein. Even after systematically testing a variety of conditions, fully reversible folding of OmpT without hysteresis between the folding and unfolding urea denaturation curves was not achieved. This phenomenon has been widely reported in the literature for different OMPs[150; 151; 165; 174]. Kinetic analysis of OmpT, PagP and HT PagP revealed a link between folding reversibility and unfolding kinetics wherein an extra exponential phase is observed in the unfolding of OMPs which fold irreversibly. It was hypothesised, based on previous work on the folding mechanisms of HT PagP[167; 240], that this additional phase may arise due to lipid detachment of the unfolded ensemble leading to aggregation. Evidence from the FRET studies presented in Chapter 4 support this view. Recent work by the Fleming group on OmpW, OmpLa and PagP at pH 3.8 suggested that the unfolded states of these proteins under conditions which promote reversible folding without hysteresis is completed solvated and cannot therefore be membrane-associated[174]. It should be noted that neither the tryptophan fluorescence emission technique used by the Fleming group[174] nor the FRET studies described in Chapter 4 provide direct evidence for the membrane-association or dissociation of these OMPs and as such it is difficult to draw firm conclusions on the behaviour of the unfolded ensemble in reversible folding.

Following the kinetic analysis of the unfolding of OmpT, PagP and HT PagP, refolding kinetics of these proteins, as well as those of OmpP, were next examined in low urea. Conditions were established under which PagP could refold reproducibly with good yield in 2–4 M urea, and these were used to facilitate the study of folding in the presence of the periplasmic chaperones, Skp and SurA. The unexpectedly slow folding

of PagP in 2 M urea compared with that in 4 M urea was attributed to a possible increase in PagP aggregation at lower urea concentrations, although the protein concentration independence of the rate constants do not fit well with this hypothesis. Studying refolding kinetics at higher PagP concentrations than those addressed in this work may show protein concentration dependence of the folding rate constants as expected in an aggregating system and should be examined. While work on SurA *in vivo* has suggested this is the main chaperone for OMP biogenesis[85; 94; 98; 291], the work presented herein alongside a lack of data implicating binding of SurA to full-length OMPs[89; 90; 92; 292] suggests that the situation *in vivo* may not be so clear cut as previously thought. Indeed, the complexity of the chaperone network in the periplasm[76; 83] makes the results of *in vivo* knock-out experiments difficult to interpret with confidence and highlights the importance for minimalistic *in vitro* assays to understand OMP:chaperone interactions.

The ability of Skp to sequester PagP and modulate PagP refolding rate in a manner dependent on electrostatic interactions is in accordance with previous data on the folding of OmpA in the presence of Skp[101; 102; 159; 161]. These data allow generic conclusions about the mechanism of action of this chaperone to be drawn, but additionally the use of a real-time spectroscopic assay allows a much higher time resolution of the refolding experiments than has been achieved previously using cold SDS-PAGE[159; 161]. A higher time-resolution not only allows more accurate determination of folding rates but also facilitates analysis of kinetic transients where multiple exponential phases may be occurring, as these processes can be difficult to identify unambiguously, especially in limited data sets. The development of this assay allows the effects of individual chaperones on OMP folding to be examined and paves the way for experiments to investigate the interactions between chaperones at different phases in the OMP biogenesis pathway. Interestingly, the results presented suggested that the presence of a holdase chaperone did not increase PagP folding efficiency over a 16 h time-course as might be expected. To address this, further work is needed to establish if this reaction has reached completion in the time-course of the experiment and if longer incubation times would yield an increase in folding efficiency. Work on this project is on-going the laboratory and aims to dissect the complex mechanisms by which OMPs traverse the periplasm *in vivo*. Additionally, the opportunity to compare the periplasmic chaperone network, which operates in an ATP-devoid environment, to

that of the cytosolic chaperone network could provide valuable insights into the similarities and differences between the folding of soluble and insoluble proteins *in vivo*.

Overall, the work in this thesis provides insight into the folding mechanisms of several OMPs in both the presence and absence of periplasmic chaperones. What is apparent is that the interaction of unfolded OMPs with the membrane environment plays a critical role in folding irrespective of whether the OMP is solubilised by a chaperone or chaotrope. Characterising and understanding these protein:lipid interactions *in vitro* will be critical for further progress in this field. Without knowledge of the role of the membrane in modulating the folding efficiency, aggregation propensity and stability of OMPs we cannot hope to fully understand the folding mechanisms of membrane proteins. It is apparent that there is much still to be learned about the folding of membrane proteins but the increasing rate of progress in recent years suggests that the gap in understanding between soluble proteins and membrane proteins is growing ever smaller. Closing this gap will allow detailed comparisons between the folding of soluble and insoluble proteins to be drawn, providing insights which may be invaluable in addressing one of the fundamental questions in biochemistry: just how does a protein's sequence confer its native structure?

REFERENCES

1. Thomas, P. J., Qu, B. H. & Pedersen, P. L. (1995). Defective protein-folding as a basis of human disease. *Trends Biochem. Sci.* **20**, 456-459.

2. Naeem, A. & Fazili, N. (2011). Defective protein folding and aggregation as the basis of neurodegenerative diseases: the darker aspect of proteins. *Cell Biochem. Biophys.* **61**, 237-250.

3. Vendruscolo, M., Knowles, T. P. J. & Dobson, C. M. (2011). Protein solubility and protein homeostasis: a generic view of protein misfolding disorders. *Cold Spring Harb. Perspect. Biol.* **3**, a010454.

4. Anfinsen, C. B., Haber, E., Sela, M. & White, F. H. (1961). Kinetics of formation of native ribonuclease during oxidation of reduced polypeptide chain. *Proc. Natl. Acad. Sci. U.S.A.* **47**, 1309-1314.

5. Anfinsen, C. B. (1973). Principles that govern the folding of protein chains. *Science* **181**, 223-230.

6. Fersht, A. R. (2008). From the first protein structures to our current knowledge of protein folding: delights and scepticisms. *Nat. Rev. Mol. Cell Biol.* **9**, 650-654.

7. Levinthal, C. (1969). Mössbauer spectroscopy in biological systems: proceedings of a meeting held at Allerton House, Monticello, Illinois. *University of Illinois bulletin* **67**, 22-24.

8. Levinthal, C. (1968). Are there pathways for protein folding? *J. Chim. Phys.* **65**, 44-45.

9. Tsong, T. Y., Baldwin, R. L. & Elson, E. L. (1972). Properties of refolding and unfolding reactions of Ribonuclease A. *Proc. Natl. Acad. Sci. U.S.A.* **69**, 1809-12.

10. Fersht, A. R. (1997). Nucleation mechanisms in protein folding. *Curr. Opin. Struct. Biol.* **7**, 3-9.

11. Brandts, J. F., Halvorson, H. R. & Brennan, M. (1975). Consideration of possibility that slow step in protein dentauration reactions is due to *cis-trans* isomerism of proline residues. *Biochemistry* **14**, 4953-4963.

12. Itzhaki, L. S., Otzen, D. E. & Fersht, A. R. (1995). The structure of the transition state for folding of Chymotrypsin Inhibitor-2 analysed by protein engineering methods - evidence for a nucleation-condensation mechanism for protein-folding. *J. Mol. Biol.* **254**, 260-288.

13. Gianni, S., Guydosh, N. R., Khan, F., Caldas, T. D., Mayor, U., White, G. W. N., DeMarco, M. L., Daggett, V. & Fersht, A. R. (2003). Unifying features in protein-folding mechanisms. *Proc. Natl. Acad. Sci. U.S.A.* **100**, 13286-13291.

14. Onuchic, J. N. & Wolynes, P. G. (2004). Theory of protein folding. *Curr. Opin. Struct. Biol.* **14**, 70-75.

15. Smith, L. J., Fiebig, K. M., Schwalbe, H. & Dobson, C. M. (1996). The concept of a random coil: residual structure in peptides and denatured proteins. *Fold. Des.* **1**, R95-R106.

16. Pashley, C. L., Morgan, G. J., Kalverda, A. P., Thompson, G. S., Kleanthous, C. & Radford, S. E. (2012). Conformational properties of the unfolded state of Im7 in nondenaturing conditions. *J. Mol. Biol.* **416**, 300-318.

17. Bartlett, A. I. & Radford, S. E. (2009). An expanding arsenal of experimental methods yields an explosion of insights into protein folding mechanisms. *Nat. Struct. Mol. Biol.* **16**, 582-588.

18. Brockwell, D. J. & Radford, S. E. (2007). Intermediates: ubiquitous species on folding energy landscapes? *Curr. Opin. Struct. Biol.* **17**, 30-37.

19. Bryngelson, J. D., Onuchic, J. N., Socci, N. D. & Wolynes, P. G. (1995). Funnels, pathways, and the energy landscape of protein folding: a synthesis. *Proteins* **21**, 167-95.

20. Ferreiro, D. U., Hegler, J. A., Komives, E. A. & Wolynes, P. G. (2007). Localizing frustration in native proteins and protein assemblies. *Proc. Natl. Acad. Sci. U.S.A.* **104**, 19819-19824.

21. Wensley, B. G., Batey, S., Bone, F. A. C., Chan, Z. M., Tumelty, N. R., Steward, A., Kwa, L. G., Borgia, A. & Clarke, J. (2010). Experimental evidence for a frustrated energy landscape in a three-helix-bundle protein family. *Nature* **463**, 685-688.

22. Neuweiler, H., Doose, S. & Sauer, M. (2005). A microscopic view of miniprotein folding: enhanced folding efficiency through formation of an intermediate. *Proc. Natl. Acad. Sci. U.S.A.* **102**, 16650-16655.

23. Tsytlonok, M. & Itzhaki, L. S. (2013). The how's and why's of protein folding intermediates. *Arch. Biochem. Biophys.* **531**, 14-23.

24. Kubelka, J., Hofrichter, J. & Eaton, W. A. (2004). The protein folding 'speed limit'. *Curr. Opin. Struct. Biol.* **14**, 76-88.

25. Schuler, B. & Eaton, W. A. (2008). Protein folding studied by single-molecule FRET. *Curr. Opin. Struct. Biol.* **18**, 16-26.

26. Lindorff-Larsen, K., Piana, S., Dror, R. O. & Shaw, D. E. (2011). How fast-folding proteins fold. *Science* **334**, 517-520.

27. Booth, P. J. & Curnow, P. (2009). Folding scene investigation: membrane proteins. *Curr. Opin. Struct. Biol.* **19**, 8-13.

28. Hartl, F. U. & Hayer-Hartl, M. (2009). Converging concepts of protein folding *in vitro* and *in vivo*. *Nat. Struct. Mol. Biol.* **16**, 574-581.

29. Bowie, J. U. (2005). Solving the membrane protein folding problem. *Nature* **438**, 581-589.

30. Hopkins, A. L. & Groom, C. R. (2002). The druggable genome. *Nat. Rev. Drug Discov.* **1**, 727-730.

31. Nam, H. J., Jeon, J. & Kim, S. (2009). Bioinformatic approaches for the structure and function of membrane proteins. *Bmb Reports* **42**, 697-704.

32. Bos, M. P. & Tommassen, J. (2004). Biogenesis of the Gram-negative bacterial outer membrane. *Curr. Opin. Microbiol.* **7**, 610-616.

33. Nguyen, U. T. T., Goody, R. S. & Alexandrov, K. (2010). Understanding and exploiting protein prenyltransferases. *ChemBioChem* **11**, 1194-1201.

34. Spiro, R. G. (2002). Protein glycosylation: nature, distribution, enzymatic formation, and disease implications of glycopeptide bonds. *Glycobiology* **12**, 43R-56R.

35. Kennedy, S. J. (1978). Structures of membrane proteins. *J. Memb. Biol.* **42**, 265-279.

36. White, S. H., Ladokhin, A. S., Jayasinghe, S. & Hristova, K. (2001). How membranes shape protein structure. *J. Biol. Chem.* **276**, 32395-32398.

37. Henderson, R. & Unwin, P. N. T. (1975). Three-dimensional model of purple membrane obtained by electron microscopy. *Nature* **257**, 28-32.

38. Schobert, B., Cupp-Vickery, J., Hornak, V., Smith, S. & Lanyi, J. (2002). Crystallographic structure of the K intermediate of bacteriorhodopsin: conservation of free energy after photoisomerization of the retinal. *J. Mol. Biol.* **321**, 715-26.

39. Raman, P., Cherezov, V. & Caffrey, M. (2006). The Membrane Protein Data Bank. *Cell. Mol. Life Sci.* **63**, 36-51.

40. MacKenzie, K. R., Prestegard, J. H. & Engelman, D. M. (1997). A transmembrane helix dimer: structure and implications. *Science* **276**, 131-133.

41. Luecke, H., Schobert, B., Richter, H.-T., Cartailler, J.-P. & Lanyi, J. K. (1999). Structure of bacteriorhodopsin at 1.55 Å resolution. *J. Mol. Biol.* **291**, 899-911.

42. Toyoshima, C. & Nomura, H. (2002). Structural changes in the calcium pump accompanying the dissociation of calcium. *Nature* **418**, 605-611.

43. Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N. & Bourne, P. E. (2000). The Protein Data Bank. *Nucleic Acids Res.* **28**, 235-242.

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

45. Wimley, W. C. (2003). The versatile beta-barrel membrane protein. *Curr. Opin. Struct. Biol.* **13**, 404-411.

46. Fairman, J. W., Noinaj, N. & Buchanan, S. K. (2011). The structural biology of β-barrel membrane proteins: A summary of recent reports. *Curr. Opin. Struct. Biol.* **21**, 523-531.

47. Haltia, T. & Freire, E. (1995). Forces and factors that contribute to the structural stability of membrane proteins. *Biochim. Biophys. Acta* **1228**, 1-27.

48. Otzen, D. E. & Andersen, K. K. (2013). Folding of outer membrane proteins. *Arch. Biochem. Biophys.* **531**, 34-43.

49. Ahn, V. E., Lo, E. I., Engel, C. K., Chen, L., Hwang, P. M., Kay, L. E., Bishop, R. E. & Prive, G. G. (2004). A hydrocarbon ruler measures palmitate in the enzymatic acylation of endotoxin. *EMBO J.* **23**, 2931-2941.

50. Pautsch, A. & Schulz, G. E. (1998). Structure of the outer membrane protein A transmembrane domain. *Nat. Struct. Biol.* **5**, 1013-1017.

51. Vandeputte-Rutten, L., Kramer, R. A., Kroon, J., Dekker, N., Egmond, M. R. & Gros, P. (2001). Crystal structure of the outer membrane protease OmpT from *Escherichia coli* suggests a novel catalytic site. *EMBO J.* **20**, 5033-5039.

52. Snijder, H. J., Ubarretxena-Belandia, I., Blaauw, M., Kalk, K. H., Verheij, H. M., Egmond, M. R., Dekker, N. & Dijkstra, B. W. (1999). Structural evidence for dimerization-regulated activation of an integral membrane phospholipase. *Nature* **401**, 717-721.

53. Bayrhuber, M., Meins, T., Habeck, M., Becker, S., Giller, K., Villinger, S., Vonrhein, C., Griesinger, C., Zweckstetter, M. & Zeth, K. (2008). Structure of the human voltage-dependent anion channel. *Proc. Natl. Acad. Sci. U.S.A.* **105**, 15370-15375.

54. Remaut, H., Tang, C., Henderson, N. S., Pinkner, J. S., Wang, T., Hultgren, S. J., Thanassi, D. G., Waksman, G. & Li, H. (2008). Fiber formation across the bacterial outer membrane by the chaperone/usher pathway. *Cell* **133**, 640-652.

55. White, S. H. http://blanco.biomol.uci.edu/mpstruc/listAll/list. *http://blanco.biomol.uci.edu/mpstruc/*, Accessed 7th October, 2013.

56. Warschawski, D. E. http://www.drorlist.com/nmr/MPNMR.html. *http://www.drorlist.com/nmr/MPNMR.html*, Accessed 7th October 2013.

57. Maslennikov, I. & Choe, S. (2013). Advances in NMR structures of integral membrane proteins. *Curr. Opin. Struct. Biol.* **23**, 555-562.

58. van Meer, G., Voelker, D. R. & Feigenson, G. W. (2008). Membrane lipids: where they are and how they behave. *Nat. Rev. Mol. Cell Biol.* **9**, 112-124.

59. Gorter, E. & Grendel, F. (1925). On bimolecular layers of lipoids on the chromocytes of the blood. *J. Exp. Med.* **41**, 439-443.

60. Singer, S. J. & Nicolson, G. L. (1972). The fluid mosaic model of the structure of cell membranes. *Science* **175**, 720-31.

61. Kusumi, A., Nakada, C., Ritchie, K., Murase, K., Suzuki, K., Murakoshi, H., Kasai, R. S., Kondo, J. & Fujiwara, T. (2005). Paradigm shift of the plasma membrane concept from the two-dimensional continuum fluid to the partitioned fluid: high-speed single-molecule tracking of membrane molecules. *Annu. Rev. Biophys. Biomol. Struct.* **34**, 351-378.

62. Fujiwara, T., Ritchie, K., Murakoshi, H., Jacobson, K. & Kusumi, A. (2002). Phospholipids undergo hop diffusion in compartmentalized cell membrane. *J. Cell Biol.* **157**, 1071-1082.

63. Suzuki, K., Ritchie, K., Kajikawa, E., Fujiwara, T. & Kusumi, A. (2005). Rapid hop diffusion of a G-protein-coupled receptor in the plasma membrane as revealed by single-molecule techniques. *Biophys. J.* **88**, 3659-3680.

64. Ruiz, N., Kahne, D. & Silhavy, T. J. (2006). Advances in understanding bacterial outer-membrane biogenesis. *Nat. Rev. Microbiol.* **4**, 57-66.

65. Bos, M. P., Robert, V. & Tommassen, J. (2007). Biogenesis of the Gram-negative bacterial outer membrane. *Annu. Rev. Microbiol.* **61**, 191-214.

66. Diedrich, D. L. & Cota-Robles, E. H. (1974). Heterogeneity in lipid composition of the outer membrane and cytoplasmic membrane of *Pseudomonas* BAL-31. *J. Bacteriol.* **119**, 1006-1018.

67. Missiakas, D., Betton, J. M. & Raina, S. (1996). New components of protein folding in extracytoplasmic compartments of *Escherichia coli* SurA, FkpA and Skp/OmpH. *Mol. Microbiol.* **21**, 871-884.

68. Ruiz, N., Kahne, D. & Silhavy, T. J. (2009). Transport of lipopolysaccharide across the cell envelope: the long road of discovery. *Nat. Rev. Microbiol.* **7**, 677-683.

69. Raetz, C. R. H. & Whitfield, C. (2002). Lipopolysaccharide endotoxins. *Annu. Rev. Biochem.* **71**, 635-700.

70. Driessen, A. J. M. & Nouwen, N. (2008). Protein translocation across the bacterial cytoplasmic membrane. *Annu. Rev. Biochem.* **77**, 643-667.

71. de Keyzer, J., van der Does, C. & Driessen, A. J. M. (2003). The bacterial translocase: a dynamic protein channel complex. *CMLS, Cell. Mol. Life Sci.* **60**, 2034-2052.

72. Rapoport, T. A. (2007). Protein translocation across the eukaryotic endoplasmic reticulum and bacterial plasma membranes. *Nature* **450**, 663-669.

73. Berg, B. v. d., Clemons, W. M., Collinson, I., Modis, Y., Hartmann, E., Harrison, S. C. & Rapoport, T. A. (2004). X-ray structure of a protein-conducting channel. *Nature* **427**, 36-44.

74. Egea, P. F. & Stroud, R. M. (2010). Lateral opening of a translocon upon entry of protein suggests the mechanism of insertion into membranes. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 17182-17187.

75. Tamm, L. K., Arora, A. & Kleinschmidt, J. H. (2001). Structure and assembly of beta-barrel membrane proteins. *J. Biol. Chem.* **276**, 32399-32402.

76. Mogensen, J. E. & Otzen, D. E. (2005). Interactions between folding factors and bacterial outer membrane proteins. *Mol. Microbiol.* **57**, 326-346.

77. Sohn, J., Grant, R. A. & Sauer, R. T. (2009). OMP peptides activate the DegS stress-sensor protease by a relief of inhibition mechanism. *Structure* **17**, 1411-1421.

78. Kennedy, M. B. (1995). Origin of PDZ (DHR, GLGF) domains. *Trends Biochem. Sci.* **20**, 350.

79. Cowburn, D. (1997). Peptide recognition by PTB and PDZ domains. *Curr. Opin. Struct. Biol.* **7**, 835-838.

80. Tormo, A., Almiron, M. & Kolter, R. (1990). *surA*, an *Escherichia coli* gene essential for survival in stationary phase. *J. Bacteriol.* **172**, 4339-47.

81. Lazar, S. W. & Kolter, R. (1996). SurA assists the folding of *Escherichia coli* outer membrane proteins. *J. Bacteriol.* **178**, 1770-1773.

82. Rouviere, P. E. & Gross, C. A. (1996). SurA, a periplasmic protein with peptidyl-prolyl isomerase activity, participates in the assembly of outer membrane porins. *Genes Dev.* **10**, 3170-3182.

83. Merdanovic, M., Clausen, T., Kaiser, M., Huber, R. & Ehrmann, M. (2011). Protein quality control in the bacterial periplasm. *Annu. Rev. Microbiol.* **65**, 149-168.

84. Johansen, J., Rasmussen, A. A., Overgaard, M. & Valentin-Hansen, P. (2006). Conserved small non-coding RNAs that belong to the σE regulon: role in down-regulation of outer membrane proteins. *J. Mol. Biol.* **364**, 1-8.

85. Vertommen, D., Ruiz, N., Leverrier, P., Silhavy, T. J. & Collet, J.-F. (2009). Characterization of the role of the *Escherichia coli* periplasmic chaperone SurA using differential proteomics. *Proteomics* **9**, 2432-2443.

86. Bitto, E. & McKay, D. B. (2002). Crystallographic structure of SurA, a molecular chaperone that facilitates folding of outer membrane porins. *Structure* **10**, 1489-1498.

87. Behrens, S., Maier, R., de Cock, H., Schmid, F. X. & Gross, C. A. (2001). The SurA periplasmic PPIase lacking its parvulin domains functions *in vivo* and has chaperone activity. *EMBO J.* **20**, 285-294.

88. Ricci, D. P., Schwalm, J., Gonzales-Cope, M. & Silhavy, T. J. (2013). The activity and specificity of the outer membrane protein chaperone SurA are modulated by a proline isomerase domain. *mBio* **4**, e00540-13.

89. Hennecke, G., Nolte, J., Volkmer-Engert, R., Schneider-Mergener, J. & Behrens, S. (2005). The periplasmic chaperone SurA exploits two features characteristic of integral outer membrane proteins for selective substrate recognition. *J. Biol. Chem.* **280**, 23540-23548.

90. Bitto, E. & McKay, D. B. (2003). The periplasmic molecular chaperone protein SurA binds a peptide motif that is characteristic of integral outer membrane proteins. *J. Biol. Chem.* **278**, 49316-49322.

91. Behrens-Kneip, S. (2010). The role of SurA factor in outer membrane protein transport and virulence. *Int. J. Med. Microbiol.* **300**, 421-428.

92. Burmann, B. M. & Hiller, S. (2012). Solution NMR studies of membrane-protein-chaperone complexes. *Chimia* **66**, 759-63.

93. Bitto, E. & McKay, D. B. (2004). Binding of phage-display-selected peptides to the periplasmic chaperone protein SurA mimics binding of unfolded outer membrane proteins. *FEBS Lett.* **568**, 94-98.

94. Sklar, J. G., Wu, T., Kahne, D. & Silhavy, T. J. (2007). Defining the roles of the periplasmic chaperones SurA, Skp, and DegP in *Escherichia coli*. *Genes Dev.* **21**, 2473-2484.

95. Rizzitello, A. E., Harper, J. R. & Silhavy, T. J. (2001). Genetic evidence for parallel pathways of chaperone activity in the periplasm of *Escherichia coli*. *J. Bacteriol.* **183**, 6794-6800.

96. Chen, R. & Henning, U. (1996). A periplasmic protein (Skp) of *Escherichia coli* selectively binds a class of outer membrane proteins. *Mol. Microbiol.* **19**, 1287-1294.

97. Tucker, C. L. & Fields, S. (2003). Lethal combinations. *Nat. Genet.* **35**, 204-205.

98. Denoncin, K., Schwalm, J., Vertommen, D., Silhavy, T. J. & Collet, J.-F. (2012). Dissecting the *Escherichia coli* periplasmic chaperone network using differential proteomics. *Proteomics* **12**, 1391-1401.

99. Schafer, U., Beck, K. & Muller, M. (1999). 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.

100. Schwalm, J., Mahoney, T. F., Soltes, G. R. & Silhavy, T. J. (2013). A role for Skp in LptD assembly in *Escherichia coli*. *J. Bacteriol.*, DOI: 10.1128/JB.00431-13

101. Qu, J., Mayer, C., Behrens, S., Holst, O. & Kleinschmidt, J. H. (2007). 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.

102. Walton, T. A., Sandoval, C. M., Fowler, C. A., Pardi, A. & Sousa, M. C. (2009). The cavity-chaperone Skp protects its substrate from aggregation but allows independent folding of substrate domains. *Proc. Natl. Acad. Sci. U.S.A* **106**, 1772-7.

103. Harms, N., Koningstein, G., Dontje, W., Muller, M., Oudega, B., Luirink, J. & de Cock, H. (2001). 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.

104. Lyu, Z.-X., Shao, Q., Gao, Y. Q. & Zhao, X. S. (2012). Direct observation of the uptake of outer membrane proteins by the periplasmic chaperone Skp. *PLoS ONE* **7**, e46068.

105. Korndorfer, I. P., Dommel, M. K. & Skerra, A. (2004). Structure of the periplasmic chaperone Skp suggests functional similarity with cytosolic chaperones despite differing architecture. *Nat. Struct. Mol. Biol.* **11**, 1015-1020.

106. Walton, T. A. & Sousa, M. C. (2004). Crystal structure of Skp, a prefoldin-like chaperone that protects soluble and membrane proteins from aggregation. *Mol. Cell* **15**, 367-374.

107. Hagan, C. L., Silhavy, T. J. & Kahne, D. (2011). β-barrel membrane protein assembly by the BAM complex. *Annu. Rev. Biochem.* **80**, 189-210.

108. Entzminger, K. C., Chang, C., Myhre, R. O., McCallum, K. C. & Maynard, J. A. (2012). The Skp chaperone helps fold soluble proteins *in vitro* by inhibiting aggregation. *Biochemistry* **51**, 4822-4834.

109. Nakamoto, H. & Bardwell, J. C. A. (2004). Catalysis of disulfide bond formation and isomerization in the *Escherichia coli* periplasm. *Biochim. Biophys. Acta - Mol. Cell Res.* **1694**, 111-119.

110. Saul, F. A., Arie, J. P., Vulliez-le Normand, B., Kahn, R., Betton, J. M. & Bentley, G. A. (2004). Structural and functional studies of FkpA from *Escherichia coli*, a cis/trans peptidyl-prolyl isomerase with chaperone activity. *J. Mol. Biol.* **335**, 595-608.

111. Krojer, T., Garrido-Franco, M., Huber, R., Ehrmann, M. & Clausen, T. (2002). Crystal structure of DegP (HtrA) reveals a new protease-chaperone machine. *Nature* **416**, 455-459.

112. Krojer, T., Sawa, J., Schafer, E., Saibil, H. R., Ehrmann, M. & Clausen, T. (2008). Structural basis for the regulated protease and chaperone function of DegP. *Nature* **453**, 885-890.

113. Sawa, J., Heuck, A., Ehrmann, M. & Clausen, T. (2010). Molecular transformers in the cell: lessons learned from the DegP protease–chaperone. *Curr. Opin. Struct. Biol.* **20**, 253-258.

114. Kim, S. & Sauer, R. T. (2012). 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.

115. 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. (2011). Genetic selection designed to stabilize proteins uncovers a chaperone called Spy. *Nat. Struct. Mol. Biol.* **18**, 262-269.

116. Volokhina, E. B., Grijpstra, J., Stork, M., Schilders, I., Tommassen, J. & Bos, M. P. (2011). Role of the periplasmic chaperones Skp, SurA, and DegQ in outer membrane protein biogenesis in *Neisseria meningitidis*. *J. Bacteriol.* **193**, 1612-1621.

117. Kim, K. H., Aulakh, S. & Paetzel, M. (2012). The bacterial outer membrane β-barrel assembly machinery. *Protein Sci.* **21**, 751-768.

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

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

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

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

122. Noinaj, N., Kuszak, A. J., Gumbart, J. C., Lukacik, P., Chang, H., Easley, N. C., Lithgow, T. & Buchanan, S. K. (2013). Structural insight into the biogenesis of β-barrel membrane proteins. *Nature* **501**, 385-390.

123. Leonard-Rivera, M. & Misra, R. (2012). Conserved residues of the putative L6 loop of *Escherichia coli* BamA play a critical role in the sssembly of β-Barrel outer membrane proteins, including that of BamA itself. *J. Bacteriol.* **194**, 4662-4668.

124. McMorran, L. M., Bartlett, A. I., Huysmans, G. H. M., Radford, S. E. & Brockwell, D. J. (2013). Dissecting the effects of periplasmic chaperones on the in vitro folding of the outer membrane protein PagP. *J. Mol. Biol.* **425**, 3178-3191.

125. Kim, K. H. & Paetzel, M. (2011). Crystal structure of *Escherichia coli* BamB, a lipoprotein component of the β-barrel assembly machinery complex. *J. Mol. Biol.* **406**, 667-678.

126. Kim, K. H., Aulakh, S. & Paetzel, M. (2011). Crystal structure of the β-barrel assembly machinery BamCD complex. *J. Biol. Chem.* **286**, 39116-21.

127. Kim, K. H., Kang, H. S., Okon, M., Escobar-Cabrera, E., McIntosh, L. P. & Paetzel, M. (2011). Structural characterization of Escherichia coli BamE, a lipoprotein component of the beta-barrel assembly machinery complex. *Biochemistry* **50**, 1081-90.

128. Robert, V., Volokhina, E. B., Senf, F., Bos, M. P., Van Gelder, P. & Tommassen, J. (2006). Assembly factor Omp85 recognizes its outer membrane protein substrates by a species-specific C-terminal motif. *Plos Biol.* **4**, 1984-1995.

129. Sandoval, C. M., Baker, S. L., Jansen, K., Metzner, S. I. & Sousa, M. C. (2011). Crystal structure of BamD: an essential component of the β-barrel sssembly machinery of Gram-negative bacteria. *J. Mol. Biol.* **409**, 348-357.

130. Webb, C. T., Selkrig, J., Perry, A. J., Noinaj, N., Buchanan, S. K. & Lithgow, T. (2012). Dynamic association of BAM complex modules includes surface exposure of the lipoprotein BamC. *J. Mol. Biol.* **422**, 545-555.

131. 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. (2011). Structure and function of BamE within the outer membrane and the β-barrel assembly machine. *EMBO Rep.* **12**, 123-128.

132. Ureta, A. R., Endres, R. G., Wingreen, N. S. & Silhavy, T. J. (2007). 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.

133. Heuck, A., Schleiffer, A. & Clausen, T. (2011). Augmenting β-augmentation: structural basis of how BamB binds BamA and may support folding of outer membrane proteins. *J. Mol. Biol.* **406**, 659-666.

134. Noinaj, N., Fairman, J. W. & Buchanan, S. K. (2011). The crystal structure of BamB suggests interactions with BamA and its role within the BAM complex. *J. Mol. Biol.* **407**, 248-260.

135. Dong, C., Yang, X., Hou, H.-F., Shen, Y.-Q. & Dong, Y.-H. (2012). Structure of *Escherichia coli* BamB and its interaction with POTRA domains of BamA. *Acta Crystallogr. Sect. D-Biol. Crystallogr.* **68**, 1134-1139.

136. Hagan, C. L., Westwood, D. B. & Kahne, D. (2013). BAM lipoproteins assemble BamA *in vitro*. *Biochemistry* **52**, 6108-6113.

137. Hagan, C. L., Kim, S. & Kahne, D. (2010). Reconstitution of outer membrane protein assembly from purified components. *Science* **328**, 890-892.

138. Kramer, R. A., Zandwijken, D., Egmond, M. R. & Dekker, N. (2000). *In vitro* folding, purification and characterization of *Escherichia coli* outer membrane protease OmpT. *Eur. J. Biochem.* **267**, 885-893.

139. Schweizer, M., Hindennach, I., Garten, W. & Henning, U. (1978). Major proteins of the *Escherichia coli* outer cell envelope membrane. Interaction of protein II with lipopolysaccharide. *Eur. J. Biochem.* **82**, 211-7.

140. Heller, K. B. (1978). Apparent molecular-weights of a heat-modifiable protein from outer membrane of *Escherichia coli* in gels with different acrylamide concentrations. *J. Bacteriol.* **134**, 1181-1183.

141. Hong, H., Joh, N. H., Bowie, J. U. & Tamm, L. K. (2009). Chapter 8: Methods for measuring the thermodynamic stability of membrane proteins. In *Methods Enzymol.*, Vol. 455, pp. 213-236.

142. Surrey, T. & Jähnig, F. (1992). Refolding and oriented insertion of a membrane protein into a lipid bilayer. *Proc. Natl. Acad. Sci. U.S.A.* **89**, 7457-7461.

143. Surrey, T. & Jähnig, F. (1995). Kinetics of folding and membrane insertion of a β-barrel membrane protein. *J. Biol. Chem.* **270**, 28199-28203.

144. Kleinschmidt, J. H. & Tamm, L. K. (1996). Folding intermediates of a β-barrel membrane protein. Kinetic evidence for a multi-step membrane insertion mechanism. *Biochemistry* **35**, 12993-13000.

145. Kleinschmidt, J. H., den Blaauwen, T., Driessen, A. J. M. & Tamm, L. K. (1999). Outer membrane protein A of *Escherichia coli* inserts and folds into lipid bilayers by a concerted mechanism. *Biochemistry* **38**, 5006-5016.

146. Kleinschmidt, J. H. (2006). Folding kinetics of the outer membrane proteins OmpA and FomA into phospholipid bilayers. *Chem. Phys. Lipids* **141**, 30-47.

147. Kleinschmidt, J. H. & Tamm, L. K. (2002). 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.

148. Hatzakis, N. S., Bhatia, V. K., Larsen, J., Madsen, K. L., Bolinger, P. Y., Kunding, A. H., Castillo, J., Gether, U., Hedegard, P. & Stamou, D. (2009). How curved membranes recruit amphipathic helices and protein anchoring motifs. *Nat. Chem. Biol.* **5**, 835-841.

149. Hong, H. D. & Tamm, L. K. (2004). Elastic coupling of integral membrane protein stability to lipid bilayer forces. *Proc. Natl. Acad. Sci. U.S.A.* **101**, 4065-4070.

150. Andersen, K. K., Wang, H. & Otzen, D. E. (2012). A kinetic analysis of the folding and unfolding of OmpA in urea and guanidinium chloride: single and parallel pathways. *Biochemistry* **51**, 8371–8383.

151. Pocanschi, C. L., Popot, J.-L. & Kleinschmidt, J. H. (2013). 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.

152. Stanley, A. M. & Fleming, K. G. (2008). The process of folding proteins into membranes: challenges and progress. *Arch. Biochem. Biophys.* **469**, 46-66.

153. Pocanschi, C. L., Patel, G. J., Marsh, D. & Kleinschmidt, J. H. (2006). Curvature elasticity and refolding of OmpA in large unilamellar vesicles. *Biophys. J.* **91**, L75-L77.

154. Hong, H., Szabo, G. & Tamm, L. K. (2006). Electrostatic couplings in OmpA ion-channel gating suggest a mechanism for pore opening. *Nat. Chem. Biol.* **2**, 627-635.

155. Hong, H., Park, S., Flores Jiménez, R. H., Rinehart, D. & Tamm, L. K. (2007). Role of aromatic side chains in the folding and thermodynamic stability of integral membrane proteins. *J. Am. Chem. Soc.* **129**, 8320-8327.

156. Killian, J. A. & von Heijne, G. (2000). How proteins adapt to a membrane-water interface. *Trends Biochem. Sci.* **25**, 429-434.

157. Wimley, W. C. & White, S. H. (1996). Experimentally determined hydrophobicity scale for proteins at membrane interfaces. *Nat. Struct. Mol. Biol.* **3**, 842-848.

158. Avanti Polar Lipids technical support. http://www.avantilipids.com/index.php?view=items&cid=5&id=10&option=com_quickfaq&Itemid=385, Accessed 19th July 2013.

159. Bulieris, P. V., Behrens, S., Holst, O. & Kleinschmidt, J. H. (2003). Folding and insertion of the outer membrane protein OmpA is assisted by the chaperone Skp and by lipopolysaccharide. *J. Biol. Chem.* **278**, 9092-9099.

160. Patel, G. J. & Kleinschmidt, J. H. (2013). 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.

161. Patel, G. J., Behrens-Kneip, S., Holst, O. & Kleinschmidt, J. H. (2009). The periplasmic chaperone Skp facilitates targeting, insertion, and folding of OmpA into lipid membranes with a negative membrane surface potential. *Biochemistry* **48**, 10235-10245.

162. Burgess, N. K., Dao, T. P., Stanley, A. M. & Fleming, K. G. (2008). 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.

163. Pocanschi, C. L., Apell, H. J., Puntervoll, P., Hogh, B., Jensen, H. B., Welte, W. & Kleinschmidt, J. H. (2006). 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.

164. Tan, A. E., Burgess, N. K., DeAndrade, D. S., Marold, J. D. & Fleming, K. G. (2010). Self-association of unfolded outer membrane proteins. *Macromol. Biosci.* **10**, 763-767.

165. Moon, C. P., Kwon, S. & Fleming, K. G. (2011). Overcoming hysteresis to attain reversible equilibrium folding for Outer Membrane Phospholipase A in phospholipid bilayers. *J. Mol. Biol.* **413**, 484-494.

166. Bishop, R. E., Gibbons, H. S., Guina, T., Trent, M. S., Miller, S. I. & Raetz, C. R. (2000). Transfer of palmitate from phospholipids to lipid A in outer membranes of Gram-negative bacteria. *EMBO J.* **19**, 5071-80.

167. Huysmans, G. H. M., Baldwin, S. A., Brockwell, D. J. & Radford, S. E. (2010). The transition state for folding of an outer membrane protein. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 4099-4104.

168. Shanmugavadivu, B., Apell, H. J., Meins, T., Zeth, K. & Kleinschmidt, J. H. (2007). Correct folding of the β-barrel of the human membrane protein VDAC requires a lipid bilayer. *J. Mol. Biol.* **368**, 66-78.

169. Sapra, K. T., Damaghi, M., Köster, S., Yildiz, Ö., Kühlbrandt, W. & Muller, D. J. (2009). One β-hairpin after the other: exploring mechanical unfolding pathways of the transmembrane β-barrel protein OmpG. *Angew. Chem. Int. Ed.* **48**, 8306-8308.

170. Damaghi, M., Bippes, C., Koster, S., Yildiz, O., Mari, S. A., Kuhlbrandt, W. & Muller, D. J. (2010). pH-dependent interactions guide the folding and gate the transmembrane pore of the β-barrel membrane protein OmpG. *J. Mol. Biol.* **397**, 878-882.

171. Mari, S. A., Köster, S., Bippes, C. A., Yildiz, Ö., Kühlbrandt, W. & Muller, D. J. (2010). pH-induced conformational change of the β-barrel-forming protein OmpG reconstituted into native *E. coli* lipids. *J. Mol. Biol.* **396**, 610-616.

172. Damaghi, M., Köster, S., Bippes, C. A., Yildiz, Ö. & Müller, D. J. (2011). One β-hairpin follows the other: exploring refolding pathways and kinetics of the transmembrane β-barrel protein OmpG. *Angew. Chem. Int. Ed.* **50**, 7422-7424.

173. Surrey, T., Schmid, A. & Jähnig, F. (1996). Folding and membrane insertion of the trimeric β-barrel protein OmpF. *Biochemistry* **35**, 2283-2288.

174. Moon, C. P., Zaccai, N. R., Fleming, P. J., Gessmann, D. & Fleming, K. G. (2013). 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.

175. Nickson, A. A. & Clarke, J. (2010). What lessons can be learned from studying the folding of homologous proteins? *Methods* **52**, 38-50.

176. Banachewicz, W., Religa, T. L., Schaeffer, R. D., Daggett, V. & Fersht, A. R. (2011). Malleability of folding intermediates in the homeodomain superfamily. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 5596-5601.

177. Nickson, A. A., Wensley, B. G. & Clarke, J. (2013). Take home lessons from studies of related proteins. *Curr. Opin. Struct. Biol.* **23**, 66-74.

178. Capaldi, A. P., Kleanthous, C. & Radford, S. E. (2002). Im7 folding mechanism: misfolding on a path to the native state. *Nat. Struct. Biol.* **9**, 209-216.

179. Friel, C. T., Capaldi, A. P. & Radford, S. E. (2003). Structural analysis of the rate-limiting transition states in the folding of Im7 and Im9: similarities and differences in the folding of homologous proteins. *J. Mol. Biol.* **326**, 293-305.

180. Friel, C. T., Beddard, G. S. & Radford, S. E. (2004). Switching two-state to three-state kinetics in the helical protein Im9 via the optimisation of stabilising non-native interactions by design. *J. Mol. Biol.* **342**, 261-273.

181. Lappalainen, I., Hurley, M. G. & Clarke, J. (2008). Plasticity within the obligatory folding nucleus of an immunoglobulin-like domain. *J. Mol. Biol.* **375**, 547-559.

182. Wright, C. F., Lindorff-Larsen, K., Randles, L. G. & Clarke, J. (2003). Parallel protein-unfolding pathways revealed and mapped. *Nat. Struct. Biol.* **10**, 658-662.

183. Religa, T. (2008). Comparison of multiple crystal structures with NMR data for engrailed homeodomain. *J. Biomol. NMR* **40**, 189-202.

184. Kusunoki, H., Minasov, G., MacDonald, R. I. & Mondragón, A. (2004). Independent movement, dimerization and stability of tandem repeats of chicken brain α-spectrin. *J. Mol. Biol.* **344**, 495-511.

185. Dennis, C. A., Videler, H., Pauptit, R. A., Wallis, R., James, R., Moore, G. R. & Kleanthous, C. (1998). A structural comparison of the colicin immunity proteins Im7 and Im9 gives new insights into the molecular determinants of immunity-protein specificity. *Biochem. J.* **333**, 183-191.

186. Improta, S., Politou, A. S. & Pastore, A. (1996). Immunoglobulin-like modules from titin I-band: extensible components of muscle elasticity. *Structure* **4**, 323-337.

187. Clark, P. L. (2011). Adding protease digestion to the membrane protein toolbox. *J. Mol. Biol.* **406**, 543-544.

188. Radford, S. E., Dobson, C. M. & Evans, P. A. (1992). The folding of hen lysozyme involves partially structured intermediates and multiple pathways. *Nature* **358**, 302-307.

189. Le Duff, C. S., Whittaker, S. B. M., Radford, S. E. & Moore, G. R. (2006). Characterisation of the conformational properties of urea-unfolded Im7: implications for the early stages of protein folding. *J. Mol. Biol.* **364**, 824-835.

190. Tafer, H., Hiller, S., Hilty, C., Fernández, C. & Wüthrich, K. (2004). Nonrandom structure in the urea-unfolded *Escherichia coli* outer membrane protein X (OmpX). *Biochemistry* **43**, 860-869.

191. Hiller, S., Wider, G., Imbach, L. L. & Wüthrich, K. (2008). Interactions with hydrophobic clusters in the urea-unfolded membrane protein OmpX. *Angew. Chem. Int. Ed.* **47**, 977-981.

192. Raschle, T., Hiller, S., Etzkorn, M. & Wagner, G. (2010). Nonmicellar systems for solution NMR spectroscopy of membrane proteins. *Curr. Opin. Struct. Biol.* **20**, 471-479.

193. Kielec, J. M., Valentine, K. G. & Wand, A. J. (2010). A method for solution NMR structural studies of large integral membrane proteins: reverse micelle encapsulation. *Biochim. Biophys. Acta Biomemb.* **1798**, 150-160.

194. Glück, J. M., Wittlich, M., Feuerstein, S., Hoffmann, S., Willbold, D. & Koenig, B. W. (2009). Integral membrane proteins in nanodiscs can be studied by solution NMR spectroscopy. *J. Am. Chem. Soc.* **131**, 12060-12061.

195. Curnow, P. & Booth, P. J. (2009). The transition state for integral membrane protein folding. *Proc. Natl. Acad. Sci. U.S.A.* **106**, 773-778.

196. Curnow, P., Di Bartolo, N. D., Moreton, K. M., Ajoje, O. O., Saggese, N. P. & Booth, P. J. (2011). Stable folding core in the folding transition state of an α-

helical integral membrane protein. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 14133-14138.

197. Otzen, D. E. (2011). Mapping the folding pathway of the transmembrane protein DsbB by protein engineering. *Protein Eng. Des. Sel.* **24**, 139-149.

198. Fersht, A. R. & Sato, S. (2004). Phi-value analysis and the nature of protein-folding transition states. *Proc. Natl. Acad. Sci. U.S.A.* **101**, 7976-7981.

199. Matouschek, A., Kellis, J. T., Serrano, L. & Fersht, A. R. (1989). Mapping the transition-state and pathway of protein folding by protein engineering. *Nature* **340**, 122-126.

200. Fersht, A. R., Matouschek, A. & Serrano, L. (1992). The folding of an enzyme. I. Theory of protein engineering analysis of stability and pathway of protein folding. *J. Mol. Biol.* **224**, 771-782.

201. Friel, C. T., Smith, D. A., Vendruscolo, M., Gsponer, J. & Radford, S. E. (2009). The mechanism of folding of Im7 reveals competition between functional and kinetic evolutionary constraints. *Nat. Struct. Mol. Biol.* **16**, 318-324.

202. Sohl, J. L., Jaswal, S. S. & Agard, D. A. (1998). Unfolded conformations of α-lytic protease are more stable than its native state. *Nature* **395**, 817-819.

203. Jaswal, S. S., Sohl, J. L., Davis, J. H. & Agard, D. A. (2002). Energetic landscape of α-lytic protease optimizes longevity through kinetic stability. *Nature* **415**, 343-346.

204. Truhlar, S. M. E. & Agard, D. A. (2005). The folding landscape of an α-lytic protease variant reveals the role of a conserved β-hairpin in the development of kinetic stability. *Proteins: Struct., Funct., Genet.* **61**, 105-114.

205. Konermann, L., Pan, Y. & Stocks, B. B. (2011). Protein folding mechanisms studied by pulsed oxidative labeling and mass spectrometry. *Curr. Opin. Struct. Biol.* **21**, 634-640.

206. Hambly, D. & Gross, M. (2005). Laser flash photolysis of hydrogen peroxide to oxidize protein solvent-accessible residues on the microsecond timescale. *J. Am. Soc. Mass Spectrom.* **16**, 2057-2063.

207. Vahidi, S., Stocks, B. B., Liaghati-Mobarhan, Y. & Konermann, L. (2012). Mapping pH-induced protein structural changes under equilibrium conditions by pulsed oxidative labeling and mass spectrometry. *Anal. Chem.* **84**, 9124-9130.

208. Chen, J., Rempel, D. L., Gau, B. C. & Gross, M. L. (2012). Fast photochemical oxidation of proteins and mass spectrometry follow submillisecond protein folding at the amino-acid level. *J. Am. Chem. Soc.* **134**, 18724-18731.

209. Pan, Y., Stocks, B. B., Brown, L. & Konermann, L. (2008). Structural characterization of an integral membrane protein in its natural lipid environment by oxidative methionine labeling and mass spectrometry. *Anal. Chem.* **81**, 28-35.

210. Pan, Y., Brown, L. & Konermann, L. (2009). Mapping the structure of an integral membrane protein under semi-denaturing conditions by laser-induced oxidative labeling and mass spectrometry. *J. Mol. Biol.* **394**, 968-981.

211. Park, C. & Marqusee, S. (2005). Pulse proteolysis: A simple method for quantitative determination of protein stability and ligand binding. *Nat. Meth.* **2**, 207-212.

212. Schlebach, J. P., Kim, M.-S., Joh, N. H., Bowie, J. U. & Park, C. (2011). Probing membrane protein unfolding with pulse proteolysis. *J. Mol. Biol.* **406**, 545-551.

213. Booth, P. J. & Curnow, P. (2006). Membrane proteins shape up: understanding *in vitro* folding. *Curr. Opin. Struct. Biol.* **16**, 480-488.

214. Oesterhelt, D. & Stoeckenius, W. (1971). Rhodopsin-like protein from the purple membrane of *Halobacterium halobium*. *Nature: New Biol.* **233**, 149-152.

215. Booth, P. J. (2000). Unravelling the folding of bacteriorhodopsin. *Biochim. Biophys. Acta Bioenergetics* **1460**, 4-14.

216. Huang, K. S., Bayley, H., Liao, M. J., London, E. & Khorana, H. G. (1981). Refolding of an integral membrane protein. Denaturation, renaturation, and reconstitution of intact bacteriorhodopsin and two proteolytic fragments. *J. Biol. Chem.* **256**, 3802-3809.

217. London, E. & Khorana, H. G. (1982). Denaturation and renaturation of bacteriorhodopsin in detergents and lipid-detergent mixtures. *J. Biol. Chem.* **257**, 7003-7011.

218. Popot, J. L., Gerchman, S. E. & Engelman, D. M. (1987). Refolding of bacteriorhodopsin in lipid bilayers. A thermodynamically controlled two-stage process. *J. Mol. Biol.* **198**, 655-676.

219. Hunt, J. F., Earnest, T. N., Bousché, O., Kalghatgi, K., Reilly, K., Horváth, C., Rothschild, K. J. & Engleman, D. M. (1997). A biophysical study of integral membrane protein folding. *Biochemistry* **36**, 15156-15176.

220. Riley, M. L., Wallace, B. A., Flitsch, S. L. & Booth, P. J. (1997). Slow α-helix formation during folding of a membrane protein. *Biochemistry* **36**, 192-196.

221. Booth, P. J., Flitsch, S. L., Stern, L. J., Greenhalgh, D. A., Kim, P. S. & Khorana, H. G. (1995). Intermediates in the folding of the membrane protein bacteriorhodopsin. *Nat. Struct. Mol. Biol.* **2**, 139-143.

222. Booth, P. J., Riley, M. L., Flitsch, S. L., Templer, R. H., Farooq, A., Curran, A. R., Chadborn, N. & Wright, P. (1997). Evidence that bilayer bending rigidity affects membrane protein folding. *Biochemistry* **36**, 197-203.

223. Curran, A. R., Templer, R. H. & Booth, P. J. (1999). Modulation of folding and assembly of the membrane protein bacteriorhodopsin by intermolecular forces within the lipid bilayer. *Biochemistry* **38**, 9328-9336.

224. Curnow, P. & Booth, P. J. (2007). Combined kinetic and thermodynamic analysis of alpha-helical membrane protein unfolding. *Proc. Natl. Acad. Sci. U.S.A.* **104**, 18970-18975.

225. Otzen, D. E. (2003). Folding of DsbB in mixed micelles: a kinetic analysis of the stability of a bacterial membrane protein. *J. Mol. Biol.* **330**, 641-649.

226. Booth, P. J. (2012). A successful change of circumstance: a transition state for membrane protein folding. *Curr. Opin. Struct. Biol.* **22**, 469-475.

227. Schlebach, J. P., Cao, Z., Bowie, J. U. & Park, C. (2012). Revisiting the folding kinetics of bacteriorhodopsin. *Protein Sci.* **21**, 97-106.

228. Joh, N. H., Min, A., Faham, S., Whitelegge, J. P., Yang, D., Woods Jr, V. L. & Bowie, J. U. (2008). Modest stabilization by most hydrogen-bonded side-chain interactions in membrane proteins. *Nature* **453**, 1266-1270.

229. Zhang, Z. & Smith, D. L. (1993). Determination of amide hydrogen exchange by mass spectrometry: a new tool for protein structure elucidation. *Protein Sci.* **2**, 522-31.

230. Bishop, R. E. (2005). The lipid A palmitoyltransferase PagP: molecular mechanisms and role in bacterial pathogenesis. *Mol. Microbiol.* **57**, 900-912.

231. 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. (2002). Solution structure and dynamics of the outer membrane enzyme PagP by NMR. *Proc. Natl. Acad. Sci. U.S.A.* **99**, 13560-13565.

232. Cuesta-Seijo, J. A., Neale, C., Khan, M. A., Moktar, J., Tran, C. D., Bishop, R. E., Pomès, R. & Privé, G. G. (2010). PagP crystallized from SDS/cosolvent reveals the route for phospholipid access to the hydrocarbon ruler. *Structure* **18**, 1210-1219.

233. Schulz, G. E. (2002). The structure of bacterial outer membrane proteins. *Biochim. Biophys. Acta Biomemb.* **1565**, 308-317.

234. Hwang, P. M., Bishop, R. E. & Kay, L. E. (2004). The integral membrane enzyme PagP alternates between two dynamically distinct states. *Proc. Natl. Acad. Sci. U.S.A.* **101**, 9618-9623.

235. Khan, M. A., Neale, C., Michaux, C., Pomes, R., Prive, G. G., Woody, R. W. & Bishop, R. E. (2007). Gauging a hydrocarbon ruler by an intrinsic exciton probe. *Biochemistry* **46**, 4565-4579.

236. Khan, M. A., Moktar, J., Mott, P. J., Vu, M., McKie, A. H., Pinter, T., Hof, F. & Bishop, R. E. (2010). Inscribing the perimeter of the PagP hydrocarbon ruler by site-specific chemical alkylation. *Biochemistry* **49**, 9046-9057.

237. Khan, M. A. & Bishop, R. E. (2009). Molecular mechanism for lateral lipid diffusion between the outer membrane external leaflet and a β-barrel hydrocarbon ruler. *Biochemistry* **48**, 9745-9756.

238. Huysmans, G. H. M., Radford, S. E., Brockwell, D. J. & Baldwin, S. A. (2007). The N-terminal helix is a post-assembly clamp in the bacterial outer membrane protein PagP. *J. Mol. Biol.* **373**, 529-540.

239. Huysmans, G. H. M. (2008). On the folding mechanism of the bacterial outer membrane protein PagP. Ph.D. Thesis.

240. Huysmans, G. H. M., Radford, S. E., Baldwin, S. A. & Brockwell, D. J. (2012). 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.

241. Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. (1990). Basic local alignment search tool. *J. Mol. Biol.* **215**, 403-10.

242. Kramer, R. A., Dekker, N. & Egmond, M. R. (2000). Identification of active site serine and histidine residues in *Escherichia coli* outer membrane protease OmpT. *FEBS Lett.* **468**, 220-224.

243. McCarter, J. D., Stephens, D., Shoemaker, K., Rosenberg, S., Kirsch, J. F. & Georgiou, G. (2004). Substrate specificity of the *Escherichia coli* outer membrane protease OmpT. *J. Bacteriol.* **186**, 5919-5925.

244. Kaufmann, A., Stierhof, Y. D. & Henning, U. (1994). New outer membrane-associated protease of *Escherichia coli* K-12. *J. Bacteriol.* **176**, 359-367.

245. Hwang, B. Y., Varadarajan, N., Li, H. X., Rodriguez, S., Iverson, B. L. & Georgiou, G. (2007). Substrate specificity of the *Escherichia coli* outer membrane protease OmpP. *J. Bacteriol.* **189**, 522-530.

246. Bartlett, A. I. (2010). Kinetic analysis of protein folding energy landscapes. Ph.D Thesis.

247. Bullock, W. O., Fernandez, J. M. & Short, J. M. (1987). XL1-Blue - A high-efficiency plasmid transforming *recA Escherichia coli* strain with beta-galactosidase selection. *Biotechniques* **5**, 376-379.

248. Studier, F. W., Rosenberg, A. H., Dunn, J. J. & Dubendorff, J. W. (1990). Use of T7 RNA-Polymerase to direct expression of cloned genes. *Methods Enzymol.* **185**, 60-89.

249. Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989). *Molecular cloning : a laboratory manual* 2nd ed edit, Cold Spring Harbor Laboratory Press.

250. Hagan, C. L. & Kahne, D. (2011). The reconstituted *Escherichia coli* BAM complex catalyzes multiple rounds of β-barrel assembly. *Biochemistry* **50**, 7444-7446.

251. Gasteiger, E., Gattiker, A., Hoogland, C., Ivanyi, I., Appel, R. D. & Bairoch, A. (2003). ExPASy: the proteomics server for in-depth protein knowledge and analysis. *Nucl. Acids Res.* **31**, 3784-3788.

252. Moon, C. P., Fleming, K. G., Michael L. Johnson, J. M. H. & Gary, K. A. (2011). Using tryptophan fluorescence to measure the stability of membrane proteins folded in liposomes. In *Methods Enzymol.*, Vol. 492, pp. 189-211.

253. Duche, D., Issouf, M. & Lloubes, R. (2009). Immunity protein protects colicin E2 from OmpT protease. *J. Biochem.* **145**, 95-101.

254. Hritonenko, V. & Stathopoulos, C. (2007). Omptin proteins: an expanding family of outer membrane proteases in Gram-negative Enterobacteriaceae. *Mol. Membr. Biol.* **24**, 395-406.

255. Sugimura, K. & Nishihara, T. (1988). Purification, characterization and primary structure of *Escherichiacoli* protease-VII with specificity for paired basic residues: identity of protease-VII and OmpT. *J. Bacteriol.* **170**, 5625-5632.

256. Kramer, R. A., Vandeputte-Rutten, L., de Roon, G. J., Gros, P., Dekker, N. & Egmond, M. R. (2001). Identification of essential acidic residues of outer membrane protease OmpT supports a novel active site. *FEBS Lett.* **505**, 426-430.

257. Neri, M., Baaden, M., Carnevale, V., Anselmi, C., Maritan, A. & Carloni, P. (2008). Microseconds dynamics simulations of the outer-membrane protease T. *Biophys. J.* **94**, 71-78.

258. Kramer, R. A., Brandenburg, K., Vandeputte-Rutten, L., Werkhoven, M., Gros, P., Dekker, N. & Egmond, M. R. (2002). Lipopolysaccharide regions involved in the activation of *Escherichia coli* outer membrane protease OmpT. *Eur. J. Biochem.* **269**, 1746-1752.

259. Ye, C., Chai, Q., Zhong, M. & Wei, Y. (2013). Effect of crowding by Ficolls on OmpA and OmpT refolding and membrane insertion. *Protein Sci.* **22**, 239-245.

260. Eswar, N., Webb, B., Marti-Renom, M. A., Madhusudhan, M. S., Eramian, D., Shen, M.-Y., Pieper, U. & Sali, A. (2002). Comparative protein structure

modeling using Modeller. In *Curr. Protoc. Bioinformatics*. John Wiley & Sons, Inc.

261. Lewis, B. A. & Engelman, D. M. (1983). Lipid bilayer thickness varies linearly with acyl chain length in fluid phosphatidylcholine vesicles. *J. Mol. Biol.* **166**, 211-217.

262. Dekker, N., Cox, R. C., Kramer, R. A. & Egmond, M. R. (2001). Substrate specificity of the integral membrane protease OmpT determined by spatially addressed peptide libraries. *Biochemistry* **40**, 1694-1701.

263. Zhou, Y. F., Lau, F. W., Nauli, S., Yang, D. & Bowie, J. U. (2001). Inactivation mechanism of the membrane protein diacylglycerol kinase in detergent solution. *Protein Sci.* **10**, 378-383.

264. Popot, J.-L. (2010). Amphipols, nanodiscs, and fluorinated surfactants: three nonconventional approaches to studying membrane proteins in aqueous solutions. *Annu. Rev. Biochem.* **79**, 737-775.

265. Bechara, C., Bolbach, G., Bazzaco, P., Sharma, K. S., Durand, G., Popot, J.-L., Zito, F. & Sagan, S. (2012). MALDI-TOF mass spectrometry analysis of amphipol-trapped membrane proteins. *Anal. Chem.* **84**, 6128-6135.

266. Ruotolo, B. T., Benesch, J. L. P., Sandercock, A. M., Hyung, S.-J. & Robinson, C. V. (2008). Ion mobility-mass spectrometry analysis of large protein complexes. *Nat. Protocols* **3**, 1139-1152.

267. Leney, A. C., McMorran, L. M., Radford, S. E. & Ashcroft, A. E. (2012). Amphipathic polymers enable the study of functional membrane proteins in the gas phase. *Anal. Chem.* **84**, 9841-9847.

268. Sonntag, Y., Musgaard, M., Olesen, C., Schiøtt, B., Møller, J. V., Nissen, P. & Thøgersen, L. (2011). Mutual adaptation of a membrane protein and its lipid bilayer during conformational changes. *Nat. Commun.* **2**, 304.

269. Granseth, E., von Heijne, G. & Elofsson, A. (2005). A study of the membrane-water interface region of membrane proteins. *J. Mol. Biol.* **346**, 377-385.

270. Monne, M., Nilsson, I., Johansson, M., Elmhed, N. & von Heijne, G. (1998). Positively and negatively charged residues have different effects on the position in the membrane of a model transmembrane helix. *J. Mol. Biol.* **284**, 1177-1183.

271. Braun, P. & von Heijne, G. (1999). The aromatic residues Trp and Phe have different effects on the positioning of a transmembrane helix in the microsomal membrane. *Biochemistry* **38**, 9778-9782.

272. Yau, W. M., Wimley, W. C., Gawrisch, K. & White, S. H. (1998). The preference of tryptophan for membrane interfaces. *Biochemistry* **37**, 14713-14718.

273. Hong, H., Rinehart, D. & Tamm, L. K. (2013). Membrane depth-dependent energetic contribution of the tryptophan side chain to the stability of integral membrane proteins. *Biochemistry* **52**, 4413-4421.

274. Gruner, S. M. (1985). Intrinsic curvature hypothesis for biomembrane lipid-composition - A role for nonbilayer lipids. *Proc. Natl. Acad. Sci. U.S.A.* **82**, 3665-3669.

275. Allen, S. J., Curran, A. R., Templer, R. H., Meijberg, W. & Booth, P. J. (2004). Controlling the folding efficiency of an integral membrane protein. *J. Mol. Biol.* **342**, 1293-1304.

276. Ishinaga, M., Kanamoto, R. & Kito, M. (1979). Distribution of phospholipid molecular species in outer and cytoplasmic membranes of *Escherichia coli*. *J. Biochem.* **86**, 161-165.

277. Capaldi, A. P., Kleanthous, C. & Radford, S. E. (2002). Im7 folding mechanism: misfolding on a path to the native state. *Nat. Struct. Mol. Biol.* **9**, 209-216.

278. Di Nardo, A. A., Korzhnev, D. M., Stogios, P. J., Zarrine-Afsar, A., Kay, L. E. & Davidson, A. R. (2004). Dramatic acceleration of protein folding by stabilization of a nonnative backbone conformation. *Proc. Natl. Acad. Sci. U.S.A.* **101**, 7954-7959.

279. Jackson, S. E. & Fersht, A. R. (1991). Folding of chymotrypsin inhibitor 2. 1. Evidence for a two-state transition. *Biochemistry* **30**, 10428-35.

280. Tanford, C. (1970). Protein denaturation. C. Theoretical models for the mechanism of denaturation. *Adv. Protein Chem.* **24**, 1-95.

281. Puorger, C., Eidam, O., Capitani, G., Erilov, D., Grutter, M. G. & Glockshuber, R. (2008). Infinite kinetic stability against dissociation of supramolecular protein complexes through donor strand complementation. *Structure* **16**, 631-642.

282. Silhavy, T. J., Kahne, D. & Walker, S. (2010). The bacterial cell envelope. *Cold Spring Harb. Perspect. Biol.* **2**, a000414.

283. Dunn, B. M., Pham, C., Raney, L., Abayasekara, D., Gillespie, W. & Hsu, A. (1981). Interaction of alpha-dansylated peptide inhibitors with porcine pepsin: detection of complex formation by fluorescence energy transfer and chromatography and evidence for a two-step binding scheme. *Biochemistry* **20**, 7206-11.

284. Lakowicz, J. R. (2006). *Principles of Fluorescence Spectroscopy, 3rd edition*, Springer.

285. Stryer, L. (1968). Fluorescence spectroscopy of proteins. *Science* **162**, 526-533.

286. Cardamone, M. & Puri, N. K. (1992). Spectrofluorimetric assessment of the surface hydrophobicity of proteins. *Biochem. J.* **282 (Pt 2)**, 589-93.

287. Ohgushi, M. & Wada, A. (1983). 'Molten-globule state': a compact form of globular proteins with mobile side-chains. *FEBS Lett.* **164**, 21-24.

288. GELifeSciences. (2013). Superdex prep grade and prepacked HiLoad columns - Manual. Data file 18-1100-52 AF.

289. Ieva, R., Tian, P., Peterson, J. H. & Bernstein, H. D. (2011). Sequential and spatially restricted interactions of assembly factors with an autotransporter β domain. *Proc. Natl. Acad. Sci. U.S.A.* **108**, E383–E391.

290. Qu, J., Behrens-Kneip, S., Holst, O. & Kleinschmidt, J. r. H. (2009). Binding regions of outer membrane protein A in complexes with the periplasmic chaperone Skp. A site-directed fluorescence study. *Biochemistry* **48**, 4926-4936.

291. Bennion, D., Charlson, E. S., Coon, E. & Misra, R. (2010). Dissection of β-barrel outer membrane protein assembly pathways through characterizing BamA POTRA 1 mutants of *Escherichia coli*. *Mol. Microbiol.* **77**, 1153-1171.

292. Wu, S., Ge, X., Lv, Z., Zhi, Z., Chang, Z. & Zhao, X. S. (2011). Interaction between bacterial outer membrane proteins and periplasmic quality control factors: a kinetic partitioning mechanism. *Biochem. J.* **438**, 505-11.

293. Blake, C. C. F., Koenig, D. F., Mair, G. A., North, A. C. T., Phillips, D. C. & Sarma, V. R. (1965). Structure of hen egg-white lysozyme: a three-dimensional fourier synthesis at 2 Å resolution. *Nature* **206**, 757-761.

294. Alam, J. M. & Yamazaki, M. (2011). Spontaneous insertion of lipopolysaccharide into lipid membranes from aqueous solution. *Chem. Phys. Lipids* **164**, 166-174.

295. Tommassen, J. (2010). Assembly of outer-membrane proteins in bacteria and mitochondria. *Microbiology* **156**, 2587-2596.

296. Booth, P. J. & Clarke, J. (2010). Membrane protein folding makes the transition. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 3947-3948.

297. Moon, C. P. & Fleming, K. G. (2011). Side-chain hydrophobicity scale derived from transmembrane protein folding into lipid bilayers. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 10174-10177.

APPENDIX

The full DNA sequence of the cloned His-tagged OmpT gene (plasmid pLMM01) is given below:

```
GCTACATATG CATCATCACC ATCACCACGA AAATCTGTAC TTCCAGGGTT CTACCGAGA
CTTTATCGTT TACTCCTGAC AACATAAATG CGGACATTAG TCTTGGAACT CTGAGCGGAA
AAACAAAAGA GCGTGTTTAT CTAGCCGAAG AAGGAGGCCG AAAAGTCAGT CAACTCGACT
GGAAATTCAA TAACGCTGCA ATTATTAAAG GTGCAATTAA TTGGGATTTG ATGCCCCAGA
TATCTATCGG GGCTGCTGGC TGGACAACTC TCGGCAGCCG AGGTGGCAAT ATGGTCGATC
AGGACTGGAT GGATTCCAGT AACCCCGGAA CCTGGACGGA TGAAAGTAGA CACCCTGATA
CACAACTCAA TTATGCCAAC GAATTTGATC TGAATATCAA AGGCTGGCTC CTCAACGAAC
CCAATTACCG CCTGGGACTC ATGGCCGGAT ATCAGGAAAG CCGTTATAGC TTTACAGCCA
GAGGTGGTTC CTATATCTAC AGTTCTGAGG AGGGATTCAG AGATGATATC GGCTCCTTCC
CGAATGGAGA AAGAGCAATC GGCTACAAAC AACGTTTTAA AATGCCCTAC ATTGGCTTGA
CTGGAAGTTA TCGTTATGAA GATTTTGAAC TCGGTGGCAC ATTTAAATAC AGCGGCTGGG
TGGAATCATC TGATAACGAT GAACACTATG ACCCGGGAAA AAGAATCACT TATCGCAGTA
AGGTCAAAGA CCAAAATTAC TATTCTGTTG CAGTCAATGC AGGTTATTAC GTCACACCTA
ACGCAAAAGT TTATGTTGAA GGCGCATGGA ATCGGGTTAC GAATAAAAAA GGTAATACTT
CACTTTATGA TCACAATAAT AACACTTCAG ACTACAGCAA AAATGGAGCA GGTATAGAAA
ACTATAACTT CATCACTACT GCTGGTCTTA AGTACACATT TTAAGGATCC GCC
```

The start codon (located within the *Nde*I site) is shown in green and the stop codon is shown in red. The restriction enzyme recognition sites are underlined.

The full protein sequence of His-tagged OmpT is given below:

```
                ↓
MHHHHHHENL YFQGSTETLS FTPDNINADI SLGTLSGKTK ERVYLAEEGG RKVSQLDWKF
NNAAIIKGAI NWDLMPQISI GAAGWTTLGS RGGNMVDQDW MDSSNPGTWT DESRHPDTQL
NYANEFDLNI KGWLLNEPNY RLGLMAGYQE SRYSFTARGG SYIYSSEEGF RDDIGSFPNG
ERAIGYKQRF KMPYIGLTGS YRYEDFELGG TFKYSGWVES SDNDEHYDPG KRITYRSKVK
DQNYYSVAVN AGYYVTPNAK VYVEGAWNRV TNKKGNTSLY DHNNNTSDYS KNGAGIENYN
FITTAGLKYT F
```

The His-tag is shown in blue, the TEV recognition site in purple, and the cut site marked by an arrow.

The full DNA sequence of the cloned His-tagged OmpP gene (plasmid pLMM02) is given below:

GCTACATATG CATCATCACC ATCACCACGA AAATCTGTAC TTCCAGGGTT
CTGATTTCTT CGGCCCGGAG AAAATAAGTA CAGAGATTAA TTTAGGTACA
CTGAGTGGCA AAACAAAAGA ACGGGTTTAT GAGCCTGAAG AAGGTGGACG
TAAAGTCAGC CAACTGGACT GGAAATACAG TAACGCCGCC ATTCTTAAAG
GCGCCGTTAA CTGGGAGCTG AATCCATGGT TATCTGTTGG TGCAGCTGGC
TGGACCACTC TCAATAGTCG GGGGGGGAAT ATGGTTGATC AGGACTGGAT
GGATTCCGGG ACTCCCGGAA CATGGACAGA TGAAAGCAGG CATCCTGATA
CACGTCTTAA TTATGCCAAC GAATTTGATT TGAACGTTAA AGGCTGGTTT
TTAAAAGAAT CTGATTATCG CCTTGCTATT ATGGCAGGTT ATCAGGAAAG
CCGTTACAGT TTTAATGCTA CAGGAGGAAC TTATATTTAT AGTGAGAATG
GTGGTTTCCG GAATGAAACG GGAGCGTTAC CTGATAAAAT AAAAGTGATT
GGTTATAAAC AACATTTTAA AATTCCTTAT GTCGGTCTGA CAGGAAACTA
CCGTTACGAT AATTTTGAGT TTGGTGGTGC ATTTAAATAC AGCGGGTGGG
TCAGGGGATC TGATAATGAT GAGCATTATG TAAGACAAAC TACATTCCGA
AGCAAAGTAA TAAACCAGAA TTACTATTCT GTTGCAGTTA ATGCCGGTTA
CTATATTACC CCAGAGGCAA AAGTGTACAT CGAGGGTGTA TGGAGTCGTC
TCACAAATAA AAAAGGGGAT ACATCTCTTT ACGACCGTAG TGATAATACT
TCGGAGCATA ATAATAACGG GGCTGGAATT GAAAATTACA ACTTCATTAC
GACGGCCGGT CTGAAGTACA CGTTTTAAGG ATCC GCC

The start codon (located within the *Nde*I site) is shown in green and the stop codon is shown in red. The restriction enzyme recognition sites are underlined.

The full protein sequence of His-tagged OmpP is given below:

MHHHHHHENL YFQG↓SDFFGP EKISTEINLG TLSGKTKERV YEPEEGGRKV SQLDWKYSNA
AILKGAVNWE LNPWLSVGAA GWTTLNSRGG NMVDQDWMDS GTPGTWTDES RHPDTRLNYA
NEFDLNVKGW FLKESDYRLA IMAGYQESRY SFNATGGTYI YSENGGFRNE TGALPDKIKV
IGYKQHFKIP YVGLTGNYRY DNFEFGGAFK YSGWVRGSDN DEHYVRQTTF RSKVINQNYY
SVAVNAGYYI TPEAKVYIEG VWSRLTNKKG DTSLYDRSDN TSEHNNNGAG IENYNFITTA
GLKYTF

The His-tag is shown in blue, the TEV recognition site in purple, and the cut site marked by an arrow.

The full DNA sequence of the cloned untagged OmpT gene (plasmid pLMM03) is given below:

```
GGAGACACAT ATACATATG  TCTACCGAGA CTTTATCGTT  TACTCCTGAC AACATAAATG
CGGACATTAG TCTTGGAACT CTGAGCGGAA AAACAAAAGA GCGTGTTTAT CTAGCCGAAG
AAGGAGGCCG AAAAGTCAGT CAACTCGACT GGAAATTCAA TAACGCTGCA ATTATTAAAG
GTGCAATTAA TTGGGATTTG ATGCCCCAGA TATCTATCGG GGCTGCTGGC TGGACAACTC
TCGGCAGCCG AGGTGGCAAT ATGGTCGATC AGGACTGGAT GGATTCCAGT AACCCCGGAA
CCTGGACGGA TGAAAGTAGA CACCCTGATA CACAACTCAA TTATGCCAAC GAATTTGATC
TGAATATCAA AGGCTGGCTC CTCAACGAAC CCAATTACCG CCTGGGACTC ATGGCCGGAT
ATCAGGAAAG CCGTTATAGC TTTACAGCCA GAGGTGGTTC CTATATCTAC AGTTCTGAGG
AGGGATTCAG AGATGATATC GGCTCCTTCC CGAATGGAGA AAGAGCAATC GGCTACAAAC
AACGTTTTAA AATGCCCTAC ATTGGCTTGA CTGGAAGTTA TCGTTATGAA GATTTTGAAC
TCGGTGGCAC ATTTAAATAC AGCGGCTGGG TGGAATCATC TGATAACGAT GAACACTATG
ACCCGGGAAA AAGAATCACT TATCGCAGTA AGGTCAAAGA CCAAAATTAC TATTCTGTTG
CAGTCAATGC AGGTTATTAC GTCACACCTA ACGCAAAAGT TTATGTTGAA GGCGCATGGA
ATCGGGTTAC GAATAAAAAA GGTAATACTT CACTTTATGA TCACAATAAT AACACTTCAG
ACTACAGCAA AAATGGAGCA GGTATAGAAA ACTATAACTT CATCACTACT GCTGGTCTTA
AGTACACATT TTAAGGATCC GCTAACAAAG CCCGAAAGGA
```

The start codon (located within the *Nde*I site) is shown in green and the stop codon is shown in red. The restriction enzyme recognition sites are underlined.

The full protein sequence of untagged OmpT is given below:

```
MSTETLSFTP DNINADISLG TLSGKTKERV YLAEEGGRKV SQLDWKFNNA AIIKGAINWD
LMPQISIGAA GWTTLGSRGG NMVDQDWMDS SNPGTWTDES RHPDTQLNYA NEFDLNIKGW
LLNEPNYRLG LMAGYQESRY SFTARGGSYI YSSEEGFRDD IGSFPNGERA IGYKQRFKMP
YIGLTGSYRY EDFELGGTFK YSGWVESSDN DEHYDPGKRI TYRSKVKDQN YYSVAVNAGY
YVTPNAKVYV EGAWNRVTNK KGNTSLYDHN NNTSDYSKNG AGIENYNFIT TAGLKYTF
```

The full DNA sequence of the cloned untagged OmpP gene (plasmid pLMM04) is given below:

GGAGACACAT ATACATATGT CTGATTTCTT CGGCCCGGAG AAAATAAGTA
CAGAGATTAA TTTAGGTACA CTGAGTGGCA AAACAAAAGA ACGGGTTTAT
GAGCCTGAAG AAGGTGGACG TAAAGTCAGC CAACTGGACT GGAAATACAG
TAACGCCGCC ATTCTTAAAG GCGCCGTTAA CTGGGAGCTG AATCCATGGT
TATCTGTTGG TGCAGCTGGC TGGACCACTC TCAATAGTCG GGGGGGGAAT
ATGGTTGATC AGGACTGGAT GGATTCCGGG ACTCCCGGAA CATGGACAGA
TGAAAGCAGG CATCCTGATA CACGTCTTAA TTATGCCAAC GAATTTGATT
TGAACGTTAA AGGCTGGTTT TTAAAAGAAT CTGATTATCG CCTTGCTATT
ATGGCAGGTT ATCAGGAAAG CCGTTACAGT TTTAATGCTA CAGGAGGAAC
TTATATTTAT AGTGAGAATG GTGGTTTCCG GAATGAAACG GGAGCGTTAC
CTGATAAAAT AAAAGTGATT GGTTATAAAC AACATTTTAA AATTCCTTAT
GTCGGTCTGA CAGGAAACTA CCGTTACGAT AATTTTGAGT TTGGTGGTGC
ATTTAAATAC AGCGGGTGGG TCAGGGGATC TGATAATGAT GAGCATTATG
TAAGACAAAC TACATTCCGA AGCAAAGTAA TAAACCAGAA TTACTATTCT
GTTGCAGTTA ATGCCGGTTA CTATATTACC CCAGAGGCAA AAGTGTACAT
CGAGGGTGTA TGGAGTCGTC TCACAAATAA AAAAGGGGAT ACATCTCTTT
ACGACCGTAG TGATAATACT TCGGAGCATA ATAATAACGG GGCTGGAATT
GAAAATTACA ACTTCATTAC GACGGCCGGT CTGAAGTACA CGTTTTAAGG
ATCCGCTAAC AAAGCCCGAA AGGA

The start codon (located within the *Nde*I site) is shown in green and the stop codon is shown in red. The restriction enzyme recognition sites are underlined.

The full protein sequence of untagged OmpP is given below:

MSDFFGPEKI STEINLGTLS GKTKERVYEP EEGGRKVSQL DWKYSNAAIL KGAVNWELNP
WLSVGAAGWT TLNSRGGNMV DQDWMDSGTP GTWTDESRHP DTRLNYANEF DLNVKGWFLK
ESDYRLAIMA GYQESRYSFN ATGGTYIYSE NGGFRNETGA LPDKIKVIGY KQHFKIPYVG
LTGNYRYDNF EFGGAFKYSG WVRGSDNDEH YVRQTTFRSK VINQNYYSVA VNAGYYITPE
AKVYIEGVWS RLTNKKGDTS LYDRSDNTSE HNNNGAGIEN YNFITTAGLK YTF