The role of native outer membrane features in the folding and function of OMPs

Jonathan Mark Machin

School of Molecular and Cellular Biology

Astbury Centre for Structural Molecular Biology

University of Leeds

Submitted in accordance with the requirements of the degree of Doctor of Philosophy

September 2023

Intellectual property and publication statement

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

Chapter 3 contains data from one jointly authored publication:

Machin, J.M., Kalli, A.C., Ranson, N.A., Radford, S.E., Protein–lipid charge interactions control the folding of outer membrane proteins into asymmetric membranes. Nature Chemistry (2023). https://doi.org/10.1038/s41557-023-01319-6

The candidate performed all the research. All authors contributed to the experimental design, analysis of the data, and all authors wrote or edited the manuscript.

Chapter 4 contains data from collaborative work and preparation of a jointly authored manuscript regarding the FusA protein is ongoing. Protein purification and detergent refolding of FusA was carried out by Khedidja Mosbahi in the Walker group, University of Glasgow. All other data presented in Chapter 4 was performed by the candidate.

Acknowledgements

Firstly, I would like to thank my supervisors Professors Neil Ranson, Sheena Radford and Antreas Kalli for their continued support, encouragement, and guidance throughout this project and as I look to the future. I am particularly grateful for the freedom they gave me to develop and pursue my own ideas and inquiries, while always being available to offer suggestions and insights.

I am highly appreciative of the working context provided by the Radford-Brockwell, Ranson and Kalli lab groups. Members of the OMPire have been invariably helpful in supporting and teaching me about the field: Bob Schiffrin, Jim Horne, Sam Haysom, James Whitehouse, Kat Fenn, Izzie Boon, Joel Crossley, Anton Calabrese. In particular Sam Haysom, who first introduced me to BAM, Bob Schifrin and Jim Horne for their encyclopaedic knowledge of the field, James Whitehouse who was always willing to have very tangential discussions and Kat Fenn whose EM advice was always useful. More broadly across the research groups: Josh White, whose knowledge and experience of EM I regularly tapped into, Dheeraj Prakaash who introduced me to the world of MD, and Kyle Le Huray. Indeed, I would like to thank all members of these groups, whose assistance and discussions have been invaluable, as well as for their persistence in making me sociable.

I must thank our collaborators at the University of Glasgow, particularly Khedidja Mosbahi and Daniel Walker for providing samples. The EM facility staff have been unwaveringly helpful in all the microscopy endeavours (and failures) throughout the project. I am grateful to the Wellcome Trust for sponsoring the PhD program and this project, and for the other students on the program: Alex Flynn, Emma Ryder, Linda Makhlouf, Upasana Sykora and Ryan Cocking. Special thanks to Ryan for being a fantastic housemate for the last three years. Finally, I am indebted to my family and friends, in Leeds, the UK and across the world, whose kindness and support have been essential to my PhD’s success.

Abstract

The outer membrane (OM) of diderm bacteria is a unique membrane, characterised by the asymmetric (outer leaflet) presence of lipopolysaccharide (LPS), a high outer membrane protein (OMP) concentration dominated by a few protein species, protein-lipid phase separation and a low lateral diffusion rate. Understanding the OM is critical in tackling the challenges and opportunities posed by antibiotic resistance and developing biotechnologies. However, much remains unclear about how the organisation, composition and interactions between components modulate the OM. Here, native features of the OM are recapitulated *in vitro* and their consequences for OMP folding and function characterised. Generation of charge asymmetric liposomes mimicking the OM's bilayer dipole revealed that the insertion and folding of OmpA and BamA is accelerated compared to symmetric liposomes (and impeded in liposomes with an inverted charge dipole). Further, a conserved patch of extracellular positive charge in OMPs is identified and shown for OmpA to play a critical role in folding. The LPS dependence of activity of the protease OmpT is well-known, and here it is further found that the activation is specific to precise regions of the LPS oligosaccharide. It is also shown that the OMP FusA strictly requires LPS to enable its structural solution by cryoEM which, combined with MD, reveals FusA:LPS binding sites. OMP-OMP interactions are poorly understood, but here it is found that the natively abundant OMP OmpA can specifically interact with and enhance OmpT and inhibit PagP (an OMP palmitoyltransferase) activity. An *in silico* screen of OMP-OMP interactions further identifies a broad diversity of interactions. Together, these results yield new insights into the mechanisms that drive OMP assembly and function, both *in vitro* and *in vivo*, pointing to the underappreciated role of native OM asymmetry, and LPS-OMP and OMP-OMP interactions to modulate the OM's form and function.

Table of Contents

List of Figures

List of Tables

Abbreviations

Lipids

Phospholipid names are formed by two-letter designation of the acyl chain identity followed by a two-letter designation of the head group.

Acyl chains:

DM Di-Myristoyl

PO 1-palmitoyl-2-oleoyl

DL Di-Lauroyl

DU Di-Undecanoyl

Head groups:

PS PhosphatidylSerine

PE PhosphatidylEthanolamine

PC PhosphatidylCholine

PG PhosphatidylGlycerol

Other lipids:

ECPL E. coli polar lipids

LPS Lipopolysaccharide

LOS Lipooligosaccharide

Other Abbreviations

ABC ATP-Binding Cassette

AFM Atomic Force Spectroscopy

AF2 Alphafold2

AMP Antimicrobial Peptides

ANS 8-Anilinonaphthalene-1-sulfonic acid

ATP Adenosine TriPhosphate

BAM	β-barrel Assembly Machinery
BSA	Bovine Serum Albumin
CD	Cyclodextrin
CDC	Centre of Disease Control
(CG)-MD	(Coarse grained)-Molecular Dynamics
Cpx	Conjugative Pilus eXpression
CTF	Contrast Transfer Function
DDM	n-dodecyl-ß-D-maltoside
DLS	Dynamic Light Scattering
EDTA	Ethylenediaminetetraacetic acid
EM	Electron Microscopy
EPR	Electron Paramagnetic Resonance
FFA	Free Fatty Acids
Gal	D-Galactose
Glc	D-Glucose
GlcN	D-glucosamine
Hep	L-glycero-D-manno-heptose
IM	Inner Membrane
Kdo	2-keto-3-deoxyoctulosonate
LDAO	Lauryldimethylamine oxide
Lol	Localisation Of Lipoproteins
LPR	Lipid:Protein Ratio
MAE	Mean Average Error
MβCD	Methyl-β-Cyclodextrin
MCE	Mammalian Cell Entry
MD	Molecular dynamics

Mla	Maintenance of Lipid Asymmetry
MS	Mass Spectrometry
MSD	Mean Squared Diffusion
NMR	Nuclear Magnetic Resonance
OA	Oleic Acid
OM	Outer Membrane
OMP	Outer Membrane Protein
OMV	Outer Membrane Vesicles
PAE	Predicted Alignment Error
PCR	Polymerase Chain Reaction
PDB	Protein Data Bank
PEP	Patch of External Positive residues
pHLIP	pH low insertion peptide
PIP2	Phosphatidylinositol 4,5-bisphosphate
pLDDT	predicted Local Distance Difference Test
PMF	Proton Motive Force
pNP	palmitate-4-nitrophenol
Rcs	Regulator of Capsule Synthesis
SDS-PAGE	Sodium Dodecyl Sulphate – PolyAcrylamide Gel Electrophoresis
SNR	Signal:Noise Ratio
TBDT	TonB Dependent Transporter
T_m	Transition temperature
TLC	Thin Layer Chromatography
TM	Transmembrane

Chapter 1: Introduction

'Thus the duty of the man who investigates the writings of scientists, if learning the truth is his goal, is to make himself an enemy of all that he reads, and, applying his mind to the core and margins of its content, attack it from every side.'

Ibn Al-Haytham, c1000

'By this means the Heavens are open'd and a vast number of new Stars and new Motions, and new Productions appear in them, to which all the ancient Astronomers were utterly strangers. By this the Earth itself, which lyes so neer to us, under our feet, shews quite a new thing to us, and in every little particle of its matter, we now behold almost as great a variety of Creatures, as we were able before to reckon up in the whole Universe itself.'

Robert Hooke, 1665

'*Nature to be commanded must be obeyed; and that which in contemplation is as the cause is in operation as the rule*'[1] mused Francis Bacon at the dawn of the European scientific revolution in his landmark *Novum Organum* as he reflected on the scientific method and its outworking. Five centuries later, vast amounts of knowledge have been generated and yet the goal of using nature's rules for holistic manipulation and control of intricately complex biological systems remains distant. Yet never has this been more urgent as the world grapples with the increasing severe challenges arising from a globally connected, industrial world: climate change, pandemics, antibiotic resistance. Biological membranes exist as liminal spaces, suspended between the abiotic environment and the life they are both a part of and encapsulate, and these thresholds thus offer vast potential to externally direct and impact that life, allowing for diverse applications including healthcare and biotechnology. Following a historical contextualisation and overview of the important features of general membrane biology, this introduction will focus on the nature and properties of the outer membrane of diderm bacteria, which the results presented centre on.

1.1 Historical context

In the midst of the European scientific revolution, and building on pioneering work from Galileo, Robert Hooke and Antonie Van Leeuwenhoek developed their microscopes and revealed the first hints of the unseen complexity of life. In 1665 when Hooke presented his *Micrographia*, he described the microscopic ultrastructure of a slice of cork as "much like a [h]oney-comb, but that the pores of it were not regular…in that these pores, or cells, were not very deep, but consisted of a great many little boxes"[2] (**Fig. 1.1a**). The defining observable feature of these 'cells' was their boundary with the extracellular space and although Hooke did not understand the nature and role of the observable structures, he intuitively grasped the necessity of communication and interaction between the enclosed 'cells' and their environment, spending extensive efforts searching for "a passage out of one of those cavities into another"[2]. Antonie van Leeuwenhoek, credited with the first unambiguous description of both aerobic and anaerobic bacteria (**Fig. 1.1b**, Fig. IV), further showed the flexibility of the cell boundary by observing single celled organisms flex and move (**Fig. 1.1b**, Fig. III)[3].

Despite this transformational early foundation, significant progress on studying and interpreting cellular boundaries did not occur until after the proposition of cell theory by Schleiden and Schawann in 1838-1839[4,5]. Although cell theory does not explicitly speak into the nature of cellular boundaries, it provides a framework on which a theory of membranes can be constructed. Indeed, in the years following, by drawing analogies to synthetic membranes Moritz Traube suggested that all cells are surrounded by semi-permeable membranes, that controllably allow ingress and egress of substances[6], thus defining the cell membrane, at its simplest conception, in 'modern' terms. Wilhelm Pfeffer's work further defined the membrane from a generic partially permeable membrane to an osmotically active, distinct part of the cell[7]. In other words, a precision biostructure. Not until the early 20th century were lipids discovered as a key membrane constituent (based on the differential effectiveness of various anaesthetics of known chemical properties[8]) and seminal work on red blood cells in 1925 identified the bilayer architecture of the membrane[9,10], although this remained controversial[11].

These insights were remarkable given that visualisations of biological membranes were not possible until the development of X-ray diffraction (high resolution but indirect data) and biological electron microscopy (EM) (artefactually stained but high resolution) in the early/mid-1900's, with the cell wall and cell

membrane typically being conflated prior to these methodologies. Even at that time it was well established that different membranes contain a huge diversity of components[12,13] and early EM images showed various architectures[14,15]. Given this, proposals of a basic, conserved architecture across all cell membranes were controversial[11], and in 1957 it remained unclear whether cells had sharp or diffuse membrane boundaries[16]. However, in the 1950-60's the Danielli-Davson model, first proposed in 1935[17], dominated thinking about membrane biology. This trilamellar membrane proposed an unstable central bilayer of phospholipids containing a central layer of lipophilic substances, scaffolded and stabilised by electrostatic interactions in a layer of proteins on either side of the bilayer (**Fig. 1.1c**). An alternative, less popular model, was the Benson-Green lipoprotein unit model[18,19], where a monolayer of lipoprotein units bound lipids and there was no continuous bilayer (**Fig. 1.1d**).

Solving the first atomic resolution protein structure, myoglobin, was transformational in how biomolecular interactions were envisaged, with myoglobin's interior being dominated with hydrophobic interactions[20]. Noting the importance of the hydrophobic effect for protein (and DNA[21]) stability[22], Singer and Nicolson developed the fluid mosaic model in 1972[23], proposing that membranes too are dominated by hydrophobic interactions, as well as emphasising the 3D asymmetry of the membrane (**Fig. 1.1e**). Although new data have rendered certain features out of date (see Section 1.2), discussions of membranes still tend to centre around this model and it retains enormous explanatory power for the interpretation of membrane properties. Three years later the first transmembrane protein structure (bacteriorhodopsin, **Fig. 1.1f**), solved *in situ* in its membrane via 2D electron crystallography[24], demonstrated many of the elements of the fluid mosaic model, including the membrane-spanning, amphiphilic nature of transmembrane proteins.

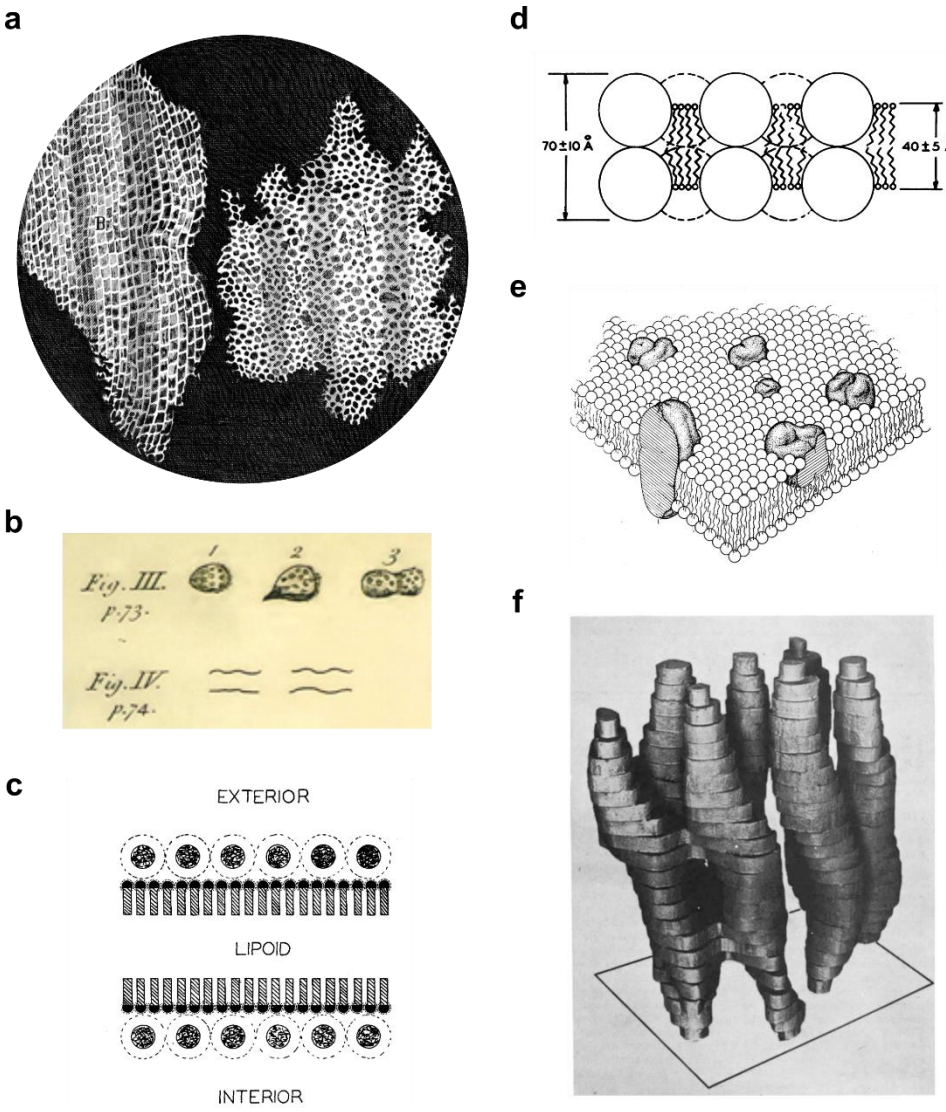


Figure 1.1: A historical perspective on cellular membranes. (a) Robert Hooke's drawing of cells observed in a slice of cork, the first publication of an observation of cells. (Plate XI from [2]). **(b)** Antonie van Leuwenhoek's sketch of (III) cells moving and possibly dividing and (IV) (probably) spiral bacterial cells (from [3]). **(c)** The Danielli-Davson model as originally proposed. Circles represent a protein layer decorating the bilayer surface (from [17]). **(d)** The Benson-Green model of cell membranes as originally proposed. Circles represent the unit lipoproteins proposed to hold the membrane together (from [19]). **(e)** Singer and Nicolson's iconic original visualisation of the fluid mosaic model (from [23]). **(f)** 3D model of the structure of bacteriorhodopsin as solved by Unwin and Hendersen via 2D electron crystallography (from [24]).

1.2 Universal features of membrane architecture

The fluid mosaic model[23] (**Fig 1.1e**) has been the dominant paradigm for understanding membrane architecture since its introduction 50 years ago, although new data have required some modifications. The model's enduring success rests upon its reliance on general, unchanging biophysical properties, notably thermodynamic considerations, lipid dynamics and protein-lipid interactions[25], and its continual explanatory power for nanoscale membrane features (e.g. [26]). Indeed, given the complexity in composition and organisation of biological membranes, with over 48, 000 unique lipid species identified to date[27] (compared to <100 in 'alphabets' for proteins and oligonucleotides), the similarity between different membranes, and particularly the consistency of its dimensions, remains surprising[28].

As proposed by Singer and Nicolson, the forces holding biological membranes together are dominated by non-covalent interactions, in particular van der Waals forces and the hydrophobic effect[21]. Amphipathic, transversely asymmetric lipid bilayers typically generate a continuous matrix into which proteins are asymmetrically intercalated with hydrophobic molecules trapped in the core of the bilayer, including the lipid-anchors of lipoproteins[25,29]. The different components will minimise their free energy via optimal interactions, for example protein-lipid hydrophobic matching and lipid phase separation, allowing manipulation of local membrane properties and the formation of distinct membrane domains via molecular sorting[30]. Lateral mobility of components in membranes varies from effectively immobilised to rapidly diffusing (measured at 0.001-0.4 $\mu m^2/s$ for bacterial transmembrane proteins[31–33]), depending on the component size, local membrane and environmental context[34,35]. Integral membrane protein (and occasionally lipid[36]) mobility may be restrained by additional interactions with membrane associated proteins and thus adjust membrane properties (most notably by the cytoskeleton and extracellular matrix or capsular features)[37,38]. Membranes contain supramolecular features, with specialised membrane domains being important for optimal lipid-protein interactions, cell signalling via protein arrays[39] and processes requiring mechanical membrane manipulation like exo/endocytosis[40]. This hierarchical organisation of the membrane, and its dynamic rearrangement, is crucial for cellular health and survival. Thus, while the modified fluid mosaic model remains a good descriptor of the current understanding membrane properties, biological membranes are typically more structurally mosaic and less fluid than originally envisioned[41] (**Fig. 1.2**).

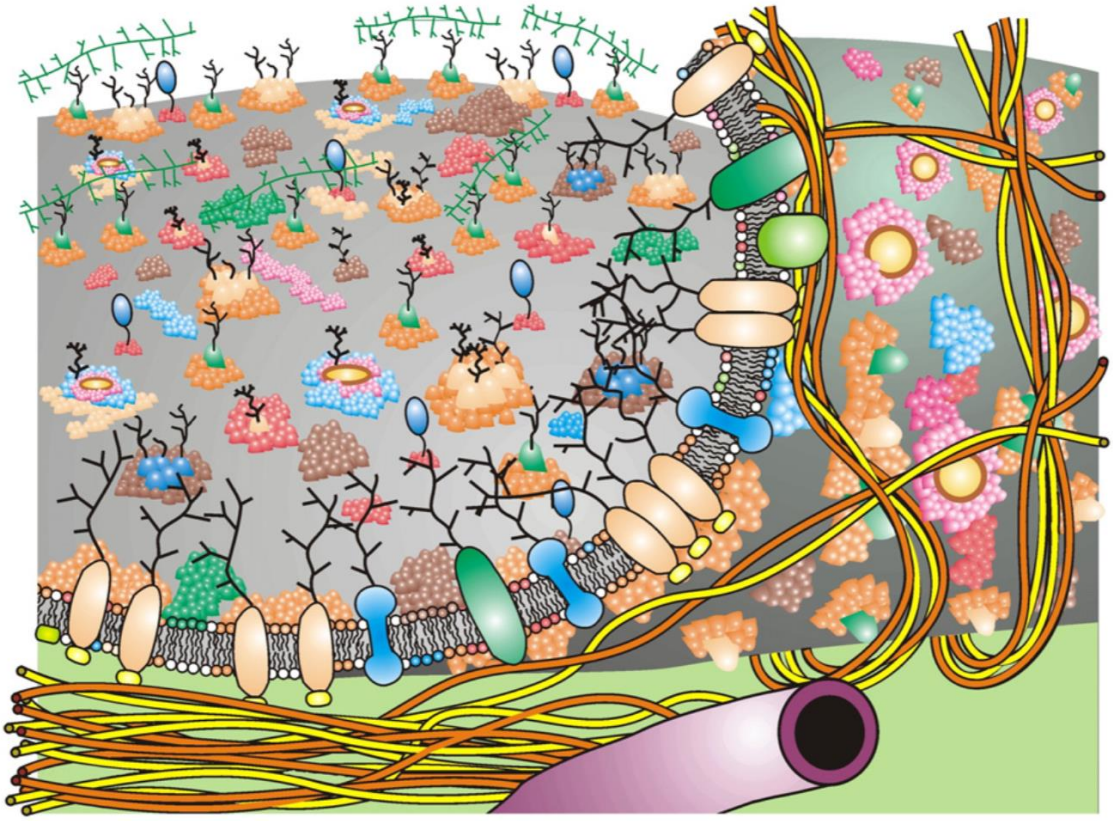


Figure 1.2: An updated model of the fluid mosaic model highlighting new discoveries around membrane domains and the communication between transmembrane and soluble cellular protein components. The cytoskeleton (orange/yellow fibres) restrict the lateral diffusion of some membrane components. Note that size, crowding, and complexity are not accurately represented here. (From [29]).

Proteins topologically linked to membranes may be integral (either fully transmembrane or, more rarely, integrated into a single leaflet of the bilayer), membrane retained by lipid anchors (lipoproteins) or only transiently associated (**Fig. 1.3**). Transmembrane protein sequence and structure is constrained by the necessity of passing through the membrane's hydrophobic core[42], with transmembrane domains adopting either an α-helical bundle or β-barrel structure (**Fig. 1.3a-c**), with the latter being restricted to the topologically outer bilayer of dual-membraned systems[43] (e.g., the outer membranes of diderm bacteria, mitochondria and chloroplasts). Depending on their function, transmembrane proteins may also have extensive soluble domains. *In vivo*, dedicated structures exist to support accurate folding of these proteins into the membrane: in *E. coli*, the SecYEG complex for inserting α-helical proteins (into the inner membrane)[44] and the BAM complex for inserting β-barrels (into the outer

membrane)[45], both of which have homologs throughout life[45,46]. The SecYEG complex is highly conserved[47] and is capable of both co-/post-translationally inserting proteins into membranes and transporting unfolded proteins through the membrane[48]. The major structural component, SecY, is a pseudo-symmetric 10-helix bundle, which forms the protein pore[49] (SecE is essential and while SecG is not strictly required it significantly enhances translocation throughput[50]). Proteins are targeted to SecYEG via a predominantly hydrophobic N-terminal signal peptide, which is recognised by cytoplasmic ATPase SecA, facilitating targeting to the SecYEG complex and providing energy for the translocation/insertion process[51].

Lipoproteins are a diverse set of proteins (e.g. **Fig. 1.3e**), with functions ranging from nutrient acquisition and mechanical roles[52], that have an acylated tail that anchors them in the membrane[52]. Similar to the transmembrane folding complexes, cellular machinery associated with lipoprotein generation and insertion is well, although not absolutely, conserved[53]. A broad range of proteins are known to transiently interact with the membrane as part of their functional cycle, including inducing vesicle formation[54], cell signalling[55] and in regulatory pathways[55] (e.g. **Fig. 1.3f**). While these proteins require a lipid interacting motif they are, along with lipoproteins, less structurally and evolutionarily constrained than transmembrane proteins, allowing them to perform otherwise challenging functions at the membrane, and as such they are frequently found as components in complex, membrane-linked cellular machinery (e.g. [56,57]).

The complexity of the membrane is formed by its varying components and their differential organisation and interactions, and specific features are highlighted below, including bilayer lipid asymmetry, protein-lipid interactions and protein-protein interactions.

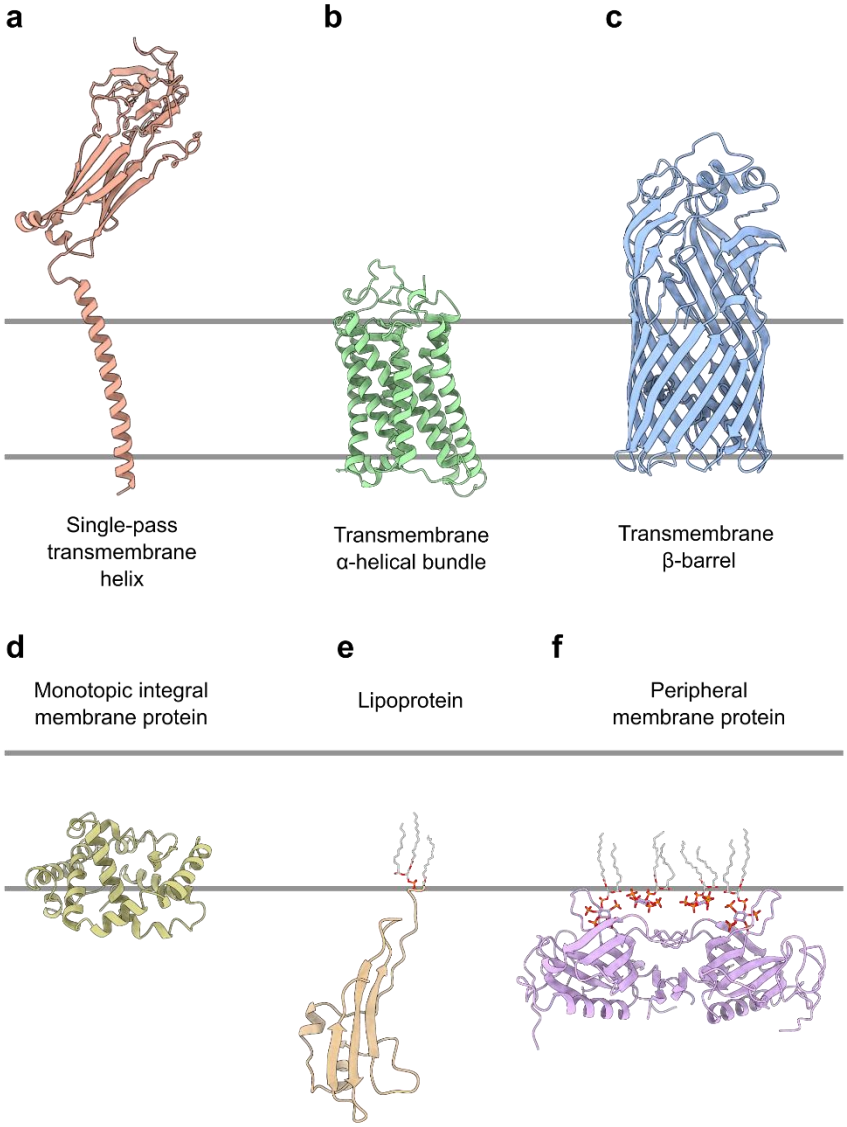


Figure 1.3: An overview of transmembrane and membrane associated proteins. Transmembrane proteins may be **(a)** monotopic (single-pass) α-helical (e.g., immune system TCR), **(b)** polytopic α-helical bundles (e.g., bacteriorhodopsin) or **(c)** transmembrane β-barrels (FadL). **(d)** A few proteins are characterised as partially inserted into the bilayer but not fully transmembrane (e.g., MlaA). **(e)** Lipoproteins are anchored in the bilayer by their acyl chains, but the proteins remain soluble (e.g., BamE). **(f)** Peripheral membrane proteins stably or transiently interact with lipids (or proteins, not shown) to associate to the membrane (e.g., PH domain, from Bruton's tyrosine kinase). (PDBs: 8ES9[58], 1DZE[59], 1T1L[60], 5NUP[61], 5LJO[56], 1BTK[62], respectively).

1.2.1 Membrane lipid asymmetry

All characterised biological membranes have transverse lipid asymmetry (i.e., differential lipid compositions in each membrane leaflet), typically consisting of lipid headgroup differences but also composed of differential acyl-chain identity, and in some cases lipid organisation[63]. (The orientation of transmembrane proteins and the different complement of proteins interacting on the membrane surface also confer asymmetry). Cells expend considerable resources, both in protein/lipid production and energy, to ensure that membranes retain their lipid asymmetry, with active processes ensuring lipids are organised[64,65], and in some cases explicitly inserted[66], into the correct leaflet. Asymmetry dysregulation is associated with multiple diseases, underlining its cellular importance[67]. Biology uses membrane asymmetry for a broad range of functions, including cell signalling[68], control of protein-lipid interactions[69], transmembrane-protein activation[70] and cell morphology[71] and it in turn can be clinically exploited[72] (**Fig. 1.4**).

Although it remains challenging to study membrane asymmetry and especially its implications on the folding and function of proteins (due to difficulties in controllably generating large quantities of asymmetric membranes), significant progress has been made in recent years[73,74]. In particular, a range of methods have been developed and/or expanded to synthesise asymmetric membranes in free-floating liposomes (**Fig. 1.5**) (recently reviewed in [75]). Predominant among these methods are those that involve lipid exchange over the external, solvent exposed leaflet of the liposome, while the inner leaflet, occluded from the bulk solution, remains unchanged. Commonly the exchange process is mediated by cyclodextrins (CDs), cyclic oligosaccharides typically 6, 7 or 8 (termed α, β or γ respectively) sugars long, which can bind and solubilise lipid acyl chains[76]. Historically used at relatively low concentrations to exchange cholesterol in and out of membranes[77,78], at higher concentrations CDs can mediate highly efficient exchange (>90 %) between donor and acceptor liposomes until an equilibrium is reached[79] (**Fig. 1.5e**). Alternatively, CD-lipid complexes can be prepared and mixed with intact acceptor liposomes and an equilibrium formed between the liposome and CD-lipid complex pool[80] (**Fig. 1.5f**). By altering the concentrations of donor and acceptor lipids/liposomes it is possible to controllably generate partial asymmetry[81]. These approaches have been particularly useful in studying how lipid asymmetry modulates membrane properties like trans-bilayer coupling[82], lipid raft formation[83] and membrane curvature and tension[84].

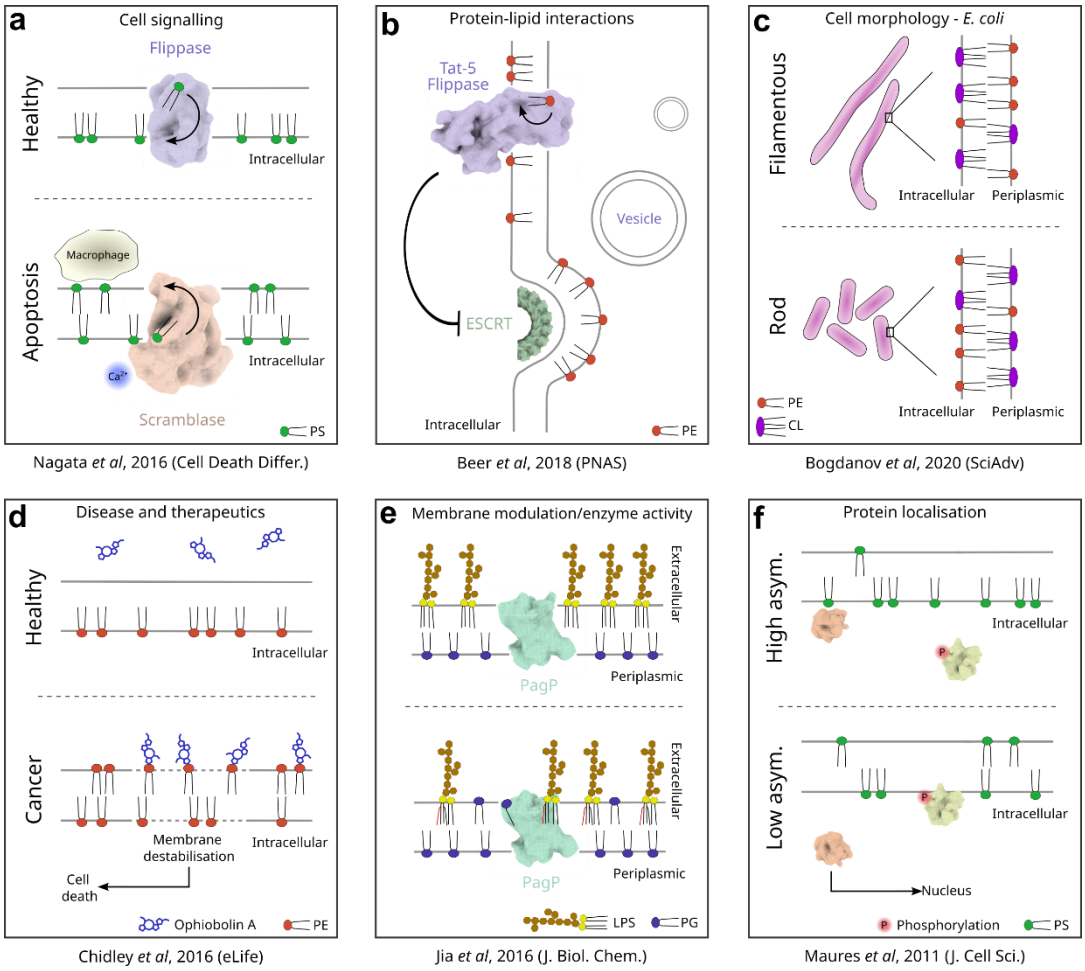


Figure 1.4: Membrane asymmetry has diverse functions. **(a)** The localisation of phosphatidylserine (PS) lipids indicate cellular health in higher eukaryotes, with excess PS in the outer cellular leaflet signalling defects and targeting the cell for apoptosis. **(b)** Controlling protein-lipid interactions by lipid localisation can regulate other biological processes like vesicles budding from membranes, which is inhibited by the presence of phosphatidylethanolamine (PE) in the interacting bilayer leaflet. **(c)** Organisation of lipids helps control the morphology of *E. coli* with excess cardiolipin in the inner leaflet of the inner membrane typical in filamentous bacteria and excess cardiolipin in the outer leaflet in rod-shaped bacteria. **(d)** Membrane asymmetry dysregulation in disease can be therapeutically exploited by designing molecules to target specifically to the dysregulated membranes, here Ophiobolin A is targeted to cancer cells due to an abnormally high concentration of PE lipid in their outer leaflet. **(e)** Lipid localisation can control membrane protein activity, here by restricting available substrate for the enzyme PagP until the inner leaflet phospholipids mislocalise to the outer leaflet. **(f)** The degree of lipid asymmetry, and the induced charge asymmetry, may regulate the association of proteins to the membrane by altering the membrane's electrostatic profile, here shown coupled to transcriptional control.

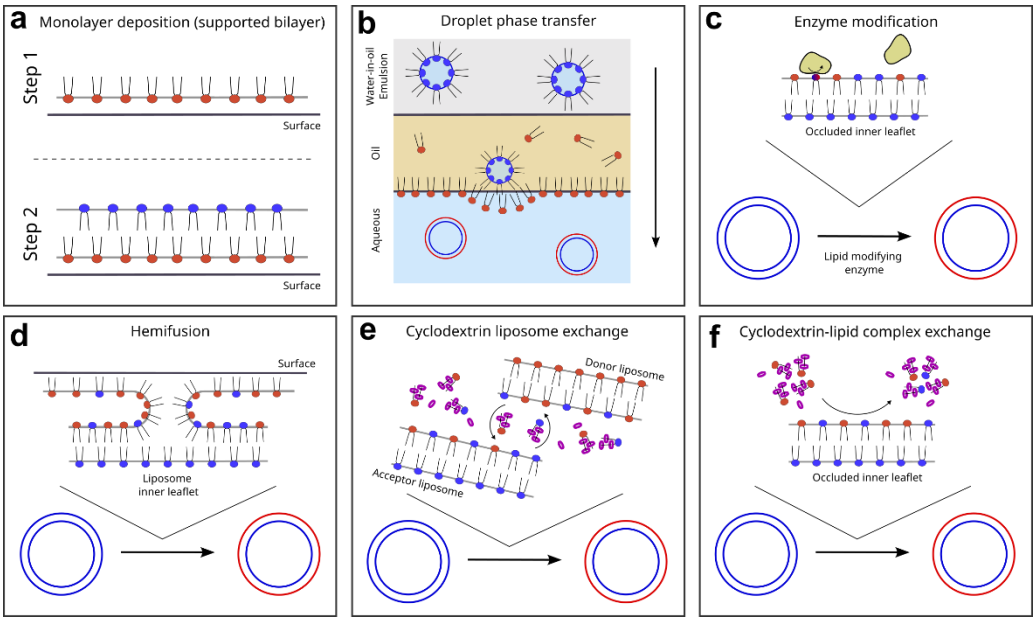


Figure 1.5: Multiple methods to generate asymmetric bilayers *in vitro* have been developed. (a) Supported lipid bilayers can be made asymmetric by separate deposition of each bilayer leaflet. **(b)** Similarly, phase transfer methodology generates asymmetric membranes by assembling the membrane leaflets separately by exploiting the amphipathic nature of lipids to align at the interface between a polar and non-polar solvent. **(c)** Lipids in the accessible surface of liposomes can be modified *in situ* by enzymatic reaction to change their identity, rendering the liposomes asymmetric. **(d)** Hemifusion can be induced between liposomes and surface deposited monolayers, facilitating exchange between the surface layer and the outer liposome leaflet. Cyclodextrins can be used to shuttle lipids either **(e)** between bilayers or **(f)** between a bilayer and a soluble cyclodextrin-lipid complex pool to reach an equilibrium between the accessible lipid surfaces and cyclodextrin lipid complexes. Inner liposome leaflet is occluded and so not exchanged.

Despite its ubiquity, less is known about how membrane asymmetry alters transmembrane protein folding and function. Perfringolysin O's rate of membrane insertion and intermediate conformation formation was found to be affected by lipid charge asymmetry[85]. The model pH low insertion peptide's (pHLIP) membrane insertion rate has also been shown to be modulated by small lipid charge asymmetries over the bilayer, with the effects being assigned to protein-lipid charge mediated interactions[86]. However, both these proteins have 'native' soluble forms, making it hard to generalise these findings. More broadly, interactions between membrane and lipid charge can modulate insertion orientation into membranes at SecYEG, and thus presumably altering the membrane charge distribution can also change the protein

topology[87]. A recent study found that differential leaflet tension can alter the function of transmembrane protein phospholipase OmpLA, with the enzyme being less active under high trans-leaflet tension in asymmetric membranes than in an equivalent symmetric liposome[88]. Much remains to be determined about the implications of membrane asymmetry on membrane protein function, but it is already clear that manipulating asymmetry can impact both folding and function of proteins.

1.2.2 Protein-lipid interactions

While the lipid environment enforces challenges for the folding and functioning of membrane proteins, it also provides a microenvironment rich in modulatory potential, with specific lipids, general membrane properties and solute/field gradients over the membrane all exploited for regulatory and functional roles[26] (**Fig. 1.6**). Underlying biophysical characteristics of membranes are used to control membrane organisation, for example protein sorting via hydrophobic mismatch[89], both laterally in a membrane and to control protein trafficking between organelles[90] (hydrophobic mismatch has also been implicated in reducing the activity of proteins localising to the wrong membrane compartment[91]). Membrane fluidity can modulate function directly by altering diffusion dynamics[92] and is tightly regulated via the mechanistically obscure homeoviscous adaptation response[93]. External mechanical stimuli that alter membrane properties are important for a range of functionalities, notably mechanosensing[94,95] and osmoregulation where lipids help transmit mechanical force to the protein, inducing conformational changes[96]. Intriguingly, in specialised membranes conformational changes in proteins are able to alter global membrane tension significantly[97]. Some lipids can spontaneously phase-separate in membranes to form lipid rafts (observable by AFM[98]) which, despite early controversy, have emerged as a key feature to modulate membrane and membrane protein functionality[30], including the local alteration of membrane properties like fluidity[99], partial sequestration of specific lipids from the bulk membrane (e.g. caveolae)[100], and lateral sorting of proteins (e.g. signalling arrays)[101]. The heterogeneous, mosaic membrane that results can be important for mediating optimal function[102]. Indeed, as noted above, a study generating lipid-asymmetric liposomes inducing differential membrane tensions without altering lipid identity demonstrated a clear functional modulation[88].

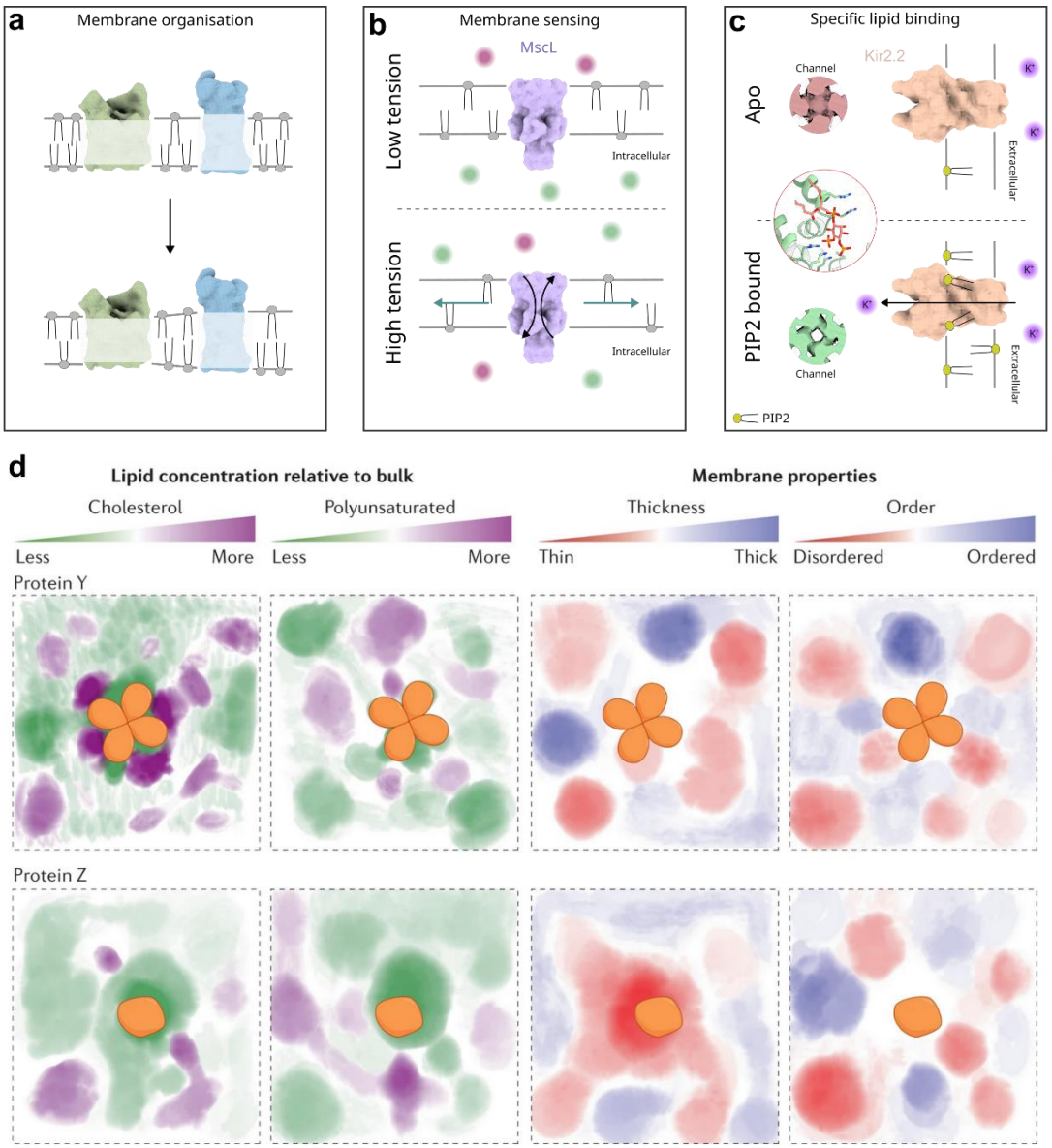


Figure 1.6: Protein lipid interactions mediate functional and regulatory roles. (a) Hydrophobic mismatch between proteins and lipids in the bilayer can lead to component organisation and membrane raft formation. **(b)** Membrane proteins can sense general membrane properties to modulate their activity, such as MscL which opens its ion channel only under high membrane tension. **(c)** Specific lipid binding can regulate membrane protein activity, here the potassium channel Kir2.2 is opened in the presence of PIP_2. (Inset PDB: 3SPI[104]) **(d)** Unlike the traditional understanding of the fluid mosaic model, most biological membranes are thought to have extreme local heterogeneities in terms of both membrane constituents and membrane properties, represented here for two potential proteins Y and Z, with the precise lipid microenvironment being important for proper regulation. (Panel (d) adapted from [25]).

In addition to protein modulation by general membrane properties, stable lipid binding via specific binding or non-specific annular lipid shells can manipulate the protein's function. Specific lipids can act like ligands to up- or down-regulate activity (e.g., PIP_2 and cholesterol interactions with Kir2.2 channels[103,104] and lipopolysaccharide (LPS) activation of the protease OmpT[105]). Lipids like PIP_2 are well-suited to this role as their size, stereochemistry and negative charge make the lipid easy to recognise and specifically bind, while the tight regulation of their cellular localisation and lateral membrane distribution ensures control[26]. While the majority of protein interacting lipids are able to exchange rapidly with the bulk membrane and are thus not strictly bound, distinct behaviour is observed for lipids adjacent and distant from proteins indicating a tight interplay between the membrane components[106]. Together, the effects on proteins of lipid induced membrane organisation, forces, and conformational changes can broadly modulate biochemical networks.

1.2.3 Protein-protein interactions

Studying protein-protein interactions and their assembly into oligomers within the context of the membrane is challenging, largely because traditional approaches (e.g., affinity pulldowns, native MS) struggle with the hydrophobic nature of membrane proteins, purification procedures removing lipids required for complex formation and the often relatively weak and/or transient nature of these interactions[107]. Interactions between single-pass TM helices are best understood, and are frequently involved in signal transduction via induced dimerisation[108]. It has been shown that membrane proteins can sense each other in a membrane up to 10 nm apart[109], allowing longer-range interactions to initiate organisation. While in some cases sequence motifs for TM helix association have been identified (e.g. GxxxG[108]) and helix-helix interactions are enriched in glycine and alanine[110], many interactions do not contain any identifiable pattern.

Protein-protein interactions in the membrane may be driven by their membrane context, for example where there is a hydrophobic mismatch between proteins and lipids it can become energetically favourable for proteins to cluster and minimise lipid interactions, thought to be one of the driving forces for the formation of signalling arrays[39,89] (**Fig. 1.7a**). Similarly, lipid raft formation can also drive protein organisation by generating local environments to either minimise the energy of different proteins or enrich proteins via specific lipid binding, for example PIP_2 enriched domains are known to be critical for clustering of the SNAP receptor

syntaxin-1A, required for exocytosis at the synapse[111] (**Fig. 1.7b**). Aquaporins, ubiquitous homo-tetrameric proteins, organise into large supramolecular assemblies in the membrane[112]. In humans, aquaporin assembly formation is dependent on the presence of alternatively spliced isoforms, which allow differential interactions and by balancing the relative concentrations, control of array size[113]; thus indicating that only small differences in an otherwise conserved sequence and fold are sufficient to substantially alter membrane protein behaviour and interactions. Different proteins often specifically interact to form functional complexes, for example in the respiratory supercomplex[114] and sugar permeation/utilisation in *E. coli*[115], although underlying mechanisms of complex formation are only partially understood[116].

In addition to internal membrane interactions, clustering of transmembrane proteins is often driven by interactions of their soluble domains, a particularly important feature of many signalling pathways (e.g. receptor tyrosine kinases[117], **Fig. 1.7c**). Intriguingly, the association of the soluble, disordered domains of transmembrane proteins are able to form protein liquid phases on the membrane, manipulating membrane rigidity and curvature[118], and suggesting additional approaches biology has exploited to mediate and control membrane processes.

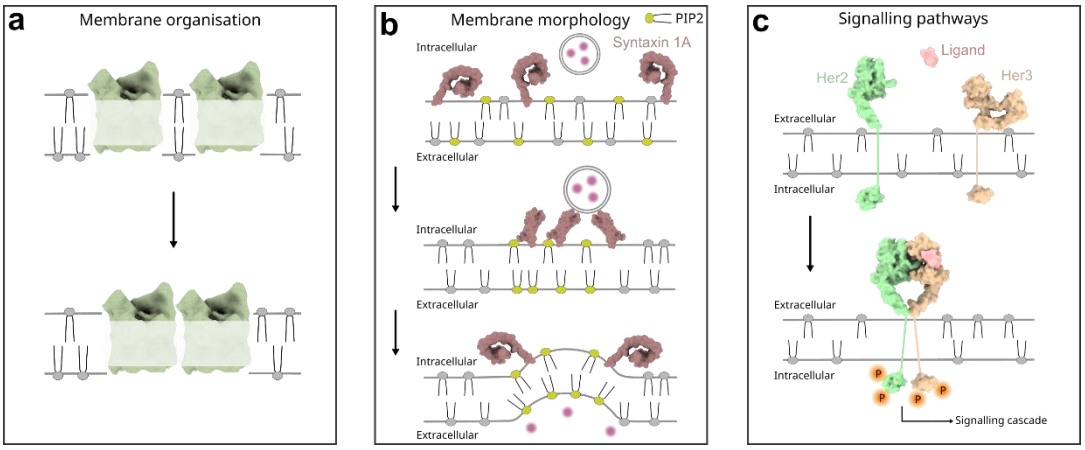


Figure 1.7: Protein lipid interactions mediate functional and regulatory roles. **(a)** Hydrophobic mismatch between lipid and protein components in the membrane can drive protein-protein interactions to minimise unfavourable protein-lipid interactions. **(b)** Protein clustering at lipid rafts, here syntaxin 1A at PIP_2 enriched lipid domains, can drive membrane morphological changes like exocytosis. **(c)** Membrane protein dimerisation via interaction of soluble domains is a key feature of many signalling pathways, for example receptor tyrosine kinases like Her2-Her3.

Thus, biological membrane composition and organisation allows membranes to perform, typically in parallel, their major biological roles of compartmentalisation, gradient generation, signal transduction and substance translocation, with different membrane components interacting and modulating each other to achieve optimal fitness and control.

1.3 Bacterial cell membranes

Since its invention in 1884, the gram stain has formed the basis of classifying and understanding bacteria by dividing them based on stain retention, gram-negative and gram-positive[119], although it took 80 years to understand the bacterial structural differences underlying the distinction[120]. Traditional characterisations of gram-negative bacteria have a cell envelope with a dual membrane architecture encasing a thin peptidoglycan layer, and gram-positive bacteria have a single cell membrane and a thicker outer layer of peptidoglycan[120]. However, while in most cases these distinctions remain accurate, they oversimply significant diversity in bacterial membrane architectures, with known counter-examples[121] (e.g. the phyla *Tenericutes* stains gram-negative, but lacks both peptidoglycan and a second, outer membrane[122]). Modern nomenclature conventions instead refer to monoderm and diderm bacteria, corresponding to those with a single or dual membrane architecture (if known), regardless of lipid composition or gram-stain response[121] (**Fig. 1.8**).

Bacterial phyla are dominated by diderms containing LPS in the outer leaflets of their outer membrane (herein referred to as LPS-diderms)[123] although sequence data indicate that about 25% of known diderm bacterial phyla do not contain LPS biosynthesis genes[124], and some other membrane architectures are identified, notably the myolate-diderm of *Mycobacterium*, where the OM consists of mycolic acids[125]. Unlike diderms, which are spread over >25 characterised phyla, only two well-characterised phyla contain the vast majority of known monoderm bacteria: *Acinetobacteria* and *Firmicutes*. Other, less defined cell envelope patterns are known, for example the diderm-phyla *Mollicutes* which lack a peptidoglycan layer[126]. Both monoderm and diderm bacteria can build an additional proteinaceous S-layer and/or oligosaccharide capsule around the cell[127,128]. Current conceptions of bacterial evolution hold that the last bacterial common ancestor was an LPS-diderm, with monoderms arising multiple times via outer membrane loss[123,129]. Particularly compelling evidence for an original diderm is found via phylogenetic analysis of the predominantly monoderm phyla *Firmicutes* with characterisations and comparisons

to its extant diderm species[47], although this has recently been challenged via a combined approach of phylogenetics and phenotypical characterisation[130].

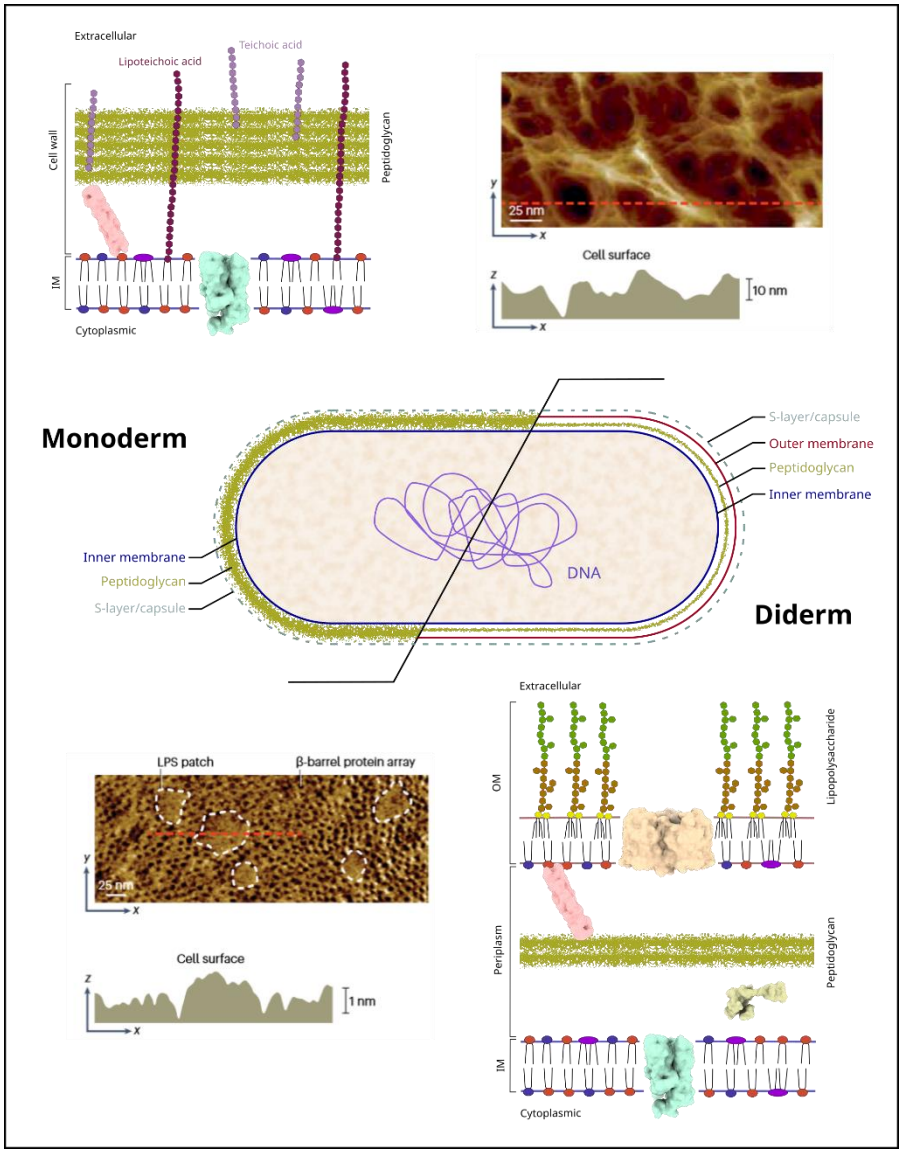


Figure 1.8: Monoderm and diderm bacteria have distinct bacterial architectures. Monoderm bacteria (upper) are characterised by an inner membrane and a thick outer peptidoglycan layer. AFM surface imaging shows a deeply grooved and pitted surface. Diderm bacteria (lower) are characterised by a dual membrane architecture enclosing a thin peptidoglycan layer. AFM imaging shows membrane domains enriched lipids (LPS) or proteins (transmembrane β-barrels, mostly forming ordered arrays). (AFM images of monoderm membranes from [134] and diderm membranes [275]).

Monoderm bacteria have a cell envelope consisting of a cell membrane, a thick (30-100 nm) peptidoglycan cell wall (which is responsible for retention of the gram-stain[131]) and often an additional polysaccharide capsule (and/or S-layer)[132] (**Fig. 1.8**, upper). The inner membrane is typically dominated by phosphatidylglycerol (PG) and cardiolipin lipids, although phosphatidylethanolamine (PE) may also be abundant, and contains the diverse set of proteins needed for signalling and substance translocation[133]. AFM and tomographic cryoEM imaging have revealed the peptidoglycan surface of the cell wall to be a network of strands with deep grooves and pores for substance access[134,135]. Additional polymers are included within the peptidoglycan network, notably teichoic and teichuronic acids which can be lipid-anchored in the membrane and help maintain the envelope integrity[136], particularly important due to the extreme turgor pressures reached in monoderm bacteria (measured up to 20 atm[137]). The capsule consists of variable oligosaccharides which may be covalently attached to peptidoglycan and, as the external facing component, it plays roles in immune evasion and phage resistance (e.g. [138]). Not all monoderms encapsulate, but it is typically observed in pathogenic strains, presumably due to the additional cellular protection offered by the extra layer[128].

In contrast, LPS-diderm cell envelopes consist of an inner and outer membrane (IM and OM), enclosing a thin peptidoglycan cell-wall (2-4 nm) in a space typically 18-22 nm wide termed the periplasm[139] (**Fig. 1.8**, lower). The IM is a phospholipid bilayer composed of a ratio typically about 75:20:5 PE:PG:cardiolipin (**Fig. 1.9**), although composition is variable, well-regulated and transversely asymmetric (the asymmetry being particularly important for regulating cell morphology[71]). The IM maintains a proton motive force (PMF), from which ATP is generated and consumed exclusively in the cytoplasm, meaning that periplasmic and OM processes have no direct access to an energy source[140] relying instead on either the intrinsic free-energy of target processes[141] or coupling to the IM via protein complex formation[142]. In addition to proteins required for IM and cytoplasmic functionality[143], the IM houses many proteins required for the transport and/or biosynthesis of substances destined for the periplasm, OM or extracellular export[144–146]. Indeed, all components of the periplasm and OM (proteins, lipids, sugars) begin their biosynthesis in the cytoplasm and must be transported through the IM before use.

The peptidoglycan layer is anchored to the OM via the lipoprotein Lpp[147], which has a critical role in determining periplasmic width, with the peptidoglycan and the OM working together as the load-bearing cellular structures[148,149]. The periplasm,

which unlike the cytoplasm is oxidising, provides an additional compartment for the cell to store potentially harmful enzymes like nucleases and phosphatases, which are beneficial for protection against exogenous threats, but would be dangerous in the cytoplasm[150,151]. The periplasm also contains the components and machinery required for peptidoglycan and OM biogenesis, maintenance, adaptation and regulation, as well as signalling pathways to transduce information from the OM to the cytoplasm[143]. Together these systems make the periplasm important for the generation and quality

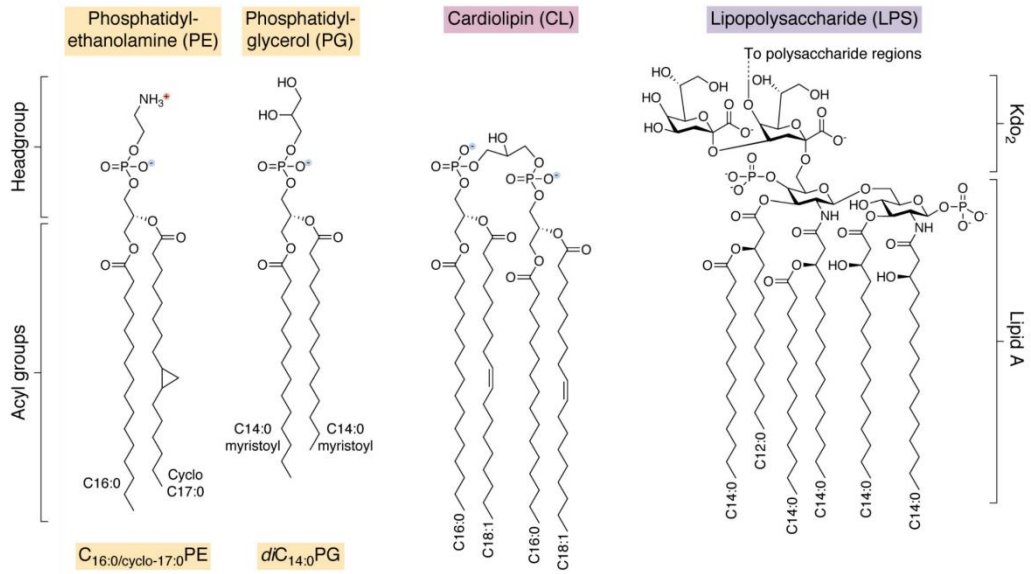


Figure 1.9: The structure of common bacterial phospholipids and Lipid A. Bacterial lipids are composed of a polar head group and hydrophobic acyl chain region. PG, cardiolipin (CL) and lipid A are negatively charged (-1, -2, -2 respectively) while PE is zwitterionic. Cardiolipin consists of two acyl tail regions connected by a glycerol linker between their phosphate groups. Some common acyl chains are shown on each lipid. Lipid A represents the most common form of lipid A-Kdo_2 found in K12 *E. coli*. (From[31]).

control of biostructures[152], and sensing and responding to cellular threats and insults[153,154].

The OM is a uniquely effectively permeability barrier due to its lipid asymmetric nature, with phospholipids (~ 80:15:5 PE:PG:cardiolipin, although lipid ratios are under regulation[155]) in the inner leaflet and LPS in the outer leaflet[156,157]. The nature of LPS, with both hydrophobic acyl chains and hydrophilic oligosaccharides, mean that when organised into the OM it is able to effectively block cellular access to both hydrophobic and hydrophilic compounds (**Fig. 1.9**, see also **Fig. 1.10a**)[158,159]. The OM has an extremely high protein (outer membrane proteins, OMPs) content compared

to the IM, with an estimated lipid:protein ratio (LPR) of ~8:1[31] and with proteins thought to cover 40-60% of the membrane surface (compared to ~32:1 LPR and 15-25% for the IM)[160]. The unique features of the OM, in particular the presence of LPS and the high protein content, combine to generate a membrane with very low lateral diffusivity[140], excellent barrier properties[159] and ordered protein arrays[161], while still allowing substance influx and efflux and (relatively) rapid OMP complement turnover, as required. OM biogenesis is next reviewed, before returning to a more detailed description of its global organisation.

1.4 The outer membrane of LPS-diderms

The OM of LPS-diderms is a protein-rich membrane with phospholipids in the inner leaflet and LPS in the outer leaflet and provides a barrier to both hydrophilic and lipophilic substances. The inner membrane leads to separation of the OM from most protein and lipid biogenesis pathways, which typically reside in the cytoplasm. Diderm bacteria have developed an array of trans-periplasm machinery to transport the often non-aqueous soluble components needed for OM biogenesis, including both continuous and discontinuous pathways (via protein complexes that bridge the periplasm and freely diffusing shuttle proteins, respectively)[162]. The biosynthesis and maintenance of the OM must occur without an adjacent energy source in the periplasm, and any energetically unfavourable process must therefore be driven by coupling to the cytoplasmic ATP pool or to the IM's PMF. The nature, transport and insertion of the major OM components (LPS, phospholipids, lipoproteins and OMPs) to the OM are discussed below.

1.4.1 Lipopolysaccharide

LPS consists of a conserved hydrophobic tail region joined to a D-glucosamine disaccharide diphosphate (Lipid A), a largely conserved, negatively charged core region of 6-7 main chain saccharides and a highly variable O-antigen oligosaccharide, with a repeating unit of 2-8 saccharides[145] (**Fig. 1.10a**). The 4-7 acyl tails of Lipid A (typically hexa-acylated in unstressed *E. coli*) are typically fully saturated and 14-16 carbons long, shorter than the average 16-18 carbons of the inner leaflet phospholipids[163,164], (although LPS variants with acyl chains 10-28 carbons long are known[165,166]). The main chain of the core saccharide is well conserved, especially the

first three sugars (the inner core; **Fig. 1.10a**, Rd1-LPS and lower), but there is significantly more diversity found in the outer core, both in the presence and identity of branching sugars and additional modifications like phosphorylation[167] (in *E. coli* there are five outer core variants[145]). The O-antigen of LPS is highly variable in length, varying from a single unit to more than 40, and in sugar identity and repeating unit structure[155,168], which can be different between closely related bacterial strains, and is used to define serotype, of which there are >190 in *E. coli*[169]. Due to their characteristic look by electron microscopy, bacteria with O-antigen containing LPS are described as having smooth membranes/LPS, while all truncated LPS below the O-antigen is generically called rough LPS and designated as Ra-Re LPS type[170] (**Fig. 1.10a**).

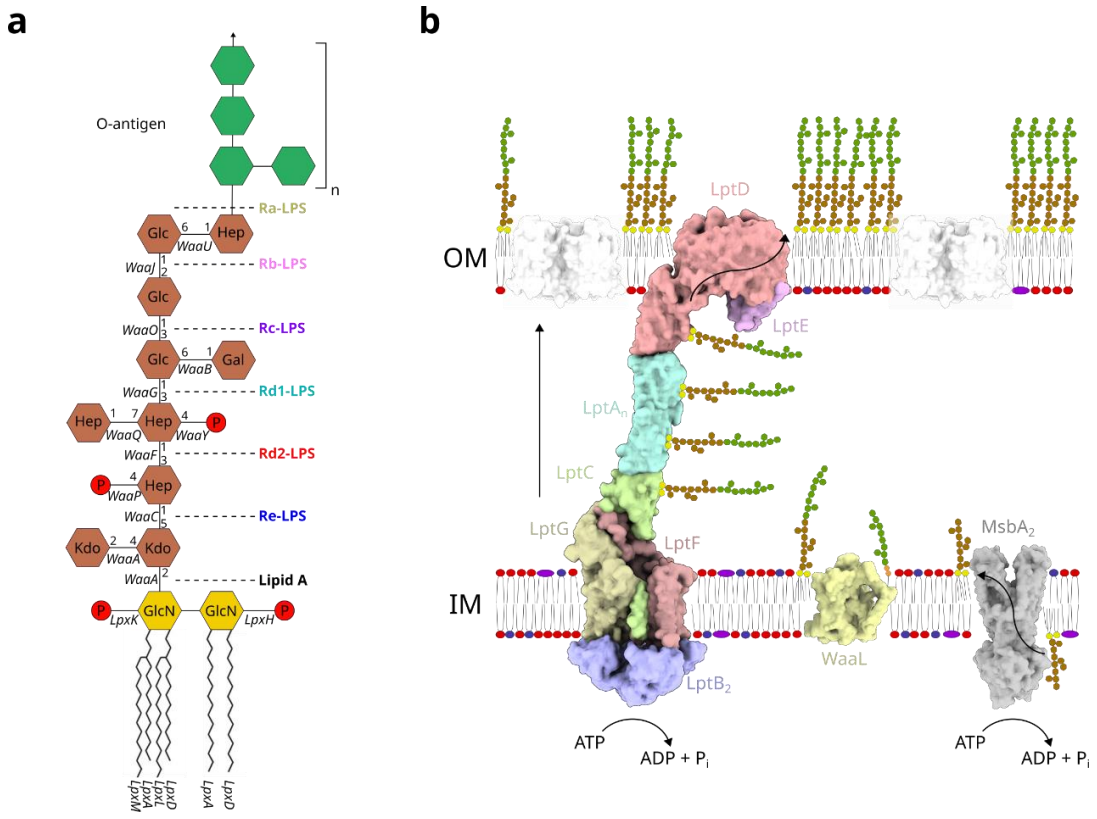


Figure 1.10: LPS structure and trans-periplasm biosynthetic machinery. (a) LPS is composed of Lipid A (yellow sugars and acyl chains), the core oligosaccharide (brown, Ra-Re-LPS naming conventions for LPS sugar truncations shown right) and the O-antigen (green). Sugar linkages and their biosynthetic enzyme names are shown on each bond, and enzymes for each acyl chain. **(b)** Overview of the LPS pathway across the periplasm from its transport across the IM by $MsbA_2$, O-antigen addition by WaaU, and IM extraction, periplasm transport and OM insertion by the Lpt pathway ($LptB_2GFC$, $LptA_n$ and LptDE, respectively).

The nature of LPS is particularly important in maintaining the barrier function of the OM, with its many saturated acyl-chains tightly packing and forming a gel-like membrane, blocking hydrophilic molecules[171]. Divalent cations coordinate the negative charges on multiple LPS molecules, 'gluing' the oligosaccharides together[172] and providing an effective barrier to both hydrophobic and larger hydrophilic molecules[173] (>~600 daltons[159,174]). *E. coli* with deep LPS truncations, in particular Re-LPS, show poor growth, increased sensitivity to antibiotics and membrane defects, emphasising the importance of LPS[175]. LPS is a potent activator of host immune systems (mostly via its lipid A component), although this can be reduced by LPS modifications[176]. Some diderms have lipooligosaccharides (LOS) rather than LPS, which are similar molecules but have an extended oligosaccharide core and no O-antigen[177].

Biosynthesis of Lipid A proceeds via the Raetz pathway at the inner leaflet of the inner membrane[145] (enzymes involved in the formation of each bond or acyl chain addition are noted in **Fig. 1.10a**). Once assembled, the nascent LPS is flipped to the outer leaflet via the ABC-transporter $MsbA_2$, binding LPS at a cavity formed by the dimer interface and flipping via protein conformational change coupled to ATP hydrolysis, observed by structural and functional characterisation of intermediate states[178,179]. The Lpt pathway mediates transport of LPS across the periplasm and insertion into the outer leaflet of the OM via a protein bridge (**Fig. 1.10b**). Following O-antigen addition at the periplasmic face of the IM, mediated by WaaL, the ABC-transporter $LptB_2FGC$ in the binds LPS, extracts it from the membrane[180,181] and passes it (via LptC[182]) into the periplasm-spanning bridge formed of up to 6/7 copies of LptA[183]. LptA forms a channel with a hydrophobic internal surface to stabilise LPS acyl-chain binding[162]. The OM LptDE complex, a 26 stranded β-barrel (LptD) plugged by the lipoprotein LptE, receives LPS at LptD's periplasmic domain which links to the LptA bridge[66,184]. A lateral gate at strands β26-β1 of LptD opens only to the outer leaflet of the OM allowing the partitioning of Lipid A to the correct leaflet[185], while the oligosaccharide is thought to be transported through the lumen of the barrel, perhaps facilitated by LptE, and gated by the large extracellular loop four[186,186]. This is termed the 'PEZ' model due to its similarities to a spring-loaded sweet dispenser, where flux through the pathway is maintained by a pushing force from new LPS molecules inserted into the continuously occupied LptA channel by $LptB_2FGC$, coupling LPS insertion at the OM to the cytoplasmic ATP energy source[187]. The Lpt pathway as described here appears largely restricted to *Proteobacteria*, with many phyla that produce LPS lacking some (e.g. *Bacteriodetes, Cyanobacteria*) or all (e.g.

Chlamydiae) of the pathway's components, indicating that other LPS transport pathways must exist[188].

Due to the linear biosynthetic pathway used to add subsequent sugars onto the growing LPS core oligosaccharide, mutations in these genes leads to controllably truncated LPS[145] (**Fig. 1.10a**, Rx-LPS phenotypes), with deeper truncations limiting growth[175]. The controlled modulation of LPS structure in response to environmental pressures can involve changes to the acyl chains[171], sugar identity or modulation of modifications, including phosphorylation and acetylation[167]. Studies have shown this diversity to be important in stress response[189], environmental and life-cycle adaptation[190], and virulence[170,191]. However, many of these processes are poorly understood and their implications for bacteria and OMPs remain largely unclear. *In situ* palmitoylation of LPS is discussed below in reference to the OM enzyme responsible for this modification, PagP.

1.4.2 Phospholipids

Unlike LPS, phospholipid transport between the inner and outer membranes is both controversial and poorly understood[162]. While there is evidence of transport in both directions across the periplasm, the transport directionality of specific pathways remains contentious[192,193]. Early electron microscopy work suggested that the outer leaflet of the inner membrane and the inner leaflet of the OM could hemifuse to facilitate lipid and lipophilic substrate transport in structures known as Bayer's bridges[194], but these are largely discounted as artefactual today (although some data suggests they may occur under certain conditions)[195]. Broadly, intermembrane lipid transport processes use proteins to shield the hydrophobic acyl chains from the cell's aqueous environment, forming either bridges, tunnels or 'boats' which shuttle lipids between membranes without being directly attached to either. These architectures are conserved across lipid transport systems generally[162], and all are implicated in phospholipid transport in the periplasm. The different known pathways in *E. coli* are briefly discussed below.

The maintenance of lipid asymmetry (Mla) system is the best characterised pathway (**Fig. 1.11**, left). It consists of the inner membrane ABC-transporter $Mla(FEB)_2D_6$, the periplasmic shuttle MlaC, and the OM MlaA, which additionally recruits an OmpC trimeric porin[64]. Deletion of Mla components causes accumulation of mislocalised phospholipids in the outer leaflet of the OM, indicating it plays a role in OM asymmetry homeostasis[196]. It is expected, therefore, that the Mla system acts as a retrograde (i.e., OM to inner membrane) transporter, although some uncertainty remains[193]. Recent work suggests that under equilibrium the system can spontaneously, but slowly, transport in both directions, but ATP hydrolysis at the inner membrane drives retrograde transport by extracting lipids from the pathway and passing them into the IM[192]. Transport thus starts at MlaA, an unusual lipoprotein in that the bulk of the protein resides in the OM forming a hydrophilic channel through

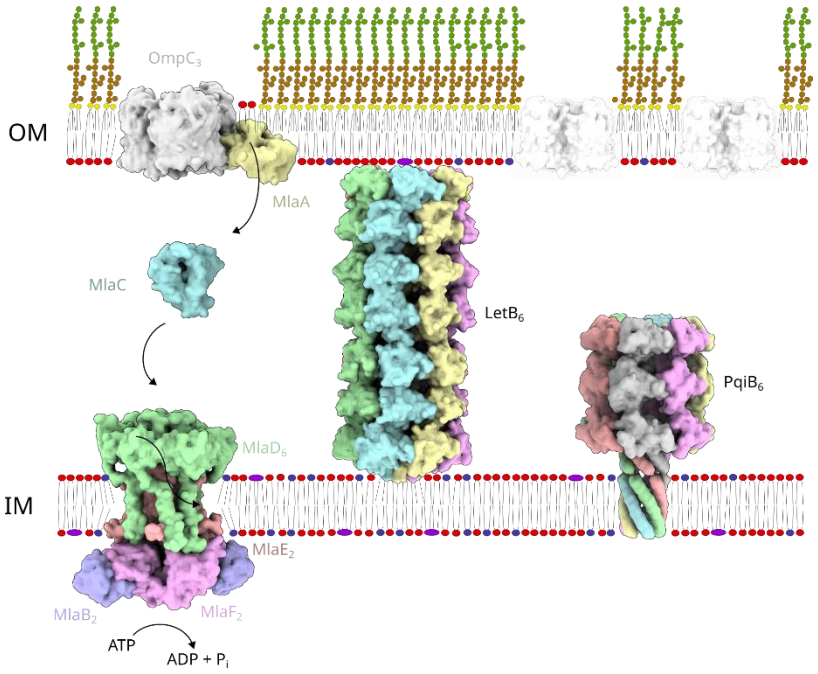


Figure 1.11: Overview of structurally defined phospholipid transport pathways in the periplasm. Left: The Mla pathway transports erroneously located phospholipids in the outer leaflet of the OM back to the IM by extracting them via $OmpC_3$ associated MlaA, passing the phospholipid to the periplasmic shuttle MlaC, which in turn passes it to the ABC-transporter $Mla(FEB)_2D_6$ to insert into the IM. The ATP hydrolysis at this complex drives the directionality of the pathway. $LetB_6$ (**centre**) and $PqiB_6$ (**right)** forms a tunnel across the membrane for lipid transport, although their attachments at the membrane and mechanisms are not well defined.

which phospholipid headgroups could feasibly be transported shielded from the membranes hydrophobic core[61], and energetically advantageously bypassing the inner leaflet. Periplasmic MlaC accepts lipids from MlaA and transports them to Mla$(FEB)_2D_6$, handing the lipid off to an $MlaD_6$ binding site, where ATP hydrolysis releases the lipid into the membrane and prevents backflow[64].

E. coli also contain MCE (mammalian cell entry[197], a protein family broadly implicated in lipid transport, of which MlaD is a member) complexes that form tunnels across the periplasm: YebS-LetB and PqiABC[198] (**Fig. 1.11**, centre and right). Both these systems play minimal roles in rich media[198,199] but are induced by oxidative stress (PqiABC)[200] or LPS defects (YebS-LetB)[198]. LetB forms a broad tunnel across the membrane, structurally resolved co-purified with lipids[201], interacting with YebS at the inner membrane and with no identified anchor at the OM[202]. In comparison, the Pqi system is held at the OM by the lipoprotein PqiC, at the inner membrane by PqiA and bridges the periplasm with a hexameric PqiB[203] (only partial structure data are available). However, the directionality, energetics and mechanisms of lipid transport in these systems is uncertain.

Other protein complexes are implicated in lipid transport, although have not yet been rigorously demonstrated. The Tol-Pal complex is important for maintaining OM stability and cell division[204]. Deletion of the complex leads to an excess of phospholipids in the OM[205], suggesting a role in retrograde transport, but structural studies show no clear binding sites or pathway for lipid translocation[206,207]. AsmA-like proteins, which have a single pass helix at the inner membrane and large periplasmic domains predicted to form β-taco domains with a putative lipid-binding hydrophobic inner channel[208], include AsmA, TamB, YdbH and YhdP[195]. Little is known about these transporters, although TamB is thought to interact with the Omp85 family member TamA at the OM, providing a plausible partner to hand lipids to[209]. YdbH, YhdP and TamB are mostly redundant, but deletion of all three leads to serious OM lipid homeostasis defects[209].

Although non-typical, some LPS-diderm bacteria also synthesise sterols and recent work has identified a transport route to the OM in *Methylococcus capsulatus* that appears structurally distinct from its eukaryotic counterparts[210], although the functional role of such lipids is a mystery. Much remains to be understood about lipid transport in the periplasm and mechanisms of efficient OM phospholipid homeostasis, but the diversity in pathways identified define it as a critical mechanism.

1.4.3 Lipoproteins

There are estimated to be ~50 OM lipoproteins in *E. coli*, implicated in diverse functions roles, from membrane biogenesis and homeostasis to substrate import and cell division, although many are of unknown function[52]. The well-defined Lol (localisation of lipoproteins) pathway transports mature lipoproteins to the OM (**Fig. 1.12**), via a mechanism analogous to the Mla pathway, but operating in reverse[211]. Following synthesis in the cytoplasm, lipoproteins are targeted to the SecYEG complex via their N-terminal signal peptide[51], which partitions into the inner membrane and holds the now periplasmic lipoprotein in place. The region adjacent to the membrane contains the lipobox motif (consensus: Leu-(Ala/Ser)-(Gly-Ala)-Cys)[52] which is recognised and the cysteine covalently linked to a diacyl-glycerol by the enzyme Lgt[212]. Following cleavage of the now redundant signal-peptide by LspA (signal peptidase II)[213], the exposed N-terminal amine is also acylated by the enzyme Lnt[214].

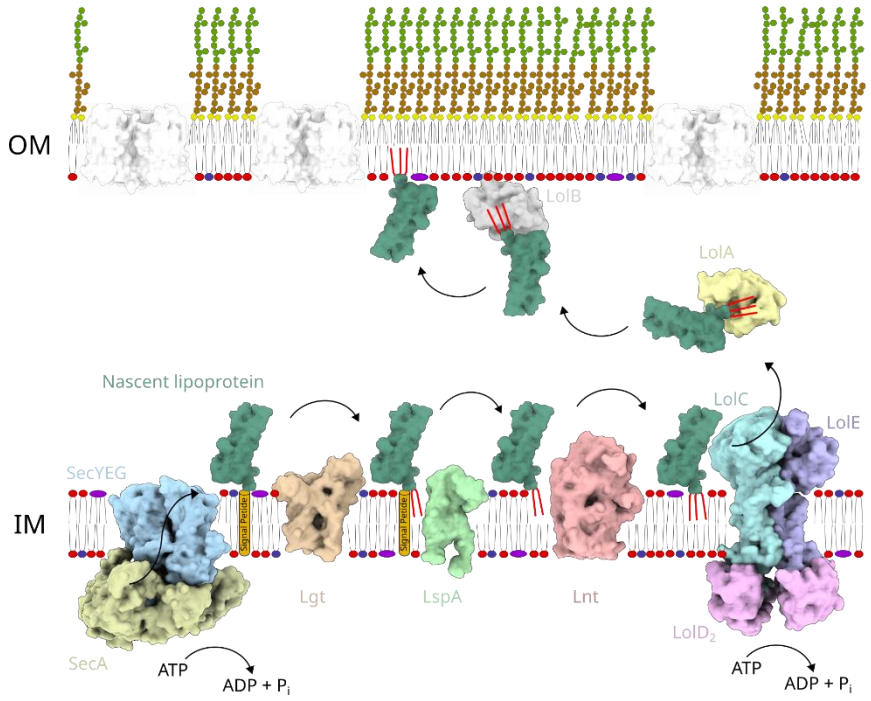


Figure 1.12: OM lipoprotein maturation and trafficking to the OM via the Lol pathway. Lipoproteins are translocated to the periplasm via the SecYEG complex (lower left) and then tri-acylated at the IM. Mature lipoproteins are extracted from the IM by the ABC-transporter $LolCD_2E$ and passed to the periplasmic shuttle LolA. LolA shields the acyl chains while crossing the periplasm and then passes the lipoprotein to LolB which inserts it into the OM. Note the similarity of the Lol pathway to the Mla pathway (**Fig. 1.10**).

Mature tri-acylated lipoproteins are trafficked to the OM (if they do not contain the IM retention signal of an aspartate at the +2 position (from N-terminal cysteine)) via recognition and binding to $LolCD_2E$ complex[215]. $LolCD_2E$ couples ATP hydrolysis at $LolD_2$ with lipoprotein extraction from the membrane and then passing, via LolC, the lipoprotein onto the periplasmic shuttle LolA[216]. To insert into the OM, LolA interacts with LolB, itself an OM lipoprotein, which can accept the translocating lipoprotein (due to its higher affinity for it than LolA)[217]. Some lipoproteins in *E. coli* are proposed to be extracellularly exposed (e.g. BamC[218]), and in other LPS-diderm bacteria this is well defined (e.g. *Bacteriodetes*[57]), but the mechanism of export is unknown.

1.4.4 Outer membrane proteins

Outer membrane proteins (OMPs) are, with very few exceptions (e.g., *E. coli* protein Wza[219]), transmembrane β-barrels, consisting of 30-60° tilted antiparallel β-strands[220]. Transmembrane strands are linked by typically shorter periplasmic turns and longer extracellular loops, with the N- and C-termini remaining in the periplasm[221] (**Fig. 1.13a**). Barrels may be monomeric or formed from homo-oligomers with each unit contributing strands to the barrel, in monomeric OMPs the first and last β-strand come together to close the barrel forming the β-seam. The number of strands per barrel is consistently even, and for monomeric OMPs ranges between 8 and 36 β-strands (SprA[222]), although oligomeric OMPs can reach 60 strands[223]. Bands of aromatic residues at the hydrophilic-hydrophobic membrane interface anchor the barrels in place (the aromatic girdle)[224], and in some cases manipulate the local membrane thickness[225]. Lumen facing residues are typically hydrophilic, and their interactions are implicated in control of the shape of the barrel, which are often ellipsoidal or kidney-bean shaped[226].

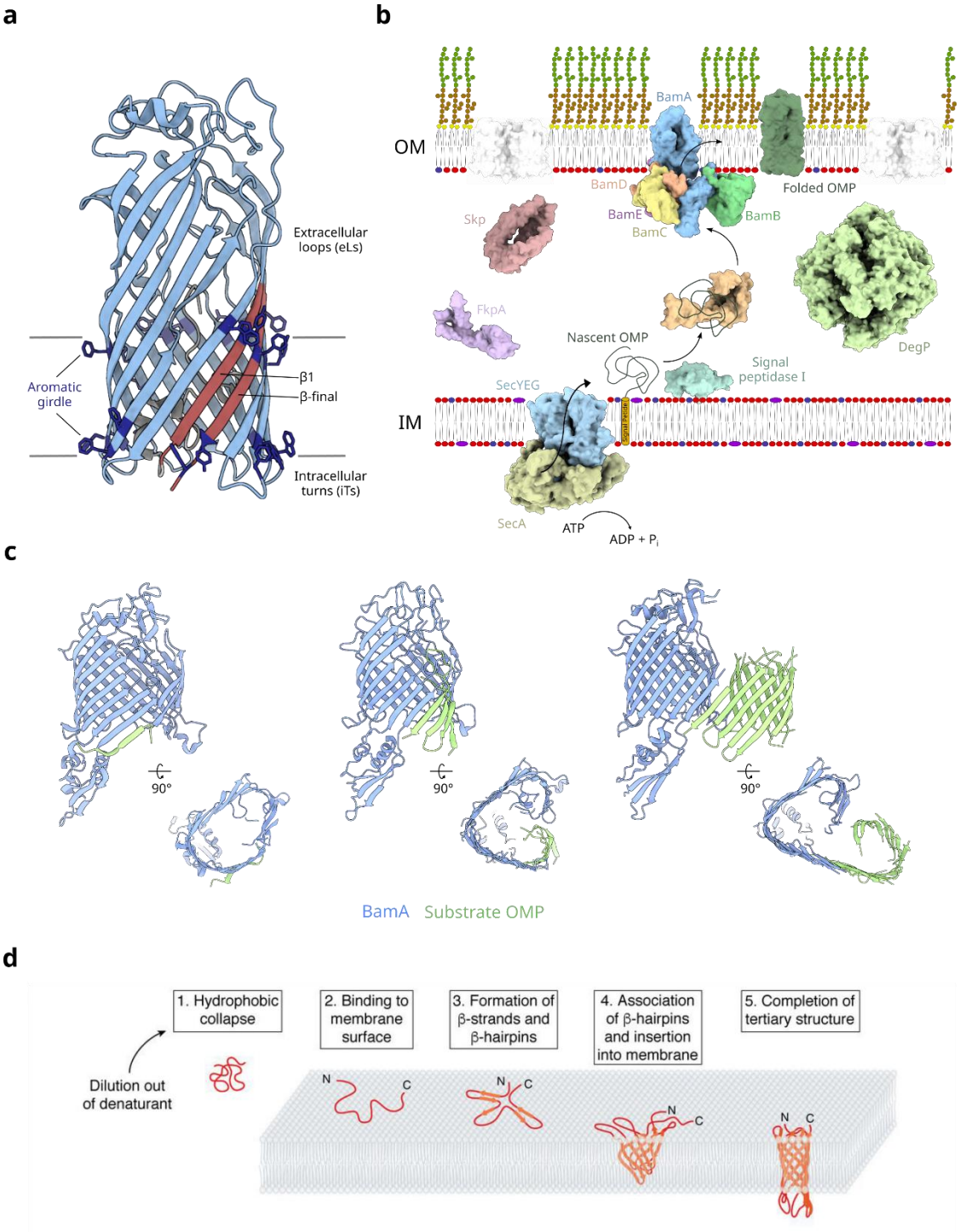


Figure 1.13: (Legend overleaf)

Figure 1.13: OMP structure, biogenesis and folding. (a) OMPs closed are transmembrane β-barrels, with a β-seam where the first and last β-strands come together (red, for monomeric barrels) and an aromatic girdle anchoring them in the membranes. OMPs typically have short intracellular turns and long extracellular loops (example OMP is FadL, PDB: 1T1L[60]). **(b)** Nascent OMPs are translocated to the periplasm via SecYEG (lower left) where signal peptidase cleaves their signal peptide. SurA delivers the unfolded OMP to the BAM complex, where folding and membrane insertion occurs. Other periplasmic chaperones (Skp, FkpA) and the dual functional chaperone-protease DegP help recover or degrade misfolded and stalled OMPs. **(c)** OMP folding via BAM proceeds via β-strand templating of the substrate C-terminal strand off BamA's β1 (left) and progressive substrate barrel elongation from the templated strand (centre and right). Barrel closure of the substrate OMP will release it from BAM and revert BAM to its apo state (PDBs: 7RI4[227], 7TT7[228], 6LYQ[229]). **(d)** Mechanism of intrinsic folding of OMPs. OMPs initially bind to membranes and from initial β-structure. The β-hairpins begin to associate as membrane insertion occurs and folding to the final secondary and tertiary structure is cooperative. ((d) adapted from [30])

OMP biogenesis is outlined in **Fig. 1.13b**. OMPs are translated in the cytoplasm, targeted to the SecYEG translocon via their signal peptide which partitions into the membrane and is cleaved following OMP transport through the inner membrane by signal peptidase-I, thus releasing the OMP into the periplasm[230]. The β-strand nature and hydrophilic lumen residues of OMPs are important for preventing partitioning into the IM by SecYEG, as the nascent chain does not present consistently hydrophobic regions (i.e., once folded in the OM only every other residue faces the membrane), allowing it to be exported[43]. Periplasmic chaperones SurA, Skp and FkpA prevent OMP aggregation[225,231,232], while the protease DegP can cleave misfolded OMPs[233]. Under some stress responses associated with increased OMP misfolding/aggregation (e.g. heat shock) DegP becomes essential[233,234]. SurA is the best studied chaperone and is known to mediate expansion of OMP polypeptides via dynamic motion of its three-domain structure[235,236], with a binding preference to aromatic-X-aromatic motifs which are enriched in OMPs[237]. SurA is also known to interact with the β-barrel assembly machinery (BAM) that inserts OMPs into the membrane[235], and thus, as well as preventing OMP misfolding, it likely retains them in a folding competent state. Skp functions as a homotrimer with three α-helical arms forming a cavity for client binding and is important for larger substrates with plug domains[238], as well as recovering[239], targeting for degradation[240] or sacrificially removing[241] OMPs that have fallen off/stalled on the SurA-BAM pathway. Multiple copies of either SurA or Skp can bind to a single OMP polypeptide, important for maintaining solubility of longer OMPs[231,238]. While FkpA, and other chaperones like

Spy, are not part of the canonical pathway, they become important under stress conditions[242].

OMP folding and partitioning into the membrane is mediated by BAM. *E. coli* BAM is a 203 kDa complex formed of the Omp85 family 16-stranded β-barrel BamA and the four lipoproteins BamBCDE[56,243,244]. BamA additionally contains five periplasmic polypeptide transport associated (POTRA) domains which support BAM function and provide a scaffold for BamBCDE binding[245,246]. BamA and BamD are essential, and excellently conserved across diderms[247]. BamD is implicated in substrate quality control and recognition[248,249]. Deletion of BamB or BamE leads to significant OMP folding defects (*in vitro* studies demonstrate that while BamE supports folding of all OMPs, BamB is involved in folding of larger substrates[250,251]), while loss of BamC is only very mildly phenotypic and the exact role of BamC remains on open question[139]. BamB is further important for organising multiple copies of BAM together into 'folding precints'[252].

The BAM complex, and particularly BamA, show significant conformational diversity in resolved structures[243,253,254]. BamA's β-seam is either open (i.e., no hydrogen bonding between β16-β1) or closed (partial or full hydrogen bonding), with large concomitant changes to the conformation and/or relative location of the POTRA domains and lipoproteins[56,243] (the open lateral gate can be seen in **Fig. 1.13b**). The lipoproteins appear to bias the conformation of BamA to be β-seam open[255], with it closed in structures of BamA alone but open in the majority of structures of the full complex[254]. This conformational flexibility seems important for function, and disulphide-locking the complexes reduces their activity *in vitro* and bacterial viability *in vivo*[255,256].

Mechanisms for BAM function focus around membrane destabilisation and direct BAM-OMP interactions. BAM reconstituted in proteoliposomes has been shown to destabilise the membrane significantly[255]. While unlikely to be sufficient to catalyse OMP folding, the destabilisation of the membrane decreases the energetic barrier for partition into the membrane. Templating of the β-strands of the nascent OMP via β1 of BamA is common across all proposed mechanisms[45], although whether OMPs partially fold in the periplasm or in the BamA lumen (supported by extracellular loop 6) remains uncertain and is likely variable for different OMPs[257]. In recent years multiple structures of stalled BAM-OMP complexes with various degrees of OMP susbtrate folding have been solved by cryoEM (by crosslinking partially folded OMP substrates to BAM) (e.g. **Fig. 1.13c**, [227,228,258,259]). These demonstrate various hybrid barrels, where BamA's lateral seam has opened to facilitate β-sheet extension via

hydrogen-bonding between BamA's β1 and the final, C-terminal strand of the nascent OMP (leaving BamA as an open barrel during folding, **Fig. 1.13c**, top-down views). However, precise mechanisms of membrane partition and insertion, including release of the newly folded β-barrel from BAM, are yet to be elucidated. Some evidence (pulldowns and very low-resolution EM) indicates a periplasm-spanning supercomplex can form between SecYEG, BAM and the periplasmic chaperones, suggesting at least transient interactions can facilitate rapid folding pathway flux, although whether a stable complex forms is an open question[260,261].

Unlike most helical transmembrane proteins, many OMPs can be refolded readily into membranes *in vitro* from a denatured state (e.g. 8 M urea) to their native conformation in a process termed intrinsic folding[262]. Although using symmetric bilayers of non-native lipids (i.e., without LPS), OMP intrinsic folding studies provide a controllable context to study mechanisms of membrane protein folding and insertion, as well as the role of the lipid bilayer. When folding into liposomes, OMPs initially bind to the membrane surface and adopt a partial β-strand structure before undergoing conformational re-arrangement and tunnelling their loops through the membrane to allow the native structure to form via a concerted mechanism[31,263] (**Fig. 1.13d**). Studies with the dominant model OMP, OmpA, indicates folding is cooperative, with either both secondary and tertiary structural formation together or neither occuring[264]. Membrane properties have large effects on OMP folding rates/yields. Folding is faster and more efficient into liposomes with shorter[225] or less saturated acyl chains[265], and those in a fluid (rather than gel) phase, although folding is fastest at the lipid transition temperature (T_m)[266], presumably due to the introduction of additional membrane defects at this temperature. PE and PG lipid headgroups have been shown to introduce a kinetic barrier to folding into $C_{10:0}$ lipid bilayers[267,268], although this is not recapitulated in $C_{14:0}$ lipids (possibly because of the additional kinetic barrier of a thicker membrane dominating the folding[269]). High protein concentrations (i.e. low LPRs) also inhibit folding[270]. OMP folding *in vitro* is much slower than *in vivo*, typically taking minutes or hours compared to seconds *in vivo* (*E. coli* doubling time at 37 °C is 20 minutes)[271,272]. Once folded, most OMPs are highly resistant to unfolding[262].

Although *E. coli* harbour 60-70 unique OMPs[31], ~70% of the protein content is composed of just three OMPs, and typically only a handful of OMPs are present at a fraction greater than 1%[273] (**Table 1.1**). ~35% of the OMP content is composed of the trimeric general porins OmpF and OmpC[273,274], formed of three 16-stranded OMPs which organise into arrays which give the OM its hexameric lattice-like look under

AFM imaging[275,276] (**Fig. 1.8**, lower). The relative abundance of OmpF and OmpC varies based on environmental context, in particular osmolarity[277]. The OM of nearly all characterised LPS-diderms are dominated by trimeric porins, although with varying details[278] (e.g., *Klebsiella* have four porins, with only one majorly expressed at a given time[279]). OmpA, the single most abundant OMP and also about ~35% of the *E.coli* OM protein content is detailed below[280], along with the OMPs FusA and MipA, and the three enzyme OMPs of *E. coli* OmpT, PagP and OmpLA, all six of which are explored extensively in Chapters 3-5.

OMP name	Strands	Absolute Abundance	Relative Abundance (%)
OmpA	8	207, 618	30.0
OmpC	16	163, 538	23.4
OmpX	8	125, 295	17.9
OmpF	16	88, 988	12.7
OmpT	10	40, 237	5.8
MipA	10	20, 925	3.0
Tsx	12	14, 911	2.1
TolC	12 (4x3)	8, 768	1.3
FadL	14	6, 912	1.0

Table 1.1: Estimated absolute and relative abundance of common OMPs in *E. coli* grown in rich media at 37 °C. Abundance estimated by absolute cellular synthesis rates converted to molecules per generation based on a growth doubling time. These will differ from final OM copy number, but the relative ratios should be approximately correct, although will vary significantly depending on growth conditions. (Data from [273] and numerical analysis adapted from [31]).

OmpA

OmpA, identified as an *E. coli* OM heat modifiable protein in 1977 (now known to be the folded-unfolded bandshift under cold (non-boiled) SDS-PAGE), has been identified at high copy numbers across many (but not all) diderms, with upwards of 100, 000 copies per cell[226,273]. Following a remarkably accurate model for the barrel of OmpA based on Ramen spectroscopy[281], its N-terminal transmembrane domain (tOmpA) was structurally resolved by crystallography as an eight-stranded barrel[282,283], and its periplasmic, soluble C-terminal domain crystallised separately[284,285] (**Fig. 1.14a**). *In vivo* the C-terminal domain binds to peptidoglycan (via residues R156 and D141 (R186 and D171 in *A. baumanni*), **Fig. 1.14b**), and along with the lipoprotein Lpp, it is crucial for maintaining cell envelope integrity[286,287], with ΔOmpA or $\Delta OmpA_{C\text{-}ter}$ having compromised cell membranes[288]. Native mass

spectrometry data indicate that OmpA likely forms homodimers in the OM, which help prevent the inter-domain flexible linker being cleaved[289].

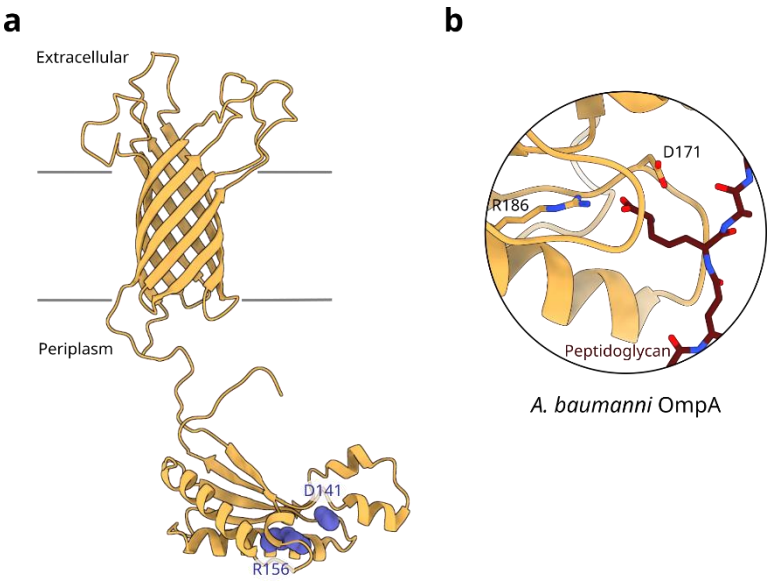


Figure 1.14: OmpA has a two-domain structure and non-covalently interacts with peptidoglycan. (a) OmpA is composed of an N-terminal eight-stranded transmembrane β-barrel and a C-terminal periplasmic domain, including the residues R156 and D141 (in *E. coli*) that bind peptidoglycan (PDBs: 1G90[290], 2MQE[284]). **(b)** Details of peptidoglycan binding by the C-terminal domain of *A. baumanni* OmpA (PDB: 4G4V (unpublished)).

Early conductance data suggested that OmpA could form an ion pore[291,292], but although the barrel structure had enclosed water cavities, it did not contain a continuous channel. Additional structural characterisation by NMR and MD identified considerable flexibility in the luminal residues[290,293], and later work found a salt-bridge switch in the core of the barrel could facilitate the opening of a small pore[294]. Alternatively, a 16-stranded structure of OmpA has been proposed where it's C-terminal domain forms eight transmembrane strands[295,296], supported by conductance data[9] suggesting forms of OmpA with both small and large pores[298] and temperature sensitive folding[298,299]. Although the large pore model is disfavoured, not least due to the important functional role of the periplasmic C-terminal domain, it remains possible that a small subset of OmpA retains this structure in the OM. Immunogold electron microscopy and immunofluorescence studies indicated that even in the presence of LPS the loops of OmpA were antibody accessible in *Mannheimia haemolytica*[300].

OmpA has a diverse range of functions, including its exploitation as a bacteriophage[301,302] or bacteriocin receptor[303], and being important for conjugation[304,305]. In *E. coli*, OmpA is required for efficient crossing of the blood-brain barrier via endothelial cell invasion[306,307] and also appears important for adhesion to cells and invasion of colonic epithelial cells[308,309]. Following macrophage phagocytosis, bacteria are able to prevent macrophage apoptosis by activating the antiapoptotic Bcl_{xl}, allowing them to replicate inside the phagosome and eventually burst the macrophage[310] (ΔOmpA bacteria do not survive phagocytosis[311], but OmpA is not solely sufficient for this process[312]). OmpA is also implicated in biofilm formation[313,314]. Given its conservation and abundance it is not surprising that OmpA is directly targeted by the immune system: serum amyloid A protein binds OmpA and induces bacterial opsonization[315], neutrophil elastase targets OmpA and thus permeabilises the OM[316], as well as being recognised by macrophages[317]. Indeed, OmpA fragments from *K. pneumoniae* are so efficiently presented as antigens that it is used as a vaccine carrier[318].

OmpA expression is controlled predominantly at the mRNA level, with mRNA half-life increasing proportionally with growth rate[319,320], mediated by an interplay between the structured 5'-untranslated region, the binding protein Hfq[321] (which destabilises and targets for the mRNA degradation[321,322]) and the small RNA MicA which targets Hfq to OmpA's mRNA. OmpA is downregulated during envelope stress σ^E response, but upregulated during polyamine exposure[323]. Modulatory factors of OmpA intrinsic folding have also been extensively studied, with many of the seminal *in vitro* folding studies using OmpA [262,263]: it was the first OMP to be intrinsically refolded into detergent[324] and membranes[325].

OmpA is pathogenically important across all the species it has been characterised in, including *Yersinia, Pseudomonas*, *Klebsiella*, *Haemophilus*, *Neisseria*, and *Chlamydia* species, typically acting via some or all of the *E. coli* mechanisms described above: enhancing adhesion, intracellular invasion and survival, cellular toxicity and increasing inflammation[326]. OmpA's role in *Acinetobacter* has been particularly well studied, where it is important for regulating adhesion, biofilm formation, aggressiveness and immune response[327,328]. Intriguingly, *A. baumanni* OmpA is known to be cytotoxically localised to the nucleus of invaded cells via its nuclear localization signal (KTKEGRAMNRR) near the C-terminus, although how the toxicity is mediated is unclear[329]. An additional survival strategy of *A. baumanni* is the release of Outer Membrane Vesicles (OMVs) with high concentrations of OmpA which

can accumulate extracellular drugs, particular β-lactams, protecting the bacterium[330,331]. Intriguingly, a few monoderm bacteria also contain OmpA homologs, with strong homology in the C-terminal domain, suggesting conserved roles of membrane anchored peptidoglycan binding[332].

MipA

Although first identified as an OMP in 2000[333], little is known about MipA (MltA Interacting Protein), despite its relatively high abundance (~2% of the OMPome[273]). Its predicted structure (**Fig. 1.15a**) reveals a 10-stranded OMP, and its conservation in pathogenic *E. coli* strains and commensal *Proteobacteria* suggest its importance[334,335]. Notably, it facilitates complex formation between a murein polymerase (PBP1B) and a murein hydrolase (MltA), possibly representing a peptidoglycan-synthesizing holoenzyme[336]. (Alphafold2 structural modelling of this complex reveals biochemically plausible structures with good electrostatic matching, **Fig. 1.15b**). It thus appears to be important in peptidoglycan biosynthesis, and possibly septation, although under rich media conditions ΔMipA shows no phenotype[337]. Several studies have linked MipA to stress response: its mild over-expression increases σ^E activity and increases expression of DegP[338], it is downregulated in response to kanamycin[339], it is upregulated in *E. coli* sessile culture[340] or in response to parathion[341], and under UVC-stress, γ-irradiation or starvation in some *vibrios*[342,343] A very recent study of *P. aeruginosa* MipAB identified a polymyxin binding site inside MipA and subsequent induction of the efflux pump MexXY-OprA, as well as inducing a more broad envelope stress response[344].

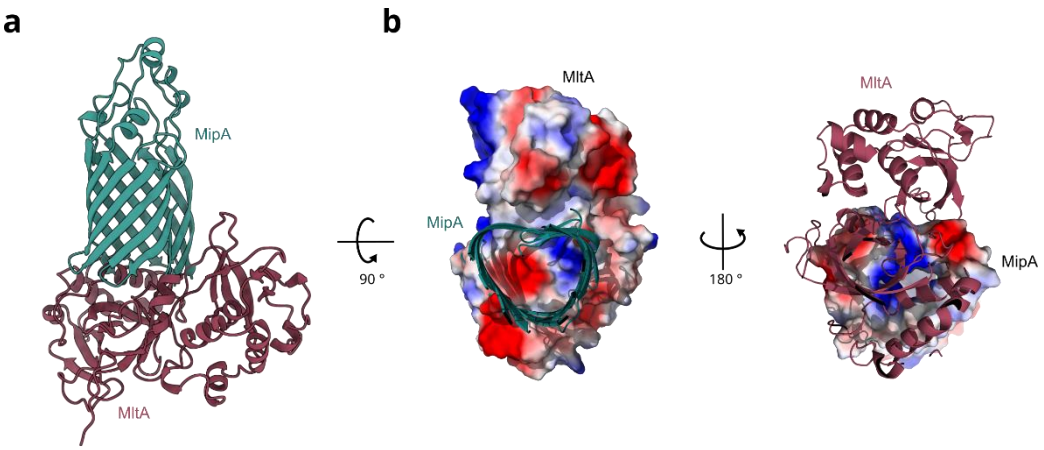


Figure 1.15: Predicted MipA-MltA complex. (a) Alphafold2 predicts MipA as a 10 stranded β-barrel (turquoise) and predicts the biochemically verified structure of the MipA-MltA (purple) complex with high confidence. **(b)** The predicted MipA-MltA complex shows good electrostatic matching between the base of the MipA barrel and the upper surface of MltA.

TonB-dependent transporters and FusA

Passive diffusion is insufficient to acquire adequate quantities of all the nutrients or import the larger substrates required for bacterial growth, and thus bacteria have developed systems that couple energy sources at the IM or cytoplasm to make active OM transporters. TonB-dependent transporters (TBDTs) are a broad, essential class of active nutrient transporters that couple the inner membrane's PMF with active substrate transport across the OM, via the periplasm spanning TonB protein and an inner membrane motor complex[142]. In *E. coli* TBDTs are typically involved in iron acquisition from different sources, but more broadly examples are known to import other nutrients including sugars and nucleotides.

TBDTs are 22-stranded barrels with a periplasmic N-terminal plug domain entirely occluding the barrel lumen in the apo-state and extracellular loops containing a substrate binding site. At the N-terminus of the plug the sequence includes a typically disordered, conserved 5-residue motif known at the TonB-box, to which TonB can bind and thus allow energetic coupling across the periplasm (**Fig. 1.16a**). It is thought that proton flow through the IM motor complex can generate a pulling force on TonB which in turn pulls the plug, leading to partial unfolding of β-sheet structures and opening a pore through the barrel lumen, supported largely by single molecule force spectroscopy, EPR (electron paramagnetic resonance) and molecular dynamics

studies[142,345]. To prevent unproductive cycling, apo-TBDTs occlude the TonB box by ordered binding at the base of the plug, with substrate binding transferring a conformational change to dislodge the TonB-box and thus expose it for TonB binding. *In vivo* EPR of the TBDT BtuB has also recently shown reordering of the C-terminal region of the plug *in vivo* upon substrate binding[345], apparently strictly dependent LPS, indicating important *in vivo* details not captured by current models. The IM motor, in *E. coli* $ExbB_5D_2$ plus TonB, couples the PMF to a pulling force via rotational motion, and shows strong homology to the MotAB flagellar motor. Exactly how rotary motion is transduced via TonB is unclear, although it has been speculated that $ExbB_5D_2$-TonB move through the IM laterally under the rotational motion and this kinetic energy is transduced to the TBDT[346].

A recently discovered, widespread TBDT family imports iron-containing host proteins for proteolytical digestion allowing iron release and utilisation[347], including the *E. coli* transporter/protease YddB/PqqL system, important for uropathogenic bacterial fitness[348,349]. FusA/FusC, from plant pathogen *Pectobacterium spp*, is the best characterised example of the family and, along with the TonB-like FusB and IM transporter FusD, imports and degrades the 12 kDa ferredoxin host protein[350] (**Fig. 1.16b**), although there are conflicting data on whether ferredoxin is unfolded for import. A crystal structure of FusA showed that the extracellular loops of FusA are particularly extensive (**Fig. 1.16c**), at least in part to help form a binding site for the large ferredoxin substrate[351] (**Fig. 1.16d**).

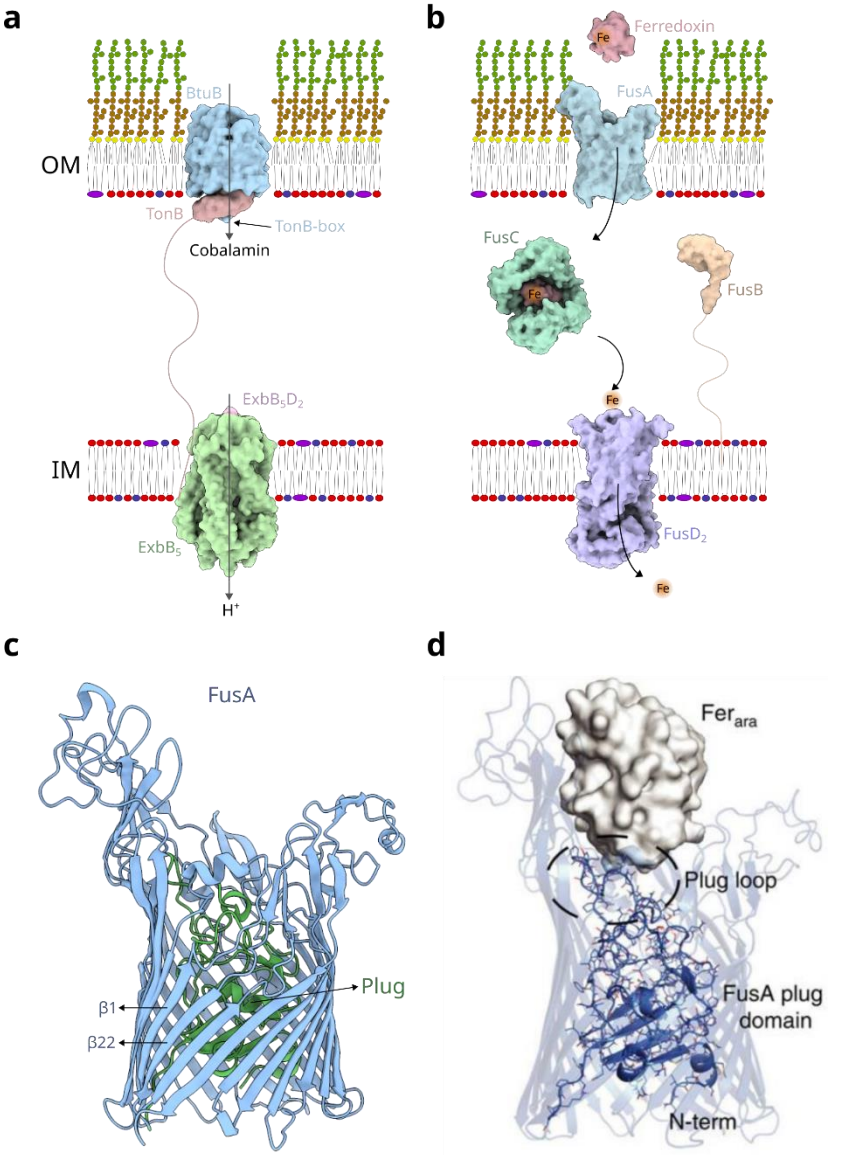


Figure 1.16: FusA is a protein importing TonB-dependent transporters (TBDTs). (a) TonB dependent transporters couple nutrient acquisition at the OM with the PMF at the IM via an IM motor complex ($ExbB_5D_2$), the periplasmic spanning TonB and the OM transporters (e.g. BtuB, shown here). **(b)** The operon FusABCD works together to import and cleave ferredoxin to access its iron (orange circles) (FusA: TonB-dependent transporter, FusB: Tonb-like protein, specific for this system, FusC: ferredoxin protease, $FusD_2$: ABC transporter that imports the liberated iron to the cytoplasm). **(c)** FusA, like all TBDTs is a 22-stranded barrel (blue) with an N-terminal plug domain (green). (PDB: 4ZGV[351]) **(d)** The large extracellular loops of FusA facilitate ferredoxin substrate binding at the top of the lumen and interact with the plug loop ((d) adapted from [351])

OmpT

OmpT was first identified in 1973 as an OMP of unknown function[352], and subsequent early research demonstrated its protease activity (via *in vivo* cleavage of FepA and inhibition by benzamidine[353]) and a substrate preference for paired basic residues (by *in vitro* characterisation[354,355]). Following its discovery, OmpT was rapidly implicated in the pathogenesis of a range of diderms such as *E. coli*, *Shigella*, and *Yersinia*[356–360]. OmpT's *in vivo* function is to cleave host-generated cationic antimicrobial peptides (AMPs) against which it has a broad action[361], and ΔOmpT cells are hypersusceptible to treatment with the AMP[362].

OmpT forms a 10-stranded barrel with large, mostly structured extracellular loops (**Fig. 1.17a**), a putative LPS binding site (based on homology from FhuA, **Fig. 1.17b**)[363] and a very electronegative central cleft to recruit its positive substrates (**Fig. 1.17c**). Development of a facile fluorescence activity assay (c.f. **Fig. 4.1c**) revealed that OmpT activity shows an LPS dependence[105], cleaves folded substrates with a consensus (Arg/Lys)↓(Arg/Lys)-Ala[364,365] (↓ indicates cleavage site) and denatured substrates at the carboxyl-side of basic residues[366] (although the P4-P2' sites are important for mediating efficient cleavage[367]). Mutational and structural analysis identified Asp210/His212 as the key catalytic dyad supported by Asp83/Asp85[368] (**Fig. 1.17d**), which act by nucleophilicly activating water to attack and cleave the substrate carbonyl[369]. Multiple simulation studies suggest that loop conformation fluctuations are important for substrate binding, likely by altering the local electric field around the catalytic site[370,371].

Steady-state OmpT expression in rich-media represents ~5% of the total OMP content[273], but expression is well regulated. Both enteropathogenic and uropathogenic *E. coli* strains have characterized thermoregulation, with OmpT expression activated at host temperature of 37 °C[372], although final quantities depend on specific niche[373,374]. OmpT and its homologs also show conserved regulation by the PhoPQ two-component signalling pathway[375]. Membrane damage by sub-lethal concentrations of AMPs induce OmpT-loaded OMV production, which further protects bacterial populations[376]. In addition to protection from AMPs, elevated OmpT expression and/or activity has been linked to bacterial adhesion, cell invasion and intracellular bacterial community formation, and upregulating proinflammatory cytokines during infection[377–379]. In addition, OmpT is active against some bacteriocidal colicins[380].

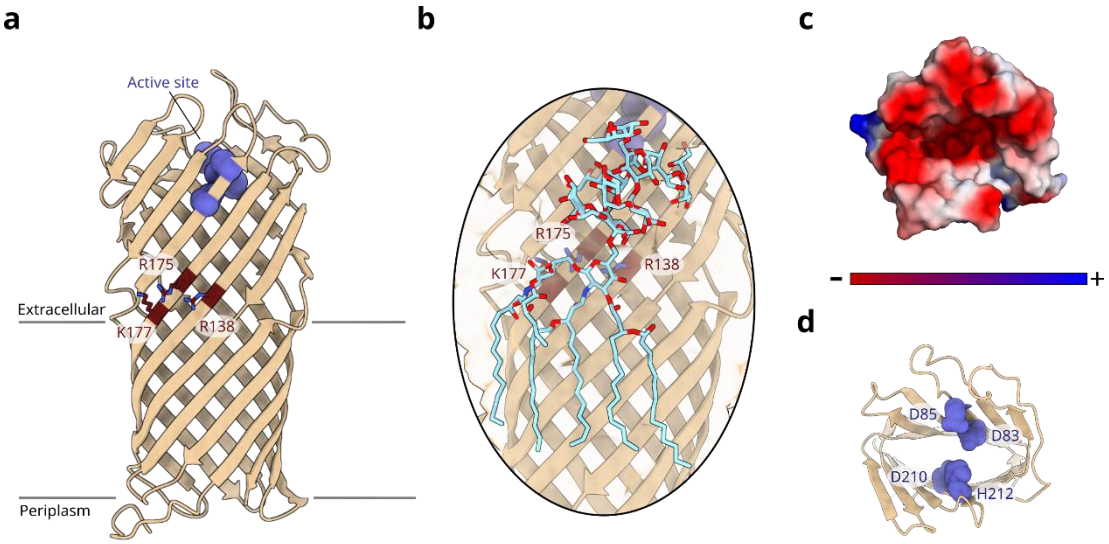


Figure 1.17: OmpT is an LPS-activated OM protease that cleaves positively charged antimicrobial peptides. (a) OmpT is a ten-stranded barrel with large extracellular loops housing the active site (blue) in the upper cleft of the barrel. The positive residues of the LPS binding motif are shown (red). **(b)** Rigidly docking the LPS molecule (cyan) resolved in the FhuA crystal structure to the equivalent residue motif on OmpT (red, three positive residues). **(c)** Electrostatic profile of the internal face of the upper cleft of OmpT. **(d)** The active site of the enzyme, showing the catalytic residues (blue). Both (c) and (d) are top-down views from the extracellular facing side of the protein. (PDB: 1I78[363], LPS from 1QFF[381]).

OmpT is the archetype of the broader bacterial omptin protease family. Indeed, clinal *E. coli* isolates typically have multiple omptin proteases including arlC and the complementation plasmid located OmpP, which have varied specificities and efficiencies[374] (although still targeting dibasic-motifs and showing LPS dependence[382,383]), yielding greater resistance benefits[383]. Although biochemically similar[384], omptin proteases exhibit a diverse range of functions depending on environmental and infection niche. *Yersinia* protease Pla, implicated (and diagnostically characteristic) in virulence of bubonic plague causing *Yersinia pestis*[385,386], cleaves blood circulating plasminogen to its active from plasmin and proteolytically inactivates plasmin's inhibitor α2-antiplasmin[387]. In addition, Pla is a lamin adhesin, targeting plasmin's protease activity to basement membranes and thus facilitating bacterial transport through tissue barriers[388]. Intriguingly LPS from *Y. pestis* grown at host 37 °C activates OmpT more than LPS when grown at 25 °C, suggesting temperature induced LPS changes help potentiate Pla-mediated proteolysis[389]. Other omptin proteases have characterised virulence importance: PgtE of *Salmonella*

enterica degrades AMPs, SopA of *Shigella flexneri* is important during intracellular phases of salmonellosis and shigellosis[390], and CroP from murine pathogen *Citrobacter rodentium* specifically targets murine AMPs[391,392]. Beyond human infections, OmpT and its homologs are known to be important for pathogenesis in a broad range of economically important organisms, including canine[393], avian[394], porcine[395], bovine[396], murine[391] and some crop plant hosts[397].

OmpLA

OmpLA (or PldA) is a widespread outer membrane phospholipase[398], responsible for degrading phospholipids that erroneously locate to the outer leaflet of the OM and thus in concert with PagP (see below) and the Mla pathway maintains OM asymmetry[399]. An early 7.4 Å 2D electron crystallography structure indicated that OmpLA oligomerises in the membrane plane[400], and an X-ray crystallography structure demonstrated a 14-stranded barrel with clear dimerization (**Fig. 1.18a**), mediated by transmembrane protein regions, although, unusually for transmembrane oligomers, the interaction strength does not correlate with occluded surface area[401]. Both enzyme dimerization[402,403] and divalent cations[404] are required for OmpLA activity, with the enzyme forming a N156-H142-S144 catalytic triad[405–407], with the divalent cation forming part of the active site and helping stabilise the oxyanion intermediate[404] (**Fig. 1.18b**). While the divalent cation bridges the dimerization interface, it contributes little to complex stability which is driven by interactions with the substrate acyl chain[408]. OmpLA solubilised in detergent has been rigorously kinetically characterised using synthetic substrates[409], but determining activity in synthetic membranes is challenging due to rapid degradation of the lipid bilayer upon enzyme activation[410]. The enzyme is promiscuous to lipid headgroups, but has a strong preference for acyl chains of 14 carbons or longer, with these longer chains better able to support dimerization[411]. Together, this suggests that OmpLA preferentially targets phospholipids that are longer than the acyl chains of LPS and hence more likely to cause hydrophobic mismatch and membrane defects. Insufficient divalent cations lead to membrane stress and induces phospholipid flip-flop[412], but the lack of ions will keep OmpLA inactive, presumably because under these conditions degrading phospholipids would further damage an already weakened membrane. Indeed, unsurprisingly, OmpLA activity in context of severe membrane disruption is harmful[413].

OmpLA expression is upregulated by a diverse range of stimuli that can disrupt OM stability in order to respond to phospholipid mislocalisation, including high-temperatures[414], AMPs[415], EDTA[416] and phage DNA insertion[417]. Intriguingly, it has been shown recently that differential leaflet stress in phospholipid asymmetric liposome models can also modulate OmpLA activity[88], perhaps sensing the changes in membrane properties when phospholipids migrate to the outer leaflet. OmpLA activity is directly coupled to upregulation of LPS biosynthesis via reuptake of the released fatty acids and lysophospholipids to the cytoplasm[418,419], where they are thioesterified to coenzyme A. The resulting acyl-CoA then interacts with FtsH, preventing degradation of LpxC, which catalyses the first committed step of LPS biosynthesis[420]. Thus, OmpLA couples with the Mla and LPS biosynthesis systems to maintain a healthy, asymmetric OM.

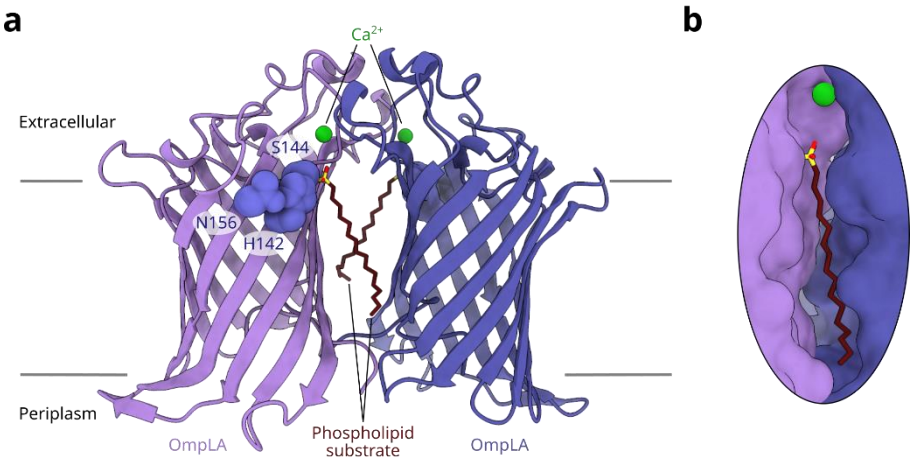


Figure 1.18: OmpLA is an OM phospholipase that requires divalent cations and homo-dimerisation for activity. (a) Overview of the 14-stranded OmpLA homodimer structure (purple and blue monomers), showing the active site residues (blue spheres), the lysophospholipid post-cleavage in its covalent acyl-enzyme intermediate (red) and the required calcium ions. Note an active site is formed on both sides of the dimer. **(b)** The substrate binding cleft is formed by the dimerisation interface (coloured as in (a)) (PDB: 7EZZ[421]).

OmpLA homologs are found across a broad range of pathogenic bacteria and are implicated in diverse behaviour. The enzyme appears particularly important in *Campylobacter*, where is has been associated with virulence[422], cell-associated hemolysis[423], cell invasion[424] and cecal colonisation[425]. Intriguingly, under anaerobic conditions the OM of *Campylobacter jejuni* contains large amounts (~35-45%) of lysophospholipids, required for efficient motility, suggesting that OmpLA activity may be upregulated for this functionality[426]. In *H. pylori* OmpLA is implicated in colonisation

and persistence[427], but it has also been shown to act as an important channel for urea influx and ammonium efflux in the acid adaptation response[428]. In addition, OmpLA is typically found upregulated in clinical isolates of *Neisseria gonorrhoeae* and meningococcal strains[429], *Arcobacter*[430], *Shigella*[431] and the animal pathogen *Riemerella anatipestifer*[432].

PagP

PagP, first identified as a PhoPQ-activated gene (Pag)[433], is a palmitoyltransferase that transfers a palmitate group from a glycerophospholipid to LPS[434] (**Fig. 1.10a**) (it also displays a mild phospholipase activity in the absence of LPS[435,436]). LPS-palmitoylation is important for mediating resistance to AMPs and other membrane stressors (such as high divalent ion concentrations) by altering the membrane properties. ΔPagP strains have increased membrane permeability under stress[437], while PagP activity generates a more robust permeability barrier[438]. PagP is an eight-stranded barrel with an amphipathic N-terminal helix and sits in the membrane at a tilt of ~30° relative to the membrane plane, with the catalytically important H33, D76 and S77 residues at the outer-leaflet membrane interface[439,440] (**Fig. 1.19a**). Importantly, there are two regions of proline-induced weaker inter-strand hydrogen-bonding (β1-2 and β6-7) around the barrel to facilitate substrate (LPS and phospholipid, respectively) approach[434,441] (**Fig. 1.19a**, dark green). While the formation of an acyl-enzyme intermediate has been proposed[442], there is no direct evidence for this and direct palmitate transfer via a ternary complex is more likely[434]. PagP maintains its high specificity to palmitate groups, important to manipulate the OM properties controllably as required, via a structural motif known as the hydrocarbon ruler (**Fig. 1.19b**). A precisely sized groove, PagP natively only accommodates the 16-carbon palmitate group, but mutation at G88 can adjust its specificity (G88A: 15-carbon, G88M: 12-carbon[443]) while site-specific alkylation allows excellent control of substrate use[444]. Although the exact mechanism is unknown, single mutations can alter the balance between the palmitoyltransferase and phospholipase activity (the W78F mutant has ~20x faster phospholipase activity[445], while K42A is only a phospholipase[444]) suggesting a fine-tuned process to stabilise the transition state of cleaved phospholipid with minimal release (i.e. phospholipase activity) in the absence of the acceptor LPS. The N-terminal helix is not required for activity, but biochemical and MD data indicate is stabilises and clamps the membrane-inserted protein after folding[440,446].

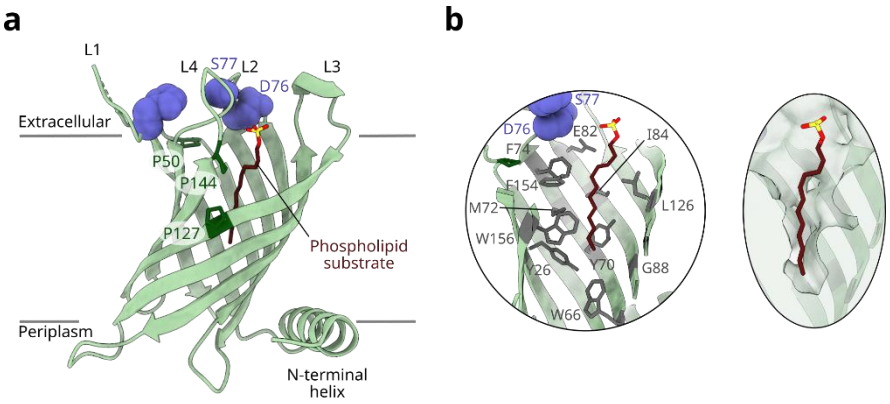


Figure 1.19: PagP is an OM LPS-palmitolytransferase with high acyl chain specificity. (a) PagP is an eight-stranded barrel that sits at a ~30° tilt in the membrane (relative to membrane normal), anchored by its N-terminal helix. Active site residues sit at the extracellular leaflet interface (blue spheres). Substrate active sites are between the extracellular loops of the protein, structured by strand breaks induced by prolines (dark green). An SDS molecule is shown in the phospholipid binding site. **(b)** The hydrocarbon ruler ensures the specificity of PagP by binding palmitate residues in a precisely formed cavity that cannot properly accommodate acyl chains with +/- one carbon. G88 (bottom right on left cutaway) can be mutated to alter specificity (coloured as in (a)) (PDB: 3GP6[441]).

Given that the access routes to the enzyme are exposed to the outer leaflet of the OM (and allows the phospholipid headgroup to remain external to the membrane core[441]), PagP activity relies on the aberrant presence of phospholipids in the outer leaflet of the OM. Thus, the enzyme remains inactive until membrane perturbations provide a phospholipid substrate[412]. This represents a key regulatory mechanism to ensure that PagP is only activated at OM regions with perturbed lipid asymmetry. Indeed the presence of phospholipids in the outer leaflet likely facilitates more rapid diffusion and thus more efficient LPS adaptation, while the less flexible hepta-acylated LPS would promote a more rigid membrane and minimise phospholipid flip-flop. PagP action is also more broadly linked to OM biogenesis, with sensing of *c*onstitutive palmitoylation of LPS causing an elevated σ^E response (without altering OMP composition/amount)[447]. In addition, the feedback mechanism outlined above for OmpLA-mediated increased LPS synthesis could also plausibly operate from the lyso-phospholipid PagP product. As well as responding to environmental challenges, during biofilm formation, lipid A palmitoylation increases *in vivo* survival[448].

PagP homologs are common across LPS-diderms but, similar to OmpT and OmpLA, have diverse virulence related functions. In *Bordetella*, PagP has been implicated in respiratory tract persistance[449] and resistance to antibody-mediated complement lysis[450], while the *Legionella* homolog, Rcp, is important for AMP resistance and intracellular infection[451]. Unlike in *Enterobacteriales*, *P. aeruginosa* palmitoylates at the 3' (rather than enterobacterial 2') position and the palmitoylated-LPS elevates the inflammatory response[452], while in *Bordetella parapertussis* PagP mediates the addition of two palmitates[453]. Although frequently beneficial for membrane properties during infection, hepta-acylated LPS activates host Toll-like Recoptor-4 mediated immune system activation 10-100 fold more than reduced acylated forms of LPS[454,455]. Indeed, some bacteria, like *Salmonella*, in addition to PagP contain PagL, an enzyme that deacylates LPS to help avoid immune system detection when required[176,456,457], while early evolutionary loss of PagP in *Y. pestis* resulted in innate immune system avoidance (synthetically adding it back in is more immune activating)[458]. However, the adaptive importance of this LPS modification is highlighted by the convergent evolution of LPS-palmitoylation by an alternative enzymatic pathway in *Acinetobacter* (via LpxLAb and LpxMAb)[459].

1.4.5 OM regulation

The complexity of the outer membrane and its biogenesis, combined with its role in first responding to cellular insults, necessitates regulatory networks. Mechanisms to maintain OM asymmetry via the Mla pathway and the actions of OmpLA and PagP are described above. Further, the actions of the three OM enzymes of *E. coli* are critical for rapid response to stress, harmful extracellular substances, and environmental changes, with them acting in concert to defend against a diverse set of possible threats[434], many of which were outlined in Section 1.4.4. Adaptive responses are mediated via stress response pathways in the cell envelope, notably the σ^E response and the Rcs and Cpx pathways[460,461]. Additional cell envelope stress responses are known, like the Bae[462] and Psp[463] systems which respond to exposure to toxic molecules and severe IM damage (including OMP localisation to the IM) respectively, but these have less direct impact on the OM.

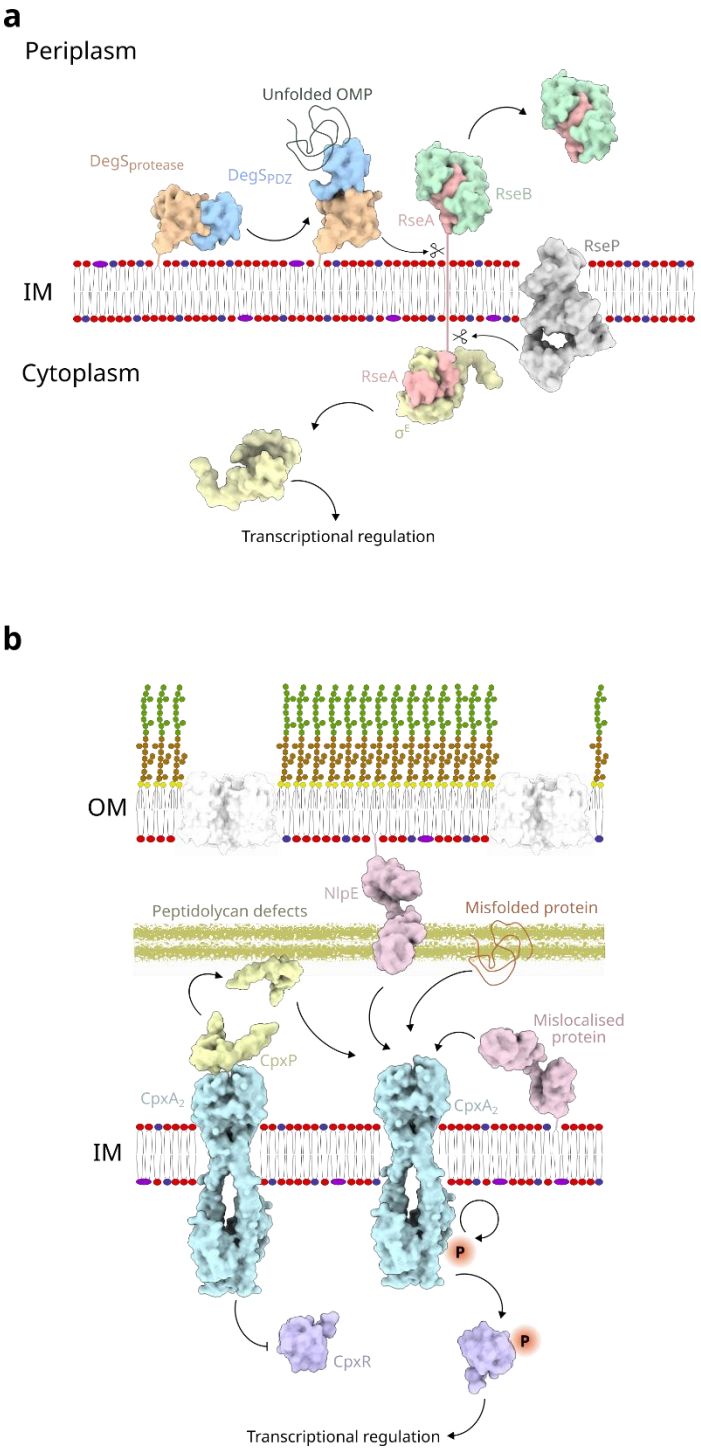


Figure 1.20: Cell envelope stress responses regulate OM composition and biogenesis. (a) The σ^E response recognises misfolded proteins or mislocalised LPS in the periplasm and activates a protease cascade that crosses the IM via RseA (not all proteolysis steps shown for clarity). Ultimately, the released σ^E transcription factor can interact with RNA polymerase and regulate gene expression. **(b)** The Cpx stress pathway is a canonical two-component signalling pathway with a diverse set of activators (including peptidoglycan defects, misfolded/mislocalised proteins and defects in lipoprotein biogenesis). Once activated $CpxA_2$ can autophosphorylate and phosphorylate effector CpxR to mediate transcriptional level regulation.

The σ^E response predominantly detects defects in OMP biogenesis via an increased concentration of unfolded/misfolded OMPs[464] or an excess of LPS in the periplasm[465] which induce a proteolytic cascade via the IM lipoprotein DegS, which ultimately leads to release of the σ^E transcription factor to interact with RNA polymerase[461] (**Fig. 1.20a**). σ^E alters transcription of a range of targets, including

downregulating general OMP expression and upregulating BAM, along with the OMP chaperones SurA, Skp and FkpA, and the protease DegP[338]. These act in combination to remove misfolded OMPs in the periplasm by degradation (DegP), chaperoning (SurA, Skp, FkpA, DegP) and increased folding rates (BAM)[466]. σ^E also downregulates the highly abundant Lpp and OmpA, which reduces demand on both the Lol and BAM folding pathways and facilitates increased production of Lpt pathway component lipoproteins, which together act to reduce the σ^E inductive event[466,467]. The σ^E response is constitutively active and both its elimination[468] and excessive upregulation[469] are lethal (in *E. coli*), indicating that OM biogenesis remains in a constant fine-tuned balance, and it is particularly important for handling the additional stresses that come during host infection[460].

The Cpx (conjugative pilus expression[470]) response is a well-conserved two-component signalling system composed of CpxA in the IM and CpxR in the cytoplasm, coupled to the (non-essential) OM stress-sensing lipoprotein NlpE[461] (**Fig. 1.20b**). Similar to σ^E, Cpx also responds to periplasmic OMP misfolding by upregulating chaperones like CpxP and Spy, the protease DegP and the oxidoreductase DsbA (and, mysteriously, downregulating σ^E), but is thought to have a broader activation profile[471,472]. For example, in distinction to σ^E, NlpE is thought to enable the Cpx response to detect errors in lipoprotein biogenesis[473]. Unlike σ^E, which is typically required for virulence across a broad range of pathogens, Cpx deletion shows mixed behaviour being strictly required in some organisms and only minorly detrimental in others[474–476].

The Rcs (regulator of capsule synthesis) pathway is an additional two-component signalling system, composed of IM RcsC/D and cytoplasmic RcsB, as well as OM lipoprotein component RcsF which is responsible for OM (and likely peptidoglycan) stress sensing[461]. The Rcs pathway is a highly flexible response as RcsB can form heterodimers with multiple partners, depending on cellular availability, and thus have various effects, as required[477]. The exact role, localisation and mechanism of RcsF in sensing remains controversial, although it appears that the protein can partially localise extracellularly, possibly allowing sensing of LPS defects, BAM function, or an (as yet) unknown lipoprotein export mechanism[478–480]. Regardless of its exact role, RcsF has defined interactions with both BamA and OMPs, and is necessary for the detection and transduction of most activators of the Rcs system[478,480].

1.4.6 OM organisation

The outer membrane components (LPS, phospholipids, OMPs/lipoproteins) are trafficked and inserted into the OM by a set of complex cellular machineries, as described above, in a highly orchestrated manner. The unique features of the different components, in particular the nature of LPS and β-barrel OMPs, and the enrichment of OMPs in the membrane creates an equally unique membrane[481], characterised by very low lateral diffusivity[140], excellent barrier properties[159] and ordered protein arrays[161] (modelled in **Fig. 1.21a**). Yet, the membrane still remains responsive to the bacteria's environmental context. Indeed, while the OM largely appears to remain a lipid bilayer matrix embedded with proteins, as postulated by the fluid mosaic model, large tracts of the membrane have essentially no lateral fluidity[274,482]. Interestingly, *Thermotoga* appear to have an outer sheath composed largely of genuinely lipid-free β-barrel proteins in an ordered array combined with lipid patches and α-helical proteins[483–485], and it has been proposed that this organisation represents the diderm ancestral form[486].

In recent years, direct visualisation of the OM via AFM has transformed our understanding of the cell surface architecture (**Fig. 1.21b-d**). Early AFM data demonstrated that the highly abundant trimeric porins could self-organise into 2D hexagonal arrays in the membrane with minimal intervening lipid[487,488], although correct formation of the base trimeric unit (probably) requires LPS binding at the subunit interfaces[489] (**Fig. 1.22a**). Whole cell imaging (mostly of *E. coli*) visualises the mosaicity of the OM, with much of the surface covered by the hexagonal array of trimeric porins[490], which, at some locations, appear to be at a higher concentration, termed OMP islands, and gaps between the porin array, interpreted as LPS-dominated patches[275,490] (**Fig. 1.21b-c**). Live-cell time-resolved imaging of the LPS patches demonstrated their ability to diffuse slowly across the bacterial surface and merge together which, together with mutational analysis, confirmed their interpretation as lipid[275]. *In vivo* crosslinking of OMPs to LPS (via inclusion of photoactive non-natural amino acids) showed that most OMPs are interacting with LPS, suggesting that even in the OMP islands there is at least a partial annulus of lipids around the proteins[491] (the low LPR of the OM means that it is probably impossible for all OMPs to be surrounded by LPS). The ultrastructure of the OM, as revealed by AFM, provided explanations both for the very slow diffusion of OMPs in the OM (typically on the order of the growth rate of *E. coli*[139,492]) and the corralled diffusion of protein components, which are mostly able to only diffuse within a small microevironment in the membrane[493]. Slow diffusion is also contributed to by the relative immobility of the

large LPS and the LPS oligosaccharides associating together via divalent cations[494], forming immobile blocks in the membrane. (Simulations considering OMP lattice formation confirmed that stable lattices only formed if LPS, not phospholipid, was present in the OM[491]).

Given how slow diffusion occurs in the OM, once inserted, macromolecular components are restricted to their membrane locale[493]. This is functionally helpful for the bacteria as it provides a mechanism for OMP turnover, without needing to construct complicated and energetically expensive machinery to extract OMPs from the OM. Indeed, the large number of hydrogen bonds between β-strands make OMPs highly stable (*in vitro* folding free energy estimated at around -13 to -32 kcal/mol (up to -140 kcal/mol reported)[43,495,496]) and the plausibility of unfolding them for extraction is questionable (certainly, no known processes can either unfold or extract OMPs from the membrane once they are inserted)[31]. Consequently, following their insertion OMPs have been shown to drift to the cellular poles during bacterial elongation and cell division, allowing complete change of the OMPome over about three generations, while the older OMPs are retained in a subset of the population (facilitating rapid population adaptation to environmental changes), in a process termed binary partitioning[140,497]. In addition, regulation of OMV formation has also been proposed as a mechanism to rapidly change the OMPome identity[498,499].

BAM itself has been shown to organise into BAM islands in the OM, termed folding precincts, where in clusters of BAM have been visualised as fluorescently labelled puncta, with their assembly dependent on the presence of BamB[252]. While this suggests an origin of OMP islands (i.e. recently inserted OMPs around a folding precinct), it is unclear where BAM is located on the OM surface relative to OMP islands and LPS patches[161,252]. Recent structural data studying OMP folding on BAM have suggested that the lipid membrane tension is important for folding[228,500], indicating that there is likely to be at least some lipids adjacent to BAM. BAM puncta are observed across the OM[140,252], but there is some evidence of increased activity of BAM at the midcell (which would enhance the rate of binary partitioning), at least for certain TBDTs[501]. This is supported by a recently identified BAM activating role for maturing peptidoglycan[501,502] which is mostly located at the zones of cell growth around the midcell, although other OMPs have been observed inserting all over the cellular membrane[161,252] (and SecYEG shows no evidence of polarisation in the IM[503]).

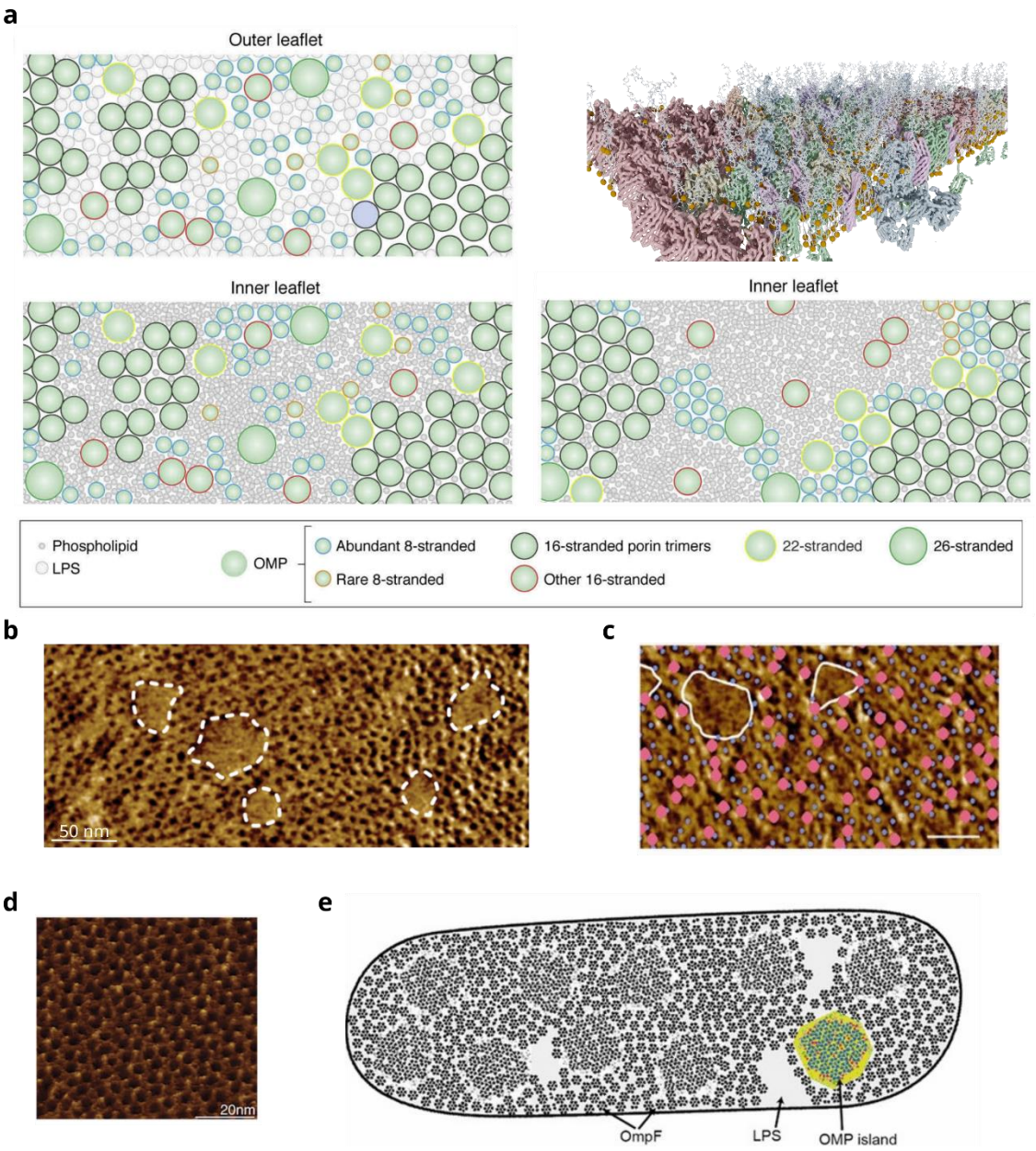


Figure 1.21: Global OM organisation is dominated by separate protein and lipid phases. (a) Two 2D-models of the OM leaflets (upper left and lower) with realistic LPRs and different degrees of protein clustering. Upper right model is 3D representation of the upper left 2D representation. **(b)** and **(c)** AFM images of the OM showing that LPS partially phase separates in the OM (white outlined regions) to form regions with minimal trimeric porins (black dots in (b), grey dots in (c)) and enriched in LPS (white outlines). Pink dots in (c) are the TBDT FepA. Scale bars are 50 nm. **(d)** High resolution AFM of the trimeric porin network in live cells is *R. denitrificans*. **(e)** Overview model of the current conceptual understanding of OM organisation, showing the matrix of trimeric porins (OmpF) that cover much of the cellular surface and regions enriched in OMPs (OMP islands) and LPS. (2D models in (a) adapted from [31], AFM images in (b) [275], (c) [491], and (d) [490] and cellular model (e) from [491]).

LPS-OMP interactions are ubiquitous[491], and as such it is expected that OMPs have evolved to function optimally given the presence of LPS and/or have specific LPS binding sites that modulate their function. Indeed, *in vivo* EPR data have revealed novel structural states for the TBDT BtuB that are not occupied in membranes *in vitro* in the absence of LPS[345,504]. This is supported by MD data showing extracellular loop ordering and thus altered substrate binding[505]. Cellular EPR has also demonstrated that most BtuB-BtuB interactions in the membrane occur via intervening LPS[506]. As noted above, OmpT has a well-defined activation in the presence of LPS, and a predicted LPS binding site[363] (Section 1.14, **Fig. 1.18b**). A few OMPs have unambiguously structurally resolved LPS, notably OmpE36 (*E. cloacae* trimeric porin, crystallography[489], **Fig. 1.22a**) and the TBDT FhuA (crystallography/cryoEM[381,507], **Fig. 1.22b**). The structural impacts of LPS binding have been determined on OprH by NMR, but the lipid itself was not observed[508]. Likely, although not unambiguously assignable, LPS density has been observed in other models, including as a component in BAM nanodiscs[228], and predicted for a broad range of OMPs via MD[505,509]. Although still relatively sparse, together these data combine to highlight LPS mediated functional modulation or optimisation of multiple OMPs, and as emerging methods make it easier to study OMP-LPS interactions *in vitro* and OMPs *in vivo*, new roles of LPS are sure to be discovered.

In addition to interacting with LPS, OMPs must interact with other OMPs, a necessity due to the low LPR, and LPS depletion from OMP islands. Given that over two thirds of the protein composition of the *E. coli* OM is formed from OmpA and OmpC/F[273], the majority of other OMPs in the membrane must be interacting with one of these proteins, either directly or mediated by intervening lipid, while the extremely limited diffusion ensures these interactions will be long-lived even if low affinity. Some OMP-OMP interactions have been described, including those known to functionally homo-oligomerise such as OmpLA[408], OmpA[289] and BAM dimers[252], and porin trimers[489], and hetero-oligomerisation, for example BtuB-OmpF[36], MlaA-OmpC[196] (Section 1.4.3) and RcsF-OmpA[478]. It is notable that all these instances include one of the abundant OMPs, which are the proteins predicted as most likely to evolve additional protein-protein interaction functionality. However, despite these examples, it is challenging to model OMP-OMP interactions in *in vitro* membrane systems where, by necessity, the protein concentrations are significantly lower than they are *in vivo* (as well as other features, like LPS, typically being absent). MD is also unable to access timescales long enough to be substantially informative regarding OMP-OMP interactions in a realistic membrane[481] and thus much remains unknown about non-

canonical OMP-OMP interactions.

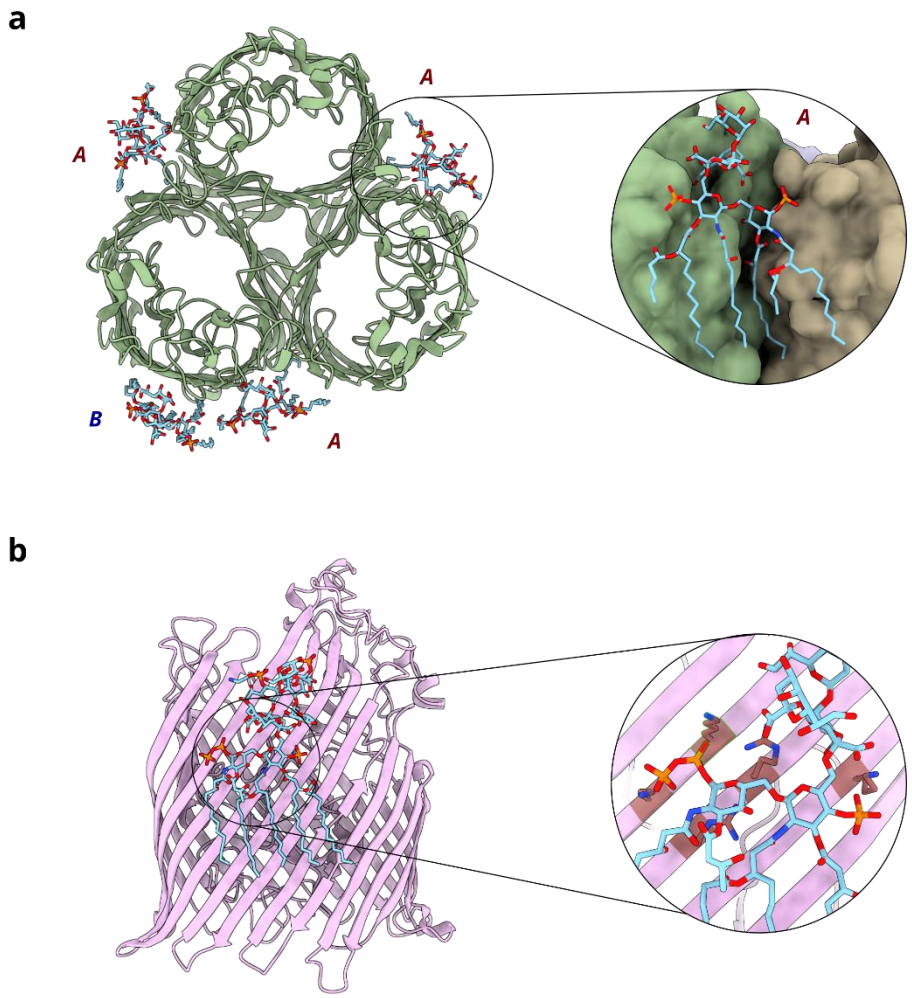


Figure 1.22: Several OMP-LPS interactions have been structurally resolved. (a) LPS was observed in a crystal structure of the *E. clocae* trimeric porin OmpE36 at two distinct sites: (A) at each dimer interface, where the LPS bridges the two proteins with one of the acyl chains sitting deeply with the groove formed by the interaction, and (B) away from the oligomerisation interfaces resolved at only one site (PDB: 5FVN[481]). **(b)** The TBDT FhuA has a well-defined LPS binding site characterised by a network of positively charged residues interacting with lipid A's phosphate groups (PDB: 1QFF[381]).

1.5 The role of the OM in health and disease

LPS-diderm bacteria play critical roles in human health and food security, both dominating pathogenic watchlists, as well as playing significant roles in gut

microbiomes and biotechnology applications. Indeed, 9 of the 12 bacteria identified by the World Health Organisation as being critically antibiotic-resistant are LPS-diderm, as well as 3 of 5 CDC (Centre for Disease Control) Urgent Threats[510], and 4 of the 6 well-known ESKAPE pathogens[511] (named by taking the first letter of the genus of the 6 pathogens). In 2019, globally antibiotic resistance was estimated to cause 1.3 million deaths (and an additional 5 million associated deaths) with about 75% attributable to diderm bacteria[512], with the scale and scope of the problem rapidly expanding due to resistance evolution. The additional protection offered by the unique OM coupled with the double membrane architecture of LPS-diderms pose significant challenges for antibiotic targeting[173,274]. Larger substances (approximately >600 Da[159]) are excluded almost entirely, and the extra compartmentalisation of the periplasm provides additional opportunities and time for bacterial machinery to degrade or export antibiotics[513–515], preventing lethal concentrations from accumulating. Although some traditional antibacterials target the OM, like polymyxin, they also damage host membranes and are thus typically reserved as antibiotics of last resort[516]. Given these challenges, antibiotics that specifically target the OM directly are highly desirable, and in recent years novel lead antibiotic compounds that target BAM have been discovered[517–519]. Perhaps most promising is darobactin and its related family of molecules[517,520], which are stapled heptapeptides that target BAM by binding BamA's β1 thus locking the barrel closed and inhibiting substrate binding[521,522], although much additional work is required to move it into the clinic.

In addition to the antibiotic resistance crisis, LPS-diderm bacteria modulate human health in non-infectious contexts via the gut microbiome. The human gut is colonised by a vast quantity ($10^{11\text{-}12}$ per gram[523]) of generally symbiotic, anaerobic bacteria and their disruption or imbalance is linked to many diseases[524]. Although by cell quantity there are estimated to be more monoderm bacteria in the gut (dominated by *Firmicutes)*, in the ~500-1000 species present there is vastly more diderm diversity represented (enriched in *Bacteroidetes, Actinobacteria and Proteobacteria)*[525,526], and they are thought to perform an equally diverse set of roles, including supporting food degradation[527], vitamin synthesis[528] and an emergent endocrine function of the population[529], particularly important in modulating the gut-brain[530] and gut-skin[531] axis. Indeed, with the gut microbiome collectively having ~100-times more unique genes than humans, their importance for adaptation and supplying additional functionality is highlighted[532]. Despite the microbiome's importance, the implications of antibiotics, other drugs, diet and lifestyle on the health and balance of these populations are poorly understood[524,533–535]. Given the role of the OM as the extracellular facing barrier

for diderms, it is clear this plays an important role both in modulating individual bacterial health and mediating population interactions and communication.

1.6 Project aims

The OM is a unique membrane that is vital for diderm bacterial survival, and remarkably is constructed and maintained despite the lack of an adjacent energy source. However, the OM is also a product of diderm bacteria's specific lifestyle. In particular, the lack of a periplasmic energy source, the single-cell organisation and tight correlation between cell division and resource availability limits many of the OM's potential dangers for bacterial populations (e.g., no active OMP turnover, minimal lateral diffusion). Despite much progress in understanding the OM, much remains unclear, in particular the synergy between its different constituents. For example, LPS-OMP interactions, the membrane structure at the junctions of LPS patches and OMP islands, OMP-OMP interactions within arrays, lipoprotein localisation and translocation, S-layer modulation and the role of the OM's superstructure and asymmetry on OMP folding. This project explores some of these open questions. Each chapter recapitulates a specific native feature of the OM *in vitro*: membrane charge asymmetry (Chapter 3), LPS-OMP interactions (Chapter 4) and OMP-OMP interactions (Chapter 5), and each characterises its implications on either the folding or the function of model OMPs (**Fig. 1.23**). Together these results yield new insights into how the unique features of the native OM modulate its form and function.

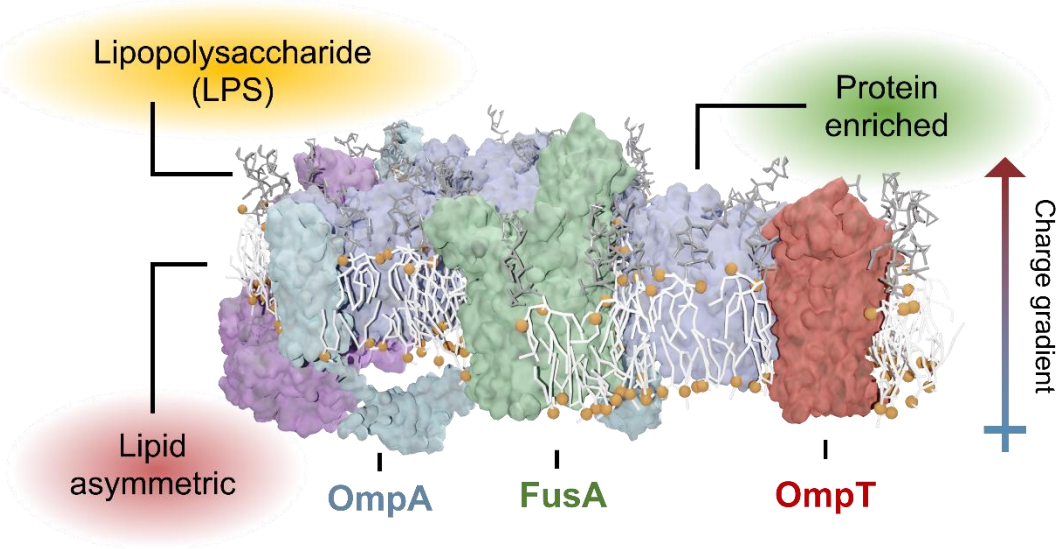


Figure 1.23: Summary of the research presented in this thesis. The OM is a unique membrane, in part due its extreme lipid asymmetry (which induces a charge gradient over the membrane), the presence of LPS in the outer leaflet of the bilayer and its very low LPR. The implications of each of these membrane features on OMP folding or function is explored, predominantly using the model OMPs OmpA, FusA and OmpT.

Chapter 2: Materials and Methods

2.1 General Materials

Table 2.1: General Chemicals and Materials

Chemical	Supplier
Acetic acid, glacial	Fisher Scientific, Loughborough, UK
Acrylamide 30 % *(w/v)*:*bis*-acrylamide 0.8 % *(w/v)*	Severn Biotech, Kidderminster, UK
Agar	Melford Laboratories, Suffolk, UK
Agarose	Melford Laboratories, Suffolk, UK
Ammonium persulphate (APS)	Sigma Life Sciences, MO, USA
β-mercaptoethanol	Sigma Life Sciences, MO, USA
Benzamidine hydrochloride	Sigma Life Sciences, MO, USA
Bromophenol blue	Sigma Life Sciences, MO, USA
Carbenicillin (disodium)	Formedium, Norfolk, UK
Chloroform	Fisher Scientific, Loughborough, UK
D-Tube dialyzer mini, 12-14 kDa MWCO	EMD Millipore, MA, USA
Dithiothreitol (DTT)	Formedium, Norfolk, UK
DNA ladder (1 kbp, 100bp)	New England Biolabs, MA, USA
n-dodecyl β-D-maltoside (DDM)	Anatrace, OH, USA
Lauryl-dimethylamine oxide (LDAO)	Anatrace, OH, USA
Ethanol	Sigma Life Sciences, MO, USA
Ethylenediaminetetraacetic acid (EDTA)	Fisher Scientific, Loughborough, UK
Mini, EDTA-free protease inhibitor cocktail tablets	Roche Applied Science
Guanidine hydrochloride	Sigma Life Sciences, MO, USA
Glycerol	Fisher Scientific, Loughborough, UK
Hydrochloric acid (HCl)	Fisher Scientific, Loughborough, UK
Imidizole	Sigma Life Sciences, MO, USA
Instant Blue Coomassie Blue Stain	Expedeon, CA, USA
Isopropyl β-D-1-thiogalactopyranoside (IPTG)	Formedium, Norfolk, UK
Butanol	Honeywell Research Chemicals, Seelze, Germany
Kanamycin	Formedium, Norfolk, UK

Laurdan	Cayman Chemical, Ann Arbor, MI, USA
LB Miller	Fisher Scientific, Loughborough, UK
Methanol	Fisher Scientific, Loughborough, UK
Protein molecular weight markers (Precision Plus Dual Xtra Standards)	Bio-Rad Laboratories, CA, USA
Phenylmethanesufonyl fluoride (PMSF)	Sigma Life Sciences, MO, USA
Polycarbonate membrane (0.1 µm)	Whatman Inc., NJ, USA
Potassium chloride (KCl)	Fisher Scientific, Loughborough, UK
Potassium hydroxide (KOH)	Fisher Scientific, Loughborough, UK
Sodium chloride (NaCl)	Fisher Scientific, Loughborough, UK
Sodium hydroxide (NaOH)	Fisher Scientific, Loughborough, UK
Sodium dodecyl sulphate (SDS)	Sigma Life Sciences, MO, USA
Triton X-100	Sigma Life Sciences, MO, USA
Tris	Fisher Scientific, Loughborough, UK
Tetramethylethylenediamine (TEMED)	Sigma Life Sciences, MO, USA
Tris(2-carboxyethyl)phosphine (TCEP)	Sigma Life Sciences, MO, USA
Tris-tricine SDS running buffer 10X, cathode buffer, pH 8.3	Alfa Aesar, Heysham, UK
Urea	MP biomedicals, Loughborough , UK
Vivaspin 20 centrifugal concentrators (3K, 5K, 30K, 50K, 100K MWCO)	Sartorius, Göttingen, Germany
Vivaspin 500 centrifugal concentrators (3K MWCO)	Sartorius, Göttingen, Germany
ZebaSpin Desalting Column 7K MWCO (0.5 ml)	Fisher Scientific, Loughborough, UK

All lipids were purchased from Avanti polar lipids (AL, USA). Kit assyas were performed according to the manufacturer's instructions: BCA protein assay kit (Pierce, Fisher Scientific), SDS-PAGE silver stain kit (Invitrogen, SilverXpress Silver staining kit), miniprep (QIAprep spin miniprep kit (Quiqagen)) and agarose gel extraction kit (New England Biolabs). Deionised 18 MΩ water was used for buffer/media preparation in all methods. All ratios are mol/mol unless otherwise stated.

2.2 Bacterial strains

Plasmid purification was performed in *E. coli* DH5α (F- φ80*lac*ZΔM15 Δ(*lac*ZYA-*arg*F) U169 *rec*A1 *end*A1 *hsd*R17(r_k^-, m_k^+) *pho*A *sup*E44 *thi*-1 *gyr*A96 *rel*A1 λ^-), originally purchased from Invitrogen. Protein expression was performed in *E. coli* BL21 (DE3) (F^- *dcm ompT hsdS*(rB^- mB^-) *gal* λ(DE3) originally purchased from Agilent (both cell lines made competent in-house).

2.3 Plasmids and protein sequences

All OMPs, OMP mutants and *de novo* designed transmembrane barrels are untagged (except FusA) and (for OMPs) are the mature sequences (no signal sequence with the addition of an initiating methionine) for cytoplasmic inclusion body expression.

2.3.1 Plasmids

OmpA, tOmpA (N-terminal transmembrane domain of OmpA, $OmpA_{1-171}$), PagP, BamA and OmpLA are in a pET11a backbone (Amp^R) and were obtained from Karen Fleming, John Hopkins University. OmpX (codon-optimised gene) was ordered and cloned into a pET11a backbone (Amp^R) by Dr Bob Schiffrin. OmpF was cloned into a pET11a backbone (Amp^R) by Dr Anton Calabrese. OmpT was cloned into a pET11a backbone (Amp^R) by Dr Lindsay McMorran. BtuB was cloned into a pBAD backbone (Amp^R) by Dr Sam Hickman. OmpTrans3 (transmembrane barrel of OmpA with short, designed loops), TMB2.17LA and TMB2.3LA (*de novo* designed barrels TMB2.3 and TMB2.17 with the extracellular loops of OmpA) in a pET29b backbone (Kan^R) was obtained from Dr Anastassia Vorobieva and Prof David Baker, University of Washington. MSP1D1 in a pET28a backbone (Kan^R) was obtained from Stephen Sligar (Addgene #20061)[536]. MipA and OmpA extracellular loop mutants (OmpA-NN/NC/NP/M3, see below table), all cloned into a pET11a backbone (Amp^R), were generated as part of this study and details are given in the relevant sections below (2.4.1).

2.3.2 Protein sequences

Table 2.2: Protein sequences of proteins expressed and used in this study

Protein	Sequence
tOmpA (OmpA1$_{1-171}$)	MAPKDNTWYTGAKLGWSQYHDTGFINNNGPTHENQLGAGAFGGYQ VNPYVGFEMGYDWLGRMPYKGSVENGAYKAQGVQLTAKLGYPITDD LDIYTRLGGMVWRADTKSNVYGKNHDTGVSPVFAGGVEYAITPEIAT RLEYQWTNNIGDAHTIGTRPDNGMLSLGVSYRFG
OmpX	MATSTVTGGYAQSDAQGQMNKMGGFNLKYRYEEDNSPLGVIGSFTY TEKSRTASSGDYNKNQYYGITAGPAYRINDWASIYGVVGVGYGKFQT TEYPTYKHDTSDYGFSYGAGLQFNPMENVALDFSYEQSRIRSVDVGT WIAGVGYRF
BamA	MEGFVVKDIHFEGLQRVAVGAALLSMPVRTGDTVNDEDISNTIRALFA TGNFEDVRVLRDGDTLLVQVKERPTIASITFSGNKSVKDDMLKQNLEA SGVRVGESLDRTTIADIEKGLEDFYYSVGKYSASVKAVVTPLPRNRVD LKLVFQEGVSAEIQQINIVGNHAFTTDELISHFQLRDEVPWWNVVGDR KYQKQKLAGDLETLRSYYLDRGYARFNIDSTQVSLTPDKKGIYVTVNIT EGDQYKLSGVEVSGNLAGHSAEIEQLTKIEPGELYNGTKVTKMEDDIK KLLGRYGYAYPRVQSMPEINDADKTVKLRVNVDAGNRFYVRKIRFEG NDTSKDAVLRREMRQMEGAWLGSDLVDQGKERLNRLGFFETVDTDT QRVPGSPDQVDVVYKVKERNTGSFNFGIGYGTESGVSFQAGVQQD NWLGTGYAVGINGTKNDYQTYAELSVTNPYFTVDGVSLGGRLFYNDF QADDADLSDYTNKSYGTDVTLGFPINEYNSLRAGLGYVHNSLSNMQ PQVAMWRYLYSMGEHPSTSDQDNSFKTDDFTFNYGWTYNKLDRGY FPTDGSRVNLTGKVTIPGSDNEYYKVTLDTATYVPIDDDHKWVVLGRT RWGYGDGLGGKEMPFYENFYAGGSSTVRGFQSNTIGPKAVYFPHQ ASNYDPDYDYECATQDGAKDLCKSDDAVGGNAMAVASLEFITPTPFI SDKYANSVRTSFFWDMGTVWDTNWDSSQYSGYPDYSDPSNIRMSA GIALQWMSPLGPLVFSYAQPFKKYDGDKAEQFQFNIGKTW

OmpA-WT	MAPKDNTWYTGAKLGWSQYHDTGFINNNGPTHENQLGAGAFGGYQ VNPYVGFEMGYDWLGRMPYKGSVENGAYKAQGVQLTAKLGYPITDD LDIYTRLGGMVWRADTKSNVYGKNHDTGVSPVFAGGVEYAITPEIAT RLEYQWTNNIGDAHTIGTRPDNGMLSLGVSYRFGQGEAAPVVAPAPA PAPEVQTKHFTLKSDVLFNFNKATLKPEGQAALDQLYSQLSNLDPKD GSVVVLGYTDRIGSDAYNQGLSERRAQSVVDYLISKGIPADKISARGM GESNPVTGNTCDNVKQRAAL IDCLAPDRRVEIEVKGIKDVVTQPQA
OmpA-NN (D41S, E53N, E89V, D126S, D137S, D180S, D189S)	MAPKDNTWYTGAKLGWSQYHSTGFINNNGPTHNNQLGAGAFGGYQ VNPYVGFEMGYDWLGRMPYKGSVVNGAYKAQGVQLTAKLGYPITDD LDIYTRLGGMVWRASTKSNVYGKNHSTGVSPVFAGGVEYAITPEIATR LEYQWTNNIGSAHTIGTRPSNGMLSLGVSYRFGQGEAAPVVAPAPAP APEVQTKHFTLKSDVLFNFNKATLKPEGQAALDQLYSQLSNLDPKDG SVVVLGYTDRIGSDAYNQGLSERRAQSVVDYLISKGIPADKISARGMG ESNPVTGNTCDNVKQRAALIDCLAPDRRVEIEVKGIKDVVTQPQA
OmpA-NP (R81S, K85T, K94S, R124S, K128G, K134S, R177S)	MAPKDNTWYTGAKLGWSQYHDTGFINNNGPTHENQLGAGAFGGYQ VNPYVGFEMGYDWLGSMPYTGSVENGAYSAQGVQLTAKLGYPITDD LDIYTRLGGMVWSADTGSNVYGSNHDTGVSPVFAGGVEYAITPEIAT RLEYQWTNNIGDAHTIGTSPDNGMLSLGVSYRFGQGEAAPVVAPAPA PAPEVQTKHFTLKSDVLFNFNKATLKPEGQAALDQLYSQLSNLDPKD GSVVVLGYTDRIGSDAYNQGLSERRAQSVVDYLISKGIPADKISARGM GESNPVTGNTCDNVKQRAALIDCLAPDRRVEIEVKGIKDVVTQPQA
OmpA-NC (D41S, E53N, E89V, D126S, D137S, D180S, D189S, R81S, K85T, K94S, R124S, K128G, K134S, R177S)	MAPKDNTWYTGAKLGWSQYHSTGFINNNGPTHNNQLGAGAFGGYQ VNPYVGFEMGYDWLGSMPYTGSVVNGAYSAQGVQLTAKLGYPITDD LDIYTRLGGMVWSASTGSNVYGSNHSTGVSPVFAGGVEYAITPEIATR LEYQWTNNIGSAHTIGTSPSNGMLSLGVSYRFGQGEAAPVVAPAPAP APEVQTKHFTLKSDVLFNFNKATLKPEGQAALDQLYSQLSNLDPKDG SVVVLGYTDRIGSDAYNQGLSERRAQSVVDYLISKGIPADKISARGMG ESNPVTGNTCDNVKQRAALIDCLAPDRRVEIEVKGIKDVVTQPQA

OmpA-M3 (R81S, R94S, R124S)	MAPKDNTWYTGAKLGWSQYHDTGFINNNGPTHENQLGAGAFGGYQ VNPYVGFEMGYDWLGSMPYKGSVENGAYSAQGVQLTAKLGYPITDD LDIYTRLGGMVWSADTKSNVYGKNHDTGVSPVFAGGVEYAITPEIATR LEYQWTNNIGDAHTIGTRPDNGMLSLGVSYRFGQGEAAPVVAPAPAP APEVQTKHFTLKSDVLFNFNKATLKPEGQAALDQLYSQLSNLDPKDG SVVVLGYTDRIGSDAYNQGLSERRAQSVVDYLISKGIPADKISARGMG ESNPVTGNTCDNVKQRAALIDCLAPDRRVEIEVKGIKDVVTQPQA
OmpTrans3	MGAPKDNTWYTGAKLGWNTDNTLGAGAFGGYQVNPYVGFEMGYD WNNSSLQGVQLTAKLGYPITDDLDIYTRLGGMVNTDNTVSPVFAGGV EYAITPEIATRLEYQWNNSSLGMLSLGVSYRFG
OmpF	MEIYNKDGNKVDLYGKAVGLHYFSKGNGENSYGGNGDMTYARLGFK GETQINSDLTGYGQWEYNFQGNNSEGADAQTGNKTRLAFAGLKYAD VGSFDYGRNYGVVYDALGYTDMLPEFGGDTAYSDDFFVGRVGGVAT YRNSNFFGLVDGLNFAVQYLGKNERDTARRSNGDGVGGSISYEYEG FGIVGAYGAADRTNLQEAQPLGNGKKAEQWATGLKYDANNIYLAANY GETRNATPITNKFTNTSGFANKTQDVLLVAQYQFDFGLRPSIAYTKSK AKDVEGIGDVDLVNYFEVGATYYFNKNMSTYVDYIINQIDSDNKLGVG SDDTVAVGIVYQF
PagP	MNADEWMTTFRENIAQTWQQPEHYDLYIPAITWHARFAYDKEKTDRY NERPWGGGFGLSRWDEKGNWHGLYAMAFKDSWNKWEPIAGYGWE STWRPLADENFHLGLGFTAGVTARDNWNYIPLPVLLPLASVGYGPVT FQMTYIPGTYNNGNVYFAWMRFQF
OmpLA	MQEATVKEVHDAPAVRGSIIANMLQEHDNPFTLYPYDTNYLIYTQTSD LNKEAIASYDWAENARKDEVKFQLSLAFPLWRGILGPNSVLGASYTQ KSWWQLSNSEESSPFRETNYEPQLFLGFATDYRFAGWTLRDVEMGY NHDSNGRSDPTSRSWNRLYTRLMAENGNWLVEVKPWYVVGNTDDN PDITKYMGYYQLKIGYHLGDAVLSAKGQYNWNTGYGGAELGLSYPIT KHVRLYTQVYSGYGESLIDYNFNQTRVGVGVMLNDLF

BtuB	MVKKASLLTACSVTAFSAWAQDTSPDTLVVTANRFEQPRSTVLAPTTV VTRQDIDRWQSTSVNDVLRRLPGVDITQNGGSGQLSSIFIRGTNASH VLVLIDGVRLNLAGVSGSADLSQFPIALVQRVEYIRGPRSAVYGSDAIG GVVNIITTRDEPGTEISGGWGSNSYQNYDVSTQQQLGDKTRVTLLGD YAHTHGYDVVAYGNTGTQAQTDNDGFLSKTLYGALEHNFTDAWSGF VRGYGYDNRTNYDAYYSPGSPLLDTRKLYSQSWDAGLRYNGELIKS QLITSYSHSKDYNYDPHYGRYDSSATLDEMKQYTVQWANNVIVGHG SIGAGVDWQKQTTTPGTGYVEDGYDQRNTGIYLTGLQQVGDFTFEG AARSDDNSQFGRHGTWQTSAGWEFIEGYRFIASYGTSYKAPNLGQL YGFYGNPNLDPEKSKQWEGAFEGLTAGVNWRISGYRNDVSDLIDYD DHTLKYYNEGKARIKGVEATANFDTGPLTHTVSYDYVDARNAITDTPL LRRAKQQVKYQLDWQLYDFDWGITYQYLGTRYDKDYSSYPYQTVKM GGVSLWDLAVAYPVTSHLTVRGKIANLFDKDYETVYGYQTAGREYTLS GSYTF
OmpT	STETLSFTPDNINADISLGTLSGKTKERVYLAEEGGRKVSQLDWKFNN AAIIKGAINWDLMPQISIGAAGWTTLGSRGGNMVDQDWMDSSNPGT WTDESRHPDTQLNYANEFDLNIKGWLLNEPNYRLGLMAGYQESRYS FTARGGSYIYSSEEGFRDDIGSFPNGERAIGYKQRFKMPYIGLTGSYR YEDFELGGTFKYSGWVESSDNDEHYDPGKRITYRSKVKDQNYYSVA VNAGYYVTPNAKVYVEGAWNRVTNKKGNTSLYDHNNNTSDYSKNG AGIENYNFITTAGLKYTF
MipA	EGKFSLGAGVGVVEHPYKDYDTDVYPVPVINYEGDNFWFRGLGGGY YLWNDATDKLSITAYWSPLYFKAKDSGDHQMRHLDDRKSTMMAGLS YAHFTQYGYLRTTLAGDTLDNSNGIVWDMAWLYRYTNGGLTVTPGIG VQWNSENQNEYYYGVSRKESARSGLRGYNPNDSWSPYLELSASYN FLGDWSVYGTARYTRLSDEVTDSPMVDKSWTGLISTGITYKF
TMB2.17LA	MEQKPGTLMVYVVVGYEQYHDTGFINNNGPTHENKVDVVGGAQYAV SPYLFLDVGYGWTGRMPYKGSVENGAYKKNFLEVGGGVSYKVSPDL EPYVKAGFEYERADTKSNVYGKNHDNRIKPTAGAGALYRVSPNLALM VEYGWKNNIGDAHTIGTRPDKQKVAIGIAYKVKD

TMB2.3LA	MQDGPGTLDVFVAAGWNQYHDTGFINNNGPTHENKIEITGGATYQLS PYIMVKAGYGWDGRMPYKGSVENGAYKRNRFEFGGGLQYKVTPDL EPYAWAGATYERADTKSNVYGKNHDNKLVPAAGAGFRYKVSPEVKL VVEYGWKNNIGDAHTIGTRPDKQFLQAGLSYRIQP
FusA	MGSSHHHHHHSSGLVPRGSHMQQNDTSADENQQKNNAESEEEQQ GDSSSGNSEDTILVRSTPTSQSMGMQIINAEQIKKMPTGNGSVTELLK NNPNVQFSNTASSSNIPGELAPENVSFHGEKFYNNNFMVDGLSNNN NINPGANNGELNQQPDGYSPADLPAGGPQSFWINSELIESLEVFDSNI SAKYGDFTGGVVDAKTMDPKLDKSSGKISYRTTRDSWTKYHISEVIS EEFYSGTNLYYQPKFKKHFYSATFNQPLSDKAGFIFAYNRQQSDIPYY HEYLQQWDDQERINETYLLKGTYLTDSGDIIRMTGMYSPHESKFYKK DVKNGGFTNSGGGYRFNMEWEHNASWGKMTSLAGYQYTEDKTEH EADSYQTWRRFSSGFVSNVIDWSSSGGANGQNSNIGGYGSFATNTS SFSLKQDYELNPVSWYGINHQIDFGWETDFYTSRYRRFSDVYTGGVL VVPTGASAGSVVCQSGDELCIPGEQYSRTRILYPERNVQVSNVNYAA YLQNSMSYGRLEVTPGVRVSYDDFLENLNIAPRFSASYDVFGDRSTR LFGGANRYYAGNILAYKMRQGIGSNIQESRISPTAPWTTPTLRTGTNY NVSDLNTPYSDELSLGLSQRVMSTVWTAKWVNRQGKEQFGRETTTI DGQSYRVMNNKGHTEGNTFSLEVEPISPHRFSFAEVNWKLGASVTK NKSNSIYYYDSANQDEQRVIFDDKLMYRGDIDAMNFNTPWRAFLNVN TYFPAVRLSWDQRVGYTAGYKGYTTSSIQVQCPGGSSACNGDPSFV GGATEYFPTQYDDFISYDWRFSYSQPVYKTQTLDITLDVLNVLDNVVE TNQTGTSNKPIVIYKPGRQFWLGVAYTW
MSP1D1	MGHHHHHHHDYDIPTTENLYFQGSTFSKLREQLGPVTQEFWDNLEK ETEGLRQEMSKDLEEVKAKVQPYLDDFQKKWQEEMELYRQKVEPLR AELQEGARQKLHELQEKLSPLGEEMRDRARAHVDALRTHLAPYSDE LRQRLAARLEALKENGGARLAEYHAKATEHLSTLSEKAKPALEDLRQ GLLPVLESFKVSFLSALEEYTKKLNTQ

2.4 General methods

2.4.1 Molecular biology

Agar plates were prepared with autoclaved and cooled 25 g/L LB and 15 g/L bacto-agar with the required antibiotic (100 µg/ml) and ~20 ml poured into plastic petri dishes. Following setting and drying under flame, plates were stored at 4 °C. Competent cells were transformed by thawing on ice, adding 1-3 µl plasmid DNA to 50 µl of cells and incubating on ice for a further 30 minutes. Cells were heat-shocked at 42 °C for 30 s, cooled on ice for 5 min and then added to 950 µl SOC media (super optimal catabolite, New England Biolabs) and allowed to outgrow for 1 h (37 °C, 200 rpm). For expression plasmid transformation ~100 µl was plated and for cloned plasmids cells were pelleted resuspended in 150 µl and ~100 µl plated. Plates were incubated overnight at 37 °C, and resulting individual colonies used to inoculate culture for plasmid purification or starter culture for protein expression (both then grown overnight, 37 °C, 200 rpm). Where required for plasmid purification, overnight cultures were pelleted and DNA extracted using a miniprep kit (Quiagen). DNA concentration was estimated by the A_{260} using a nanodrop 2000 (ThermoFisher).

DNA was analysed via agarose gel electrophoresis (1.5% (w/v) agarose in TAE buffer (40 mM Tris-Cl, pH 8.0, 20 mM acetic acid, 1 mM EDTA) mixed with 0.1% (v/v) of 1000x SYBR Safe DNA dye (Invitrogen). Samples were mixed with 6x purple loading dye (without SDS, New England Biolabs) and run at 100 V alongside an appropriate ladder (1 kb or 100b, New England Biolabs), the imaged with an Alliance Q9 Advanced gel dock (UVITEC, Cambridge, UK).

OmpA extracellular loop mutants

Sequence alignment of OmpA homologs identified residue substitutions of charged residues within its extracellular loops. The most common alternative residue was used for generating the OmpA variants, or for residues that are completely conserved, they were substituted with serine. OmpA-NP: R81S, K85T, K94S, R124S, K128G, K134S, R177S; OmpA-NN: D41S, E53N, E89V, D126S, D137S, D180S, D189S; OmpA-NC: combination of both OmpA-NP and OmpA-NN; OmpA-M3: R81S, K94S, R124S. Genes encoding mutants of OmpA were ordered from GeneWizz, ligated into a pET11a vector using flanking BamHI and NdeI restriction sites and validated by sequencing.

MipA

The MipA gene (without signal peptide, residues 1-22) was colony PCRed (introducing flanking BamHI and NdeI restriction sites from K12 WT *E. coli* (forward primer: GCCGCCATATGGAAGGTAAATTTTC, reverse: GCAGCGGATCCTCAGAATTTGTAG, both 5' to 3'). Following digestion of a pET11 backbone and the MipA PCR product with BamHI and NdeI the DNA was purified via agarose gel and extraction. The complementary backbone and MipA fragment were mixed and ligated (T4 DNA ligase) in a 1:1 ratio, incubated for 2 hours and transformed into DH5α *E. coli* and plated. Colonies were grown overnight and mini-prepped for DNA sequencing validation of successful MipA insertion.

2.4.2 Protein expression and purification

All OMPs, OMP mutants and *de novo* designed transmembrane barrels (untagged, other than a his-tag on FusA) were prepared in a similar manner via inclusion body expression and purification (by Dr Bob Schiffrin (OmpLA, OmpA, BamA), Dr Anton Calabrese (OmpF), Dr James Whitehouse (tOmpA, TMB2.17LA, TMB2.3LA), Dr Samuel Haysom (OmpX) and Dr Khedidja Moshabi (FusA), all others by the author). Unless otherwise stated, all proteins were grown in 1x LB (25 g/L).

Inclusion body expression and purification

Competent BL21(DE3) *E. coli* cells were transformed with the relevant plasmid (carbenicillin resistant), grown overnight at 37 °C on agar plates, a single colony picked and grown overnight in ~20 ml of LB containing 100 µg/ml carbenicillin (37 °C, 200 rpm). 5 ml of culture was added to 500 mL of LB, grown to an OD_{600} of ~0.6 and protein expression was then induced with 1 mM IPTG. Three hours post-induction, cells were harvested (5000 *g*, 15 min, 4 °C) and the cell pellet frozen. After thawing the pellet was resuspended in 20 ml buffer (50 mM Tris-HCl pH 8.0, 5 mM EDTA, 1 mM PMSF, 2 mM benzamide), and the cells lysed via sonication. Following centrifugation (25,000 *g*, 30 min, 4 °C) the pellet was resuspended in 20 ml buffer (50 mM Tris-HCl pH 8.0 2% (*v/v*) Triton-X-100) and incubated for 1 h (room temperature, 50 rpm). Following centrifugation (25,000 *g*, 30 min, 4 °C) the supernatant and cell debris were removed from the resulting inclusion body pellet. The inclusion bodies were washed twice by resuspending in 50 mM Tris-HCl pH 8.0 and incubating for 1 h (room temperature, 50 rpm) before pelleting by centrifugation (25,000 *g*, 30 min, 4 °C). Inclusion bodies were solubilised in 25 mM Tris-HCl and 6 M Gdn-HCl (pH 8.0)

for 1 h (60 rpm stirring), and following a final centrifugation (25,000 *g*, 30 min, 4 °C), the supernatant was loaded onto a Superdex 75 HiLoad 26/60 SEC column (GE Healthcare), equilibrated in 25 mM Tris-HCl pH 8.0, 6 M Gdn-HCl. Protein fractions were collected and concentrated to ~100 µM (Vivaspin concentrators) and flash frozen for -80 °C storage. Prior to folding, proteins were buffer exchanged into Tris-buffered saline (TBS, 20 mM Tris-HCl, 100 mM NaCl, pH 8.0), 8 M urea using Zeba™ Spin Desalting Columns, 7k MWCO, 0.5 ml (Thermo Scientific).

FusA

FusA was purified by Khedjida Mosbahi (Walker group, University of Glasgow). FusA was recombinantly produced in *E. coli* and purified from inclusion bodies as previously described[537] with minor modifications. N-terminally His_6-tagged FusA from *P. atrosepticum* SCRI1043 (uniprot: YP_048987) lacking the 20 amino acid signal sequence was overexpressed from *E. coli* BL21(DE3) carrying the pET28a-based expression plasmid pFusA1043Δ20. Cells were grown in LB broth at 37°C to an OD_{600} of 0.6, induced with 1 mM IPTG and grown overnight at 30°C before harvesting by centrifugation. After cell lysis by sonication, the insoluble fraction was collected by centrifugation (22,000g, 30 min, 4°C) and the inclusion bodies resuspended by homogenisation in 50mM Tris, pH 7.5 containing 1.5% LDAO. After incubation at room temperature for 30 min the inclusion bodies were pelleted by centrifugation and then sequentially washed by homogenisation and centrifugation in 50mM Tris, pH 7.5 containing 1.5% LDAO and 50 mM Tris, pH 7.5. The pelleted inclusion bodies were then solubilised by homogenisation in 10 mM Tris, 8 M urea, 1 mM EDTA, 1 mM DTT, pH 7.5 and incubated at 56°C for 30 min with insoluble material removed by centrifugation (8,000g, 10 min 4°C). Refolding of FusA was achieved by drop wise addition, with stirring, to an equal volume of 20 mM Tris, 1 M NaCl, pH 7.9 containing 5% LDAO. Refolded FusA was then dialysed at 4 °C overnight into 20 mM Tris, 0.5 M NaCl, pH 7.9 containing 0.1% LDAO, applied to a 5 ml HisTrap FF column equilibrated in the same buffer and eluted with a 0-500 mM imidazole gradient. The eluted protein was applied to a HiLoad 26/600 Superdex 200 column equilibrated in 50 mM Tris, 200 mM NaCl, pH 7.9 containing 0.1% LDAO and purified protein stored at -80°C.

MSP1D1

MSP1D1 was prepared by Anton Calabrese, as descibed[253]. *E. coli* BL21 (DE3) were transformed with a his-tagged MSP containing vector (pET-28a). Following cell growth to OD_{600} (37 °C, 200 rpm), expression was induced for 3 h with 1 mM IPTG and then the cells centrifugally harvested. Pellets were resuspended in 15 ml/L culture of 20 mM phosphate, pH 7.4 (with 1 mM PMSF, 1% v/v Triton-X 100 and 5 mg of DNase I). Cells were lysed on ice by sonication (6-times cycles of 1 min on and 1 min off), centrifuged at 30,000 *g*, 10 min, 4 °C and the supernatant applied to a pre-equilibrated (20 mM sodium phosphate buffer, pH 7.4) HisTrap FF column (5 ml, GE Healthcare). Then 100 ml of 40 mM Tris-HCl, 0.3 M NaCl and 1% (v/v) Triton X-100, pH 8.0 was passed down the column, followed by 100 ml 40 mM Tris-HCl, 0.3 M NaCl, 50 mM sodium cholate and 20 mM imidazole, pH 8.0, and 100 ml of 40 mM Tris-HCl, 0.3 M NaCl and 50 mM imidazole, pH 8.0. MSP1D1 was eluted with 25 ml of 40 mM Tris-HCl, 0.3 M NaCl and 0.4 M imidazole. To cleave the His-tag: 1 mg of his-tagged TEV protease per 18 mg of MSP1D1 was added in TBS, pH 8.0, 14.3 mM β- mercaptoethanol and incubated for 24 h at 4 °C. The protease and cleaved his-tag were removed by binding to a 5-ml HisTrap FF column, and the MSP1D1 flowthrough was dialysed overnight against TBS, pH 8.0 at 4 °C. Following an additional dialysis against 20 mM Tris-HCl, 0.1 M NaCl and 0.5 mM EDTA, pH 7.4, overnight at 4 °C, the protein was concentrated, filtered, snap frozen and stored at -80 °C.

2.4.3 SDS-PAGE

Samples were mixed 1:3 with loading dye (50 mM Tris-HCl pH 6.8, 6% (*w/v*) SDS, 0.3% (*w/v*) bromophenol blue, 40% (*v/v*) glycerol), boiled if required (>10 min, 100 °C) and ~14 µl loaded onto the gel. The molecular weight markers used were Precision Plus ProteinTM Dual Xtra Standards (BioRad). 15% Tris-tricine gels were homemade and contained 0.1 % (*w/v*) SDS, 1 M Tris-HCl at pH 8.45 and 13.3% (*v/v*) glycerol was included in the resolving layer. The cathode buffer was 100 mM Tris-HCl, 100 mM tricine, 0.1% (*w/v*) SDS, pH 8.25 and the anode buffer, 200 mM Tris-HCl, pH 8.9. A constant current of 30 mA (stacking) and 60 mA (resolving) was used. Following staining (InstantBlue® Coomassie, Abcam) gels were imaged with a Q9 alliance imaging system (Uvitec) and densitometry analysed using ImageJ.

Cold SDS-PAGE makes use of the resistance of natively folded OmpA to

denaturation by SDS in the absence of heat, enabling the fraction of folded/unfolded OmpA (the apparent stability) at different urea concentrations to be determined using gel densitometry. Fraction folded was calculated using only the monomer bands as folded/(folded+unfolded). (For OmpA, inclusion of the higher order bands as unfolded species or by normalising folding against the boiled sample made no significant difference to the fraction folded, compared to [80], possibly due to the use of full length OmpA here). OmpT-DDM was mixed with 1-8 M urea at 1 M increments and 0.5 M urea at a final concentration of 2 µM and incubated overnight. SDS-PAGE loading buffer was added to OmpT-DDM-urea in a 1:3 ratio and 15 µl loaded onto the gel. Fraction folded was quantified by densitometry between the folded and unfolded bands.

2.4.4 Dynamic light scattering (DLS) and ζ-potential

ζ-potentials and DLS (Chapter 3) were measured on a Zetasizer Nano ZS (Malvern) using DTS1070 cells at 25 °C (60 sec incubation), 10-100 measurements were made in a water dispersant. Each sample was measured in triplicate, and cells were cleaned with 2% (*v/v*) Hellmanex, 18MΩ-H_2O and then ethanol and dried under nitrogen. The cell quality was ensured every ~5 measurements with a reference standard (DTS1235, Malvern).

Samples for which only DLS was performed (Chapter 4/5) were measured on a Wyatt miniDawnTreos® system (equipped with an additional DLS detector). Filtered (0.22 µm) buffer was used to obtain ~5 min baselines before and after sample injection. A three-minute sample window was used for the analysis by the software. System was washed with 1 M Nitric acid and MQ-H2O after each run, followed by 1 mL of buffer. Correlation curves were analysed, using the Astra 6.0.3® software, by regularisation. For proteoliposomes, samples were diluted to ~5-10 µg/ml lipid and 300 µl injected.

2.4.5 OMP Intrinsic folding and stability by fluorescence

Intrinsic folding kinetics

The kinetics of intrinsic folding were measured using a QuantaMaster Fluorimeter (Photon Technology International (PTI)) controlled by FelixGX software v4.3, including a peltier-controlled temperature unit. Excitation/emission wavelengths of 280/335 nm were used. OMPs were buffer exchanged from 25 mM Tris-HCl and 6

M Gdn-HCl (pH 8.0) into 10 mM Tris-HCl and 8 M urea (pH 7.4) (Zeba spin desalting columns (Thermo Scientific)). Folding was initiated by rapid dilution of a 3.3 μM unfolded OMP stock in 8 M urea to a final concentration of 0.2 μM OmpA and 0.48 M urea in the presence of 0.32 mM liposomes (LPR of 1600:1 (*mol/mol*)) in 10 mM Tris-HCl and 100 mM NaCl at 30 °C. A minimum of three biological samples were measured for each liposome environment, typically with multiple technical repeats of each preparation, and the kinetics fitted to one-phase exponentials with a custom python script using SciPy to derive the observed rate constants which were then used for further analysis. Kinetics for DMPS containing liposomes were fitted to a two-phase exponential model based on high residual error in one-phase fits. Kinetic traces which folded to an amplitude of < ~75% were not fitted. OmpA folding into PO-lipid liposomes is less efficient than into shorter chain DM-lipid analogues (folding yields ~ 80% and ~ 30% for DMPC and POPC respectively).

OmpA mutants (-NP/NC/NN) were folded into DMPC liposomes at a nominal LPR of 320:1 (folding efficiency assessed by SDS-PAGE bandshift). Proteoliposomes were subject to high speed pelleting (105,000 *g*, 4 °C,30 min, Beckman Coulter, Optima MAX-XP) and then low speed (5,000 *g*, 5 min) to remove unfolded protein and aggregates respectively. The folding kinetics of OmpA-WT was assessed by Trp-fluorescence as above. Folding kinetics were compared by their T_{50} (time to half maximum amplitude).

Systematic OMP folding into DMPG (Chapter 5): protein folding was performed in 20 mM Tris-Cl pH 8.5, 100 mM NaCl buffer. Folding of each protein was assayed by monitoring the changing trp-fluorescence over time when 0.05-0.2 μM (dependent on number of trp residues in the protein) were rapidly mixed with DMPG at an LPR of 1600:1 at 24 °C and a final urea of 0.5 M. Spectra of the unfolded protein (280 nm excitation) in 8 M urea and the post-folding reaction were also collected and blanked against 8 M urea buffer and DMPG + 0.5 M urea buffer. The same assays were conducted with DMPG-LPS liposomes and DMPG in 2 M urea as controls for the presence of LPS and a higher urea concentration respectively.

Protein stability by urea titration

Tryptophan fluorescence emission spectra (300-400 nm) with excitation at 280 nm were measured on samples that had been incubated overnight in different concentrations of urea at 30 °C to ensure equilibrium was reached. The fraction folded protein was then determined by taking the 335/350 nm ratio. OMP

concentration was 0.2 µM with an LPR 1600:1. For OmpT-DDM, 0.2 µM OmpT in 0.05% DDM (20 mM Tris-Cl, pH 8.5, 100 mM NaCl).

2.4.6 General statistics

For all kinetic data significant differences were determined using permutation testing[81], which only assumes data exchangeability under the null-hypothesis (i.e. it makes no assumption about the underlying distribution of the data). The test statistic was defined as the average difference between a pair of datasets. All permutations of the data in these datasets are randomly sampled (without replacement), and the p-value determined as the proportion of samples with a test statistic larger than that of the measured data. For urea-stability data significance was tested using a two-tailed paired t-test. All statistical tests were conducted via pythons scripts, typically using scipy.

2.5 Lipids and liposomes

2.5.1 Liposome preparation

DMPC, DMPG, DMPE, DMPS, POPC, POPG, DLPC and ECPL (*E.* coli polar lipid) lipids (Avanti polar lipids) were prepared as stocks with concentrations of 25 mg/ml in chloroform. Liposome preparations were all made to ~40 mM lipid concentration. Lipids were placed into amber glass vials and dried under N_2, vacuum desiccated for >3 h and resuspended in buffer (10 mM Tris-HCl, 100 mM NaCl, pH 8.5). Following complete resuspension, samples were freeze-thawed five times in liquid N_2 and a 42 °C water bath, then extruded 31-times through 100 nm nucleopore polycarbonate track-etched membranes (Whatman; GE Healthcare) (Avanti extruder) at a temperature ~10 °C higher than the T_m. DMPE lipids were sonicated rather than extruded. 1% (*mol/mol*) DPPE-rhodamine (Avanti) was introduced as a fluorescent label where indicated. Liposomes were used within 48 hours of their synthesis. Lipid concentrations of DMPC and DMPG liposomes were determined using absorbance, calibrated by the Stewart assay[77]: samples were dissolved in 750 µl chloroform, to which 750 µl of guanidine ferric-thiocyanate were added (0.4 M guanidine thiocyanate, 0.1M iron (III) chloride hexahydrate). Samples were vortexed vigorously for 1 min. Following phase separation, the chloroform phase was removed with an 18-gauge needle, its absorbance at 448 nm measured and lipid concentration determined using the calibration prepared. Where required, the desired amount of Ra-

LPS (10 mg/ml aqueous stock) was desiccated with the other lipid, and liposome preparation proceeded identically.

2.5.2 Liposome phase transition temperature (laurdan)

Lipid transition temperatures were measured by laurdan fluorescence using a method adapted from[68]. Laurdan, dissolved in DMSO, was added to preformed liposomes at a ratio 3200:1 *mol/mol* (lipid:laurdan) and a final DMSO concentration of 0.1 % (*v/v*). Liposomes were incubated near their transition temperature overnight. Laurdan fluorescence was excited at 340 nm and its emission at 440 and 490 nm measured for 10 s using a PTI fluorimeter as described above. Spectra were acquired at either 1 °C or 0.25 °C intervals at temperatures spanning approximately +/- 10 °C around the transition temperature, with 3 min equilibration at each temperature. General Polarisation (GP) was determined from the intensity at 440 nm and 490 nm (I, averaged over 10 s acquisition) using the equation

$$GP = (I440 - I490)/(I440 + I490)$$

Midpoints were determined by numerically taking the first differential of the data. At 30 °C (temperature of folding and stability assays) all liposomes used in this study were in the fluid phase. They should thus have similar mechanical properties, as the Youngs and bending moduli are dominated by lipid phase[61–63], although some changes could occur based on the distribution of charged lipids[64–66].

2.5.3 Lipid exchange

The following protocol was adapted from the previous publication ([19]). Concentrations of donor (Cd) and acceptor (Ca) lipid were determined by:

$$Cd = a * Ca * asym / (1 - asym)$$

where *a* is the fraction of lipid accessible (~0.5) and *asym* is the desired asymmetry (up to about 0.5). The concentration of MβCD (Cm) was determined by:

$$Cm = n * Cd + (Cd * K) * 1 / n;$$

where n (set as 4) is the stoichiometry of the CD:lipid complex and K is an experimentally derived value (set as 292 M^{-3} for DMPC, DMPG and DMPS donation, and empirically adjusted to 150 M^{-3} for DMPE donation). These values are sensitive to MβCD activity and phospholipid-specific differences can be significantly reduced by using intermediate MβCD-lipid saturation (fixed at 70 %). Donor liposomes (or resuspended lipid for DMPE) were first solubilised with MβCD (Sigma) at 50 °C, 1000 rpm for >20 min. Acceptor liposomes were then added and incubated at 35 °C, 400

rpm for >20 min to allow for exchange. Liposomes were purified by two rounds of ultracentrifugation (105,000*g*, 4 °C,30 min, Beckman Coulter, Optima MAX-XP). Following resuspension, liposomes were centrifuged at 5,000*g* for 5 min to remove aggregates. To ensure high sample yields only a single round of exchange was carried out, limiting asymmetry to ~30% DMPC/PG and ~ 50% DMPG/PC. Generated asymmetric liposomes were grouped to the nearest 10% (+/- 3%) for analysis. Exchanged liposomes were quality checked and used the same day they were made. Stable symmetric and asymmetric DMPE/PC membranes were successfully created with up to 20% DMPE. Attempts at making DMPS/PG liposomes consistently resulted in aggregate formation. As DMPS and DMPE have T_m values of 35 °C and 50 °C, respectively[78], liposomes always had <20% of these to ensure they remained in the fluid phase.

2.5.4 Determination of sugar concentration using anthrone

Samples (30 µl) were mixed with 100 µl anthrone reagent (0.2% (*w/w*) anthrone in 50% (*v/v*) H_2SO_4), heated at 95 °C for exactly 10 min and then quenched by cooling on ice. Absorbance of the samples was measured at 630 nm as soon after quenching as possible and in all cases <5 minutes. A calibration curve of 0-200 µM substrate (e.g., MβCD, gluocose) at 25 µM intervals was measured every time samples were assayed and used to calculate the sugar concentration.

2.5.5 Liposome-fluorophore absorbance deconvolution

Liposome absorbance was measured between 300-600 nm using quartz cuvettes. Absorbance traces were deconvoluted using a custom script, that found the liposome and fluorophore concentrations that minimised the following function

$$\sum_{\lambda=300}^{\lambda=600} (A_{reconvoluted} - A_{raw})^2$$

where A_{raw} is the raw absorbance trace and $A_{reconvoluted}$ is the theoretical absorbance from the deconvoluted data, using reference spectra of fluorophore alone and unlabelled liposomes. Minimisation was reached via an iterative brute-force method, searching around starting parameters estimated at 560 nm and 440 nm for fluorophore and liposome concentration respectively.

2.5.6 Thin layer chromatography

Liposome samples were diluted to ~0.5-2 mM and 5 µl dried under nitrogen. The sample was resuspended in chloroform and spotted on a TLC plates (Silica gel 60 F_{254}, Sigma 1.6834) and run with 40:9:6:3 (*v/v*) chloroform:methanol:ethanoic acid:water (DMPG-DMPC, POPG-POPC), 60:20:1 (*v/v*) chloroform:methanol:water

(DMPG-DMPE) or 130:20:2 (*v/v*) chloroform:methanol:water (DMPC-DMPS). Plates were dried at 50 °C, dip stained into phosphomolybdic acid and developed by heating at 200 °C for exactly 20 min. Plates were imaged with a Q9 alliance imaging system (Uvitec) and densitometry analysed using ImageJ.

2.5.7 Imaging liposomes using cryo-EM

Samples (3 µl) of ~0.5 mM liposomes were put onto glow discharged (PELCO easiGlow, Ted Pella Inc.) quantifoil grids (1.2/1.3) and incubated for 30 s. Grids were then blotted for 6 s with Whatman #1 filter paper at 4 °C and ~ 90% relative humidity and plunge-frozen in liquid ethane using a Vitrobot Mark IV (Thermofisher). Grids were imaged on a 300 keV Titan Krios (ThermoFisher) electron microscope using EPU software and a K2 detector.

2.5.8 DMPE/PC liposome generation and FRET asymmetry assay

DMPC liposomes doped with different concentrations of NBD-DPPE were generated and the fluorescence spectra measured at excitation/emission wavelengths of 457/530 nm and 375/530 nm in the absence or presence of BSA-ANS (pre-incubated for 1 h at 37 °C, final concentrations 10 µM (BSA) and 30 µM (ANS)). DMPE doped with 1% NBD-DPPE was exchanged into outer leaflets of DMPC liposomes as described above and the ANS-NBD FRET determined by subtracting the fluorescence spectra from the backgrounds (ANS-BSA alone and NBD-DPPE fluorescence excited at 375 nm) and normalised to the concentration of NBD. The FRET of the exchanged samples was significantly larger than the expected symmetric FRET, indicating retention of asymmetry.

2.6 FusA Methods

2.6.1 FusA Grid preparation

Purified FusA (50 mM Tris, 200 mM NaCl and 0.1% (v/v) LDAO, pH 7.9) was diluted to ~ 3 mg/ml. For the DDM-FusA sample the protein was first buffer exchanged to 50 mM Tris, 200 mM NaCl and 0.05% (w/v) DDM, pH 7.9) on a Superdex 200 10/300 column and concentrated to ~ 3 mg/ml. For samples with LPS: Re-LPS (Sigma L9764) or Ra-LPS (Sigma L9641) was diluted in FusA buffer to 20x the final FusA concentration and then mixed 1:1 ratio with 2x concentrated FusA, and then left at room temperature for ~30 minutes to allow for LPS equilibration. CryoEM grids were prepared using an optimised protocol (1.2/1.3 Quantifoil copper 300 mesh unless otherwise stated). Grids were glow-discharged for 20 s at 60 mA in a GlowQube Plus (Electron Microscopy Sciences) under amylamine vapour. 3 µl of sample was applied

to each grid, incubated for 10 s, then blotted for 6 s with Whatman #1 filter paper at 4 °C and >80 % humidity and plunge frozen into liquid ethane (Vitrobot Mark IV, Thermofisher). Collection 1 (K2 detector) was glow-discharged in air vacuum for 30 s at 60 mA, and the grids were blotted immediately after sample application.

2.6.2 FusA CryoEM data collection

Data was collected automatedly on a 300 keV Titan Krios (ThermoFisher) TEM using EPU (Thermofisher) equipped with a range of different detectors (K2 + Gatan energy filter (10 eV slit width), Falcon4, Falcon4i, Falcon4i + Selectris energy filter (5 eV slit width)) as indicated for each sample. A 100 µm objective aperture was used. Full data collection parameters for each sample are shown in **Table A1-A2.**

2.6.3 FusA CryoEM image processing and model building

Image processing was carried out in RELION 3.1[538]/4.0[539] unless otherwise stated. Dose-fractionated micrographs were motion-corrected by RELION and the CTF estimated by CtfFind4[540]. Collections 1-5 were initially picked using crYOLOs[541] general model, and then subjected to either (A) 1-3 rounds of 2D classification and the best protein-like classes selected, or (B) 1-3 rounds of 2D classification followed by initial model generation and one round of 3D classification and particles from the best model selected. The 'good' particles selected from these approaches were used to train both crYOLO and topaz[542] models, the datasets picked with a low-confidence threshold and the particle stacks combined and deduplicated. Particle counts for each of these stages are reported in Supplementary Table 1. As described in the main text, the reconstructions for model 1-6 showed varying extents of rotational averaging, and thus the reported resolutions are rough estimates only. All map sharpening and local resolution estimates used RELION.

For collection 1: 361, 633 particles from trained picking models were passed through two rounds of 2D classification, leaving 68, 263 good protein particles, which were unbinned, in a 280 pixel box and used to generate an initial model. Multiple rounds of iterative 3D classification and 3D refinement resulted in a final particle stack of 10 623, and a model of ~13 Å.

Collection 2: 417, 907 particles from trained picking models were passed through three rounds of 2D classification leaving 103, 092 particles which were

extracted and unbinned into a 288 pixel box. Following initial model generation, a single round of 3D classification yielded 31, 357 particles which refined to a model of ~8.6 Å. Bayesian polishing and CTF refinement made no improvements to the map.

Collection 3: 896, 632 particles from trained picking models were 2D classified once using both VDAM and EM algorithms and the good classes from each recombined and further 2D classified to give 152, 767 particles, which were extracted into a 288 pixel box. 3D refinement gave the best model of ~9.7 Å. Further splitting of the particle stack by 3D classification did not improve the model. 3D refinement of topaz denoised data yielded a map of similar quality.

Collection 4: 377, 986 particles from trained picking models were 2D classified once using VDAM algorithm to remove junk, and all okay particles then 2D classified with multiple parameters (VDAM or EM, T value (2-8), with/without ignoring CTF to first peak) in parallel. The best classes were recombined to 94, 277 particles which were extracted into a 240 pixel box. Following 2 rounds of 3D classification 46, 189 particles were 3D refined (to ~ 9 Å), and an additional round of 3D classification yielded 31, 872 particles. A single round of polishing and masked refinement gave the best model of ~8.6 Å.

Collection 5: 214, 571 particles from trained picking models were 2D classified twice, giving 34, 105 particles. Particles were extracted unbinned in a 288 pixel box and 3D classified once. The resultant 32, 329 particles refined to final model of ~18 Å resolution.

Collection 6: 711, 757 particles were picked by crYOLO's general model and extracted with 2x binning for two rounds of 2D classification. Good classes were grouped into side, top and tilt views and each group used to train a new crYOLO and topaz model, to repick the data at a low threshold (finding 1, 237, 542 unique particles). Following one round of 2D classification to remove junk, all okay particles then subject to two rounds of 2D classification with multiple parameters (T value (2 or 8), with/without ignoring CTF to first peak) in parallel. The best classes were recombined and deduplicated to 275, 399 particles, 3D classified twice, leaving 124, 599 good particles. (An additional lower resolution class was identified that appeared to be missing the plug domain entirely). These were iteratively refined and classified to 85, 859 particles and ~10 Å reconstruction, which were then filtered to 74, 343 particles by excluding all the particles whose angular assignment had rotated by more than 10° in any Euler angle during the final iteration of 3D refinement. An additional

round of 3D refinement on these particles gave a final model of ~8.3 Å resolution. Polishing and CTF refinement yielded no improvement to the model.

Collection 7: 545, 970 particles were picked by crYOLO's general model, extracted with 2x binning and 2D classified twice. Good class averages were grouped as top, tilt or side views, each subject to additional round of 2D classification and then the good classes from each group used to train a new crYOLO and topaz model. These were each subject to two rounds of 2D classification with multiple parameters (T value (2 or 8), with/without ignoring CTF to first peak, EM or VDAM algorithm) in parallel. The protein-like classes were recombined and deduplicated to yield 261, 382 particles. Following two rounds of 2D classification a final high-resolution 37, 826 particle stack was passed into 3D refinement, giving a model of ~8.6 Å. Following one round of polishing and CTF refinement, a final reconstruction at ~7.2 Å was obtained. Map was sharpened with b-factor -187.6 $Å^2$. Neither additional 3D classification or cryoSPARCs Non-uniform[543] refinement improved the model.

Collection 8: 754, 788 particles were picked by crYOLO's general model, extracted with 2x binning and passed through 2D classification, giving 183, 330 protein-like particles. These were used to train a crYOLO model which gave 360, 633 particles, which were subject to iterative, divisive 2D classification (where groups of similar views were classified on individually through multiple rounds). Top, side and tilt views were then re-grouped and used individually to train crYOLO and topaz models, picked with a low threshold and the deduplicated particle stack from each of the view groups subject to iterative, divisive 2D classification. This eventually yielded 38, 494 excellent particles for initial model generation, and were then passed directly into 3D refinement, yielding a model of ~6 Å. This model was used for multiple rounds of 3D classification with all the crYOLO/topaz picked particles, giving a final particle number of 84, 061. Following box expansion to 400 pixels, these were then subjected to 5x rounds of cryoSPARC Non-Uniform[543] refinement and RELION polish and CTF refinement, giving a final model of 2.4 Å.

Additional FusA:Ra-LPS collection: 441, 637 particles were picked by crYOLO's general model, extracted with 4x binning and passed through 2D classification. The resulting best 70, 499 particles were used to train a new crYOLO model, which yielded 401, 942 particles. Following iterative 2D classification, initial model generation and a single round of 3D classification, the 39, 264 good particles were passed to cryoSPARC for Non-Uniform refinement which yielded a model of ~10 Å resolution (post processed in relion).

3D classification/refinement of the optimised particle stacks from collections 1-6 against the high resolution Fusa:Ra-LPS model from collection 8 did not yield improved reconstructions.

2.6.4 FusA Model building

FusA crystal structure (4ZGV)[351] was used as a starting point and following model optimisation in ISOLDE[544] the model was passed through real-space refinement in PHENIX[545] (v1.20) with secondary structure restraints and manually optimised in COOT[546]. Geometry was assessed using molprobity[547]. For LPS fitting (and display), non-protein density was segmented from the map, smoothed using vop-gaussian in chimeraX[548] and then used to fit the LPS into separately before co-refinement with the protein in the unprocessed density.

2.6.5 FusA Molecular dynamics

Coarse-grained models of FusA were generated from the crystal structure (PDB: 4ZGV[351]) using the martinize script and an elastic spring network of 1000 KJ/mol/nm^2 (upper distance cut-off of 0.7 nm). CG-MD was conducted using gromacs[549] (v5.0.7 with martini (v2.2) forcefield[550,551]). Bilayers were built around the transmembrane regions of the protein (which was placed in the centre of the xy plane) using the insane script to randomly place phospholipids[552]. A modified version of insane was written to handle placement of LPS for the 25% LPS simulations by randomly distributing squares of either 2x2-LPS moieties or 3x3-phopholipids over the membrane area. More than two LPS blocks were forbidden from being placed adjacently. The LPS was placed in the sparse-LPS simulations with a custom script that removed 3 phospholipids from the membrane and replaced them with a single LPS moiety. The system was neutralised with 0.1 M NaCl and 0.025 M $CaCl_2$, energy minimised (steepest descent algorithm) and equilibrated with the protein backbone position-restrained for 3 ns. The equilibrated system was used to generate production systems for 3/6 μs (**Table 2.3**), with a 20 fs time-step (simulations with Ra-LPS used a 10 fs time-step), frames were generated at 200 ps intervals. The barostat and thermostat were Parinello-Rahman[553] (1 bar) and V-rescale[554] respectively. A compressibility of 3×10^{-4} bar^{-1} was used. The LINCS algorithm constrained bond lengths[555].

Protein	Membrane composition	Box size /nm	Time /μs	Replicas
-	Outer: 25:75 **Re-LPS**:DLPE Inner: 80:15:5 DLPE:DLPG:CDL2	40x40x11	3	3
-	Outer: 25:75 **Ra-LPS**:DLPE Inner: 80:15:5 DLPE:DLPG:CDL2	40x40x12	3	3
FusA	Outer: 25:75 **Re-LPS**:DLPE Inner: 80:15:5 DLPE:DLPG:CDL2	20x20x20	6	10
FusA	Outer: 25:75 **Ra-LPS**:DLPE Inner: 80:15:5 DLPE:DLPG:CDL2	20x20x20	6	10
FusA	Outer: 1.5:98.5 **Re-LPS**:DLPE Inner: 80:15:5 DLPE:DLPG:CDL2	20x20x20	6	10
FusA	Outer: 1.5:98.5 **Ra-LPS**:DLPE Inner: 80:15:5 DLPE:DLPG:CDL2	20x20x20	6	10

Table 2.3: Summary of all LPS only and FusA:LPS simulations

Lipid-protein contact analysis used a 0.55 nm distance cutoff to define contacts, performed on merged data from all replicas using gmx mindist. All lipid-protein contacts were normalised to lipid number and simulation time. For lipid density analysis the trajectories of all simulation replicas were concatenated and the protein orientation centered and fixed (gmx trjconv), gmx densmap was used to calculate densities. MSD and RDF analysis used gmx msd and gmx rdf repectively. Residence time was calculated using the pyLIPID[556] module with short and long distance cutoffs of 0.475 and 0.8 nm. Convergence of protein-lipid contacts was assessed by comparing convergence between repeats and over the simulation time. Simulations were visualised using VMD[557] and Blender with a set of custom scripts (github.com/JonMarks29/Biomolecular-Blender).

2.7 OmpT Methods

2.7.1 OmpT detergent refolding

Unfolded OmpT in 8 M urea, 20 mM Tris-Cl pH 7.5 was dripped into stirred refolding buffer (4 °C) of 0.5 % DDM, 20 mM Tris-Cl pH 7.5, 50 mM NaCl of 10-times the volume of added OmpT and incubated overnight. Aggregates were pelleted by centrifugation at 25,000 *g*, 30 min, 4 °C and the supernatant diluted 10x in 20 mM Tris-Cl pH 7.5, 50 mM NaCl (i.e. final DDM 0.05% and urea 0.08 M). Sample was then concentrated and purified on a SEC column (Superdex 75 10/300 GE), and concentrated. Protein concentration as measured by A_{280} and then snap-frozen and stored at -80 °C. Refolding into LDAO was identical, with initial and final LDAO concentrations of 0.5% and 0.1% respectively.

2.7.2 OmpT proteoliposome preparation (Chapter 4)

For OmpT-LPS (chapter 4) characterisation all liposomes were made with pre-folded OmpT and lipid by the dialysis method[255]. Lipid (20 mg/ml stock in chloroform) was dried under N_2 gas and vacuum desiccated for >3 h. Refolded OmpT (20 mM Tris-Cl pH 7.5, 100 mM NaCl, 5 mM $MgCL_2$, 0.05% DDM) were mixed with the lipid films to final concentrations of 4 mM lipid and 2.5 µM OmpT and made up to 250 µl in OmpT buffer. Detergent was removed by dialysing against 2 L (20 mM mM Tris-Cl pH 7.5, 100 mM NaCl, 5 mM $MgCL_2$) in 12–14 kDa MWCO D-Tube™ Mini Dialyzers (Merck) at room temperature for 48 h with a total of six buffer changes. Proteoliposomes were then pelleted twice by ultracentrifugation (105,000 *g*, 30 min, 4 °C) and resuspended in 20 mM Tris-Cl pH 7.5, 100 mM NaCl, 5 mM $MgCL_2$. Protein concentration measured via the BCA assay (ThermoScientific) and reconstitution via SDS-PAGE. Final LPRs were approximately 5000:1. Where required, LPS was included via desiccation with the lipid.

2.7.3 OmpT detergent/liposome activity assay

Unless otherwise stated, all OmpT activity assays were performed at 30 °C using 50 µM ARRAY (Abz-Ala-Arg-Arg-Ala-Tyr(NO2)-NH2 , Protein Synthetics) synthetic peptide, which increases in fluorescence when cleaved. Samples were measured on a platereader (BMG Clariostar, 325/430 nm excitation/emission) using 96-well plates (Corning Costar, black plates with transparent bottoms) and 100 µl volume, with fluorescent readings every 90 s and 10 s of gentle shaking (100 rpm) prior to each reading. Unless otherwise stated buffers are: OmpT-DDM buffer: 20 mM mM Tris-Cl pH 7.5, 100 mM NaCl, 5 mM $MgCL_2$, OmpT-liposome buffer: 20 mM Tris-Cl pH 7.5, 100 mM NaCl, 5 mM $MgCL_2$.

OmpT-DDM LPS titrations used 0.025 µM OmpT in 20 mM mM Tris-Cl pH 7.5, 100 mM NaCl, 5 mM $MgCL_2$ and an addition of 0.05, 0.1, 0.5, 1, 5, 10, 50, 100 LPS:OmpT molar ratio of Re-LPS or Ra-LPS. Samples were incubated for 20 minutes at room temperature prior to activity measurements.

OmpT-DDM/-ECPL activity against unfolded BtuB was assayed in 2 M urea with 2 µM BtuB and 0.1 µM OmpT, either apo or with 10 µM Re-/Ra-LPS. Activity against LL37 (synthetically produced, Protein Synthetics) at 3.75 µM LL37 and 0.1 µM OmpT, either apo or with 10 µM Re-/Ra-LPS. For both LL37 and BtuB samples

were incubated overnight (15 h) for degradation and then assessed by the loss of intensity of the LL37 or BtuB band on SDS-PAGE gels.

2.7.4 OmpT activity with LPS variants

0.01 µM OmpT-DDM was mixed with approximately 50x molar excess of each LPS variant and incubated for 20 minutes at room temperature prior to activity measurements with 50 µM ARRAY. Following measurements, the initial activity was determined from the linear region of the activity trace.

OmpT-ECPL+LPS liposomes were generated by dialysis as in Section 2.7.2with 0.1% of the desired LPS. For determining Michaelis-Menten kinetics: 5 µM, 25 µM, 50 µM, 100 µM, 150 µM, 250 µM, 500 µM or 1000 µM ARRAY were added to 0.005 µM OmpT-ECPL-LPS and the activity assayed. Following measurements, the initial activity was determined from the linear region of the activity trace and fitted to a Michaelis-Menten exponential curve. Errors were estimated by pooling all data and determining the 95% CL by bootstrapping.

2.7.5 OmpT product inhibition methods

For adding Ra-LPS to apo-OmpT-DDM mid-plateux: 0.05 µM OmpT-DDM + ARRAY was run to completion overnight (with OmpT-DDM +Re-LPS/Ra-LPS, 50-times molar ratio). After ~78,000 s, fresh OmpT (making a final 0.1 µM OmpT concentration) or Ra-LPS was added to the apo-sample (Ra-LPS at 50-times molar excess compared to OmpT) and the reaction measured for a further 20,000s.

For reactions with cleaved ARRAY: cleaved ARRAY was generated by incubating 750 µM ARRAY with 1 µM OmpT-DDM overnight (fluorescence readout indicated reaction completion) and then purifying the ARRAY fragments by centrifugal filtration (3 KDa MWCO) and taking the flow-through. Assuming complete cleavage, final concentration of the cleaved products was approximately 500 µM. 10 µl of the cleaved product mixture (or buffer blanks) was added at the mid-exponential phase of 0.05 µM OmpT-DDM reactions, and the reactions followed until completion. Alternatively, 10 µl of the cleaved products were included in the starting reaction volume of 0.01 µM OmpT-DDM in the absence and presence of Ra-LPS (50-times molar excess to OmpT). Samples without the cleaved products or without OmpT-DDM

were measured in parallel. Initial reaction rates in each case were then extracted and compared.

2.7.6 OmpT MD and processing

OmpT MD simulations were performed as for FusA, but with OmpT protein model (PDB: 1I78[363]). LPS-protein contacts for the different LPS sugars were determined by counting the protein contacts only for the beads in each sugar group (**Fig. A12**), with the contacts counted and normalised as for FusA.

2.7.7 OmpT nanodisc preparation

Nanodiscs were prepared with ECPL with 10% Ra-LPS was dried under N_2 gas and vacuum dessicated for >3 h, and then solubilised at 25 mM in 100 mM sodium cholate in 20 mM Tris-Cl pH 7.5, 100 mM NaCl, 5 mM $MgCL_2$. OmpT-DDM, MSP1D1 and ECPL-LPS mixture were combined in 1:3:60 ratio, with final sodium cholate concentration of 14 mM. Detergent was removed via BioBead (SM-2 (BioRad)) incubation (4 °C) with 4 replacements over 24 h. Nanodiscs were gel filtrated (10/300 GL, equilibrated in in 20 mM Tris-Cl pH 7.5, 60 mM NaCl, 5 mM $MgCL_2$ and 5% (v/v) glycerol) and protein containing fractions concentrated. The presence of protein and LPS was assessed by silver stain SDS-PAGE, and nanodiscs by DLS.

2.7.8 OmpT cryoEM methods

All grids were prepared in a similar manner. Refolded OmpT, OmpT-nanodiscs or OmpT-OmpA dimers were diluted to ~2.5 mg/ml. For the refolded OmpT, a 20x molar excess of Ra-LPS was added (relative to OmpT) and incubated at room temperature for ~30 minutes to allow LPS equilibration. CryoEM grids (1.2/1.3 Quantifoil copper 300 mesh) were glow-discharged for 20 s at 60 mA in a GlowQube Plus (Electron Microscopy Sciences). 3 µl of sample was applied to each grid, then immediately blotted for 6 s with Whatman #1 filter paper at 4 °C and >80 % humidity and plunge frozen into liquid ethane (Vitrobot Mark IV, Thermofisher).

OmpT-DDM: Data was collected automatedly on a 300 keV Titan Krios (ThermoFisher) TEM using EPU (Thermofisher) equipped with Falcon4i + Selectris energy filter (10 eV slit width) at a pixel size of 0.74 Å. A 100 µm objective aperture

was used. 5399 micrographs were collected, motion corrected and the CTF estimated. 559,003 particles were initially picked by crYOLO's general model, which were subject to multiple rounds of 2D classification, splitting them into different view-groups (tops, sides and tilts). New crYOLO models were trained using particles from each view-group, and the resulting particles stacks subject to extensive additional, iterative and divisive 2D classification (with particular focus on collecting as many different tilt views as possible). The best 64,394 particles were used for initial model generation, and then passed directly into 3D refinement, yielding a model of ~12 Å resolution. This model was 3D classified against all initially picked particles, which resulted in 97,833 good particles which were pre-aligned in a one-class 3D classification and then 3D refined (tau2: 4, low_resol_join_halves 12), resulting in ~10 Å map. Despite extensive additional 3D classification and varying parameters for 3D reconstructions and attempting processing in cryoSPARC, no significant further map improvements were obtained.

OmpT-MSP1D1: Data was collected automatedly on a 300 keV Titan Krios (ThermoFisher) TEM using EPU (Thermofisher) equipped with Falcon4i at a pixel size of 0.68 Å. A 100 µm objective aperture was used. 4740 micrographs were collected, motion corrected and the CTF estimated. 838,151 particles were picked with crYOLO's general model at a low threshold and subject to multiple rounds of 2D classification. 2D classes were grouped into views and used to train new crYOLO models and the results subject to multiple rounds of iterative, divisive 2D classification. The best 35,079 particles were used for initial model generation, and then all particles picked from the trained picking models (543,081, deduplicated) 3D classified against the initial model. The best 74,367 particles were taken forwards for 3D refinement which yielded a model of approximately the correct shape and size but remained very noisy.

OmpT-LDAO: Data was collected automatedly on a 300 keV Titan Krios (ThermoFisher) TEM using EPU (Thermofisher) equipped with Falcon4i + Selectris energy filter (10 eV slit width) at a pixel size of 0.58 Å. A 100 µm objective aperture was used. 20,567 micrographs were collected. A low threshold pick with crYOLOs general model yielded 2,031,841 particles, which initial rounds of 2D classification resulting in 127,861 excellent particles. These were used to train new crYOLO and topaz models in different view-groups (top, side, tilt), which in turn were subject to multiple rounds of divisive and iterative 2D classification to identify the best particles

in each view-group (with particular focus on collecting as many different tilt views as possible). The best 95,351 particles were used for initial model generation (although different initial models were reconstructed from a diverse set of particles stacks), with an ascending T value (3 to 7) over 200 iterations, with 2 classes and a tight spherical mask (15 nm diameter). This model was 3D classified against all the non-junk particles from the initial pick (882,981 particles) multiple times in tandem with 3D refinement of the best resulting class (i.e., better 3D models were used for 3D classification in each case). This yielded a final particle stack of 227,466 good particles and 94,942 excellent particles which were used in the final reconstructions presented in chapter 4. For final 3D reconstruction, the particle stack was first pre-aligned in a one-class 3D classification, these alignments set as angular priors and then passed to 3D refinement where angular shifts were limited to 45° and the number of beyond resolution fourier shells considered expanded to 15 (flag: incr_size). Final models were post-processed using an optimised b-factor (-180). Relion consistently overestimated the resolution, with the final reconstructions reported resolution at 5.0 Å, while it is manually estimated ~6.5 Å.

A broad array of additional processing techniques were attempted, including extensive 2D/3D classification and initial model generation in Relion and cryoSPARC, both with and without alignment and attempting different parameters, particle filtering based on metadata features (e.g., assigned angular accuracy, absolute contrast, probability distribution function, ice thickness estimation), flattening of 2D class distribution, novel 3D classification approaches based on statistically bootstrapping the 3D-reconstruction, various particle picking approaches, micelle subtraction, initial model generation from the solved crystal structure (PDB: I178), optimising CTF estimation by combining multiple methods. However, these approaches yielded none or only nominal model improvements. Processing in cryoSPARC generally yielded worse results than relion.

OmpT-OmpA dimer: A 1:1 ratio of OmpT-OmpA were refolded into DMPG at a combined LPR of 600:1 incubated at 24 °C overnight. Liposomes were solubilised by the addition of LDAO to a final concentration of 0.2%, and gel filtrated (Superdex 75 10/300 GE) and concentrated. The presence of OmpT and OmpA was confirmed by SDS-PAGE. Following grid prep as above, data was collected automatedly on a 300 keV Titan Krios (ThermoFisher) TEM using EPU (Thermofisher) equipped with Falcon4i at a pixel size of 0.68 Å. A 100 µm objective aperture was used. 3099

micrographs were collected, motion corrected and the CTF estimated. 99,441 particles were picked with crYOLO's general model and subject to multiple rounds of 2D classification. Additional crYOLO/topaz model training and repicking the micrographs did not yield better particle stacks. 3D initial model generation failed to generate sensible reconstructions.

2.7.9 OmpT-OMP proteoliposome preparation and activity assays

For Chapter 5 all OmpT proteoliposomes were generated by intrinsic folding OmpT-OMP assays were prepared and performed in a 20 mM Tris-Cl pH 8.5, 100 mM NaCl buffer. DMPG OmpT proteoliposomes were prepared by adding 1 µM OmpT (unfolded in 8 M urea, final urea for folding 0.5 M) to 6 mM DMPG at 24 °C for ~30 min (i.e. 1:6000 LPR). The required ratio (typically 0, 0.2, 0.4, 0.6, 0.8, 1.2, 1.6, 2.4 times OmpT) of another OMP or OMP pair was added to the OmpT-DMPG proteoliposomes with a 2x dilution of OmpT-DMPG, maintaining final urea concentration at 0.5 M, and incubated at 24 °C for ~30 min. For activity assays, the refolded OmpT-OMP mixture was diluted 10x (final OmpT 0.05 µM) and measured as described above for the OmpT alone. Initial rates were extracted by fitting to the linear region of the activity trace. For determining Michaelis-Menten kinetics: 5 µM, 25 µM, 50 µM, 100 µM, 150 µM, 250 µM, 500 µM or 1000 µM ARRAY were added to preformed DMPG 1:1 OmpT-OmpA proeteoliposomes and activity measured. Following measurements, the initial activity was determined from the linear region of the activity trace and fitted to a Michaelis-Menten exponential curve. Errors were estimated by pooling all data and determining the 95% CL by bootstrapping. For OmpT activity in DMPG-LPS liposomes, DMPG liposomes with 0.5% or 0.1% Ra-LPS were prepared as described in Section 2.5.1, and then (following validation of folding by trp-fluorescence) used to prepare proteoliposomes and measure OmpT activity identically to DMPG only liposomes.

For activity in different lipids: OmpT or 1:1 OmpT-OmpA or 1:1/1:2 OmpT:BamA/MipA mixture were first refolded into DMPC or DLPC liposomes (6000:1 LPR) by incubation overnight at 24 °C or DMPS by incubation at 35 °C for 30 minutes, and then OmpT activity measured as described above.

2.8 OmpLA Methods

2.8.1 OmpLA activity assay (DLS and light scattering)

OmpLA and OmpLA+OMP (always 1:1 ratio) DMPG proteoliposomes were prepared identically to OmpT and OmpT+OMP proteoliposomes. DLS of OmpLA refolded into DMPG at LPR 6000:1, DMPG-OmpLA after incubation at 30 °C for 24 h, DMPG-OmpLA in the presence of 10 mM EDTA and DMPG only blank was performed on the Wyatt instrument as in Section 2.4.4. For the time course: DMPG-OmpLA liposomes (6000:1 LPR) were prepared and the reaction initialised by the addition of 10x molar excess $CaCl_2$ (compared to OmpLA), and quenched at each time point with 10x molar excess (compared to $CaCl_2$) of EDTA. Samples were then diluted and measured as in Section 2.4.4.

Light scattering of OmpLA and OmpLA+OMP was performed with a final liposome concentration of 1 mM and an LPR of 6000:1. Reaction was initialised by addition of 10x molar excess of $CaCl_2$ (compared to OmpLA). Reactions were measured in paired (i.e. baseline and reaction) QS quartz cuvettes on a room temperature spectrophotometer monitoring the A_{560} every 1 s, which acts as a proxy for increased light scattering as liposomes collapse and aggregate.

2.8.2 OmpLA activity assay (fluorescence)

The fluorescence of ANS (200 µM), ANS+BSA (20 µM, 1 µM) and ANS+BSA+oleic acid (20 µM, 1 µM and 0.1% v/v respectively) was compared on a Horiba PTI with 377/470 nm excitation/emission wavelengths at 30 °C and with 5 nm slit widths. Relative concentrations of BSA and ANS to generate a maximal response upon the addition of oleic acid were optimised. OmpLA kinetics were measured with 9 µM ANS, 0.5 µM BSA, 0.02 µM OmpLA in DMPG at a 6000:1 LPR. Reactions were initialised by the addition of 1 mM $CaCl_2$ and monitored with the same fluorescent settings as above. Pre-/Post-reaction spectra were also taken.

The assay throughput was improved by optimising for a platereader (Fluostar Omega, BMG Labtech, 430 nm filter) performed in 96 well Corning Costar, black plates with transparent bottoms) with final concentrations of 9 µM ANS, 0.5 µM BSA, 150 µM DMPG, 0.025 µM OmpLA (LPR 12000:1); 1 mM $CaCl_2$ added to start the reaction; ~10 s reaction deadtime. Samples were measured at 30 °C, with a 1:1 ratio of OmpLA:OMP, with measurements every 4 s. Following measurements, the initial activity was determined from the linear region of the activity trace and the T_{50} as the half-maximum intensity drop.

2.9 PagP Methods

2.9.1 PagP refolding

PagP refolding into DMPG (3000:1 LPR) was optimised to overnight in 6 M urea at 24 °C in 20 mM Tris-Cl pH 8.5, 100 mM NaCl buffer. Circular dichroism spectra were measured in 1 mm QS quartz cuvette using a Chirascan plus CD spectrometer (Applied Photophysics) with bandwidth of 2.5 nm and adaptive sampling. Three scans were averaged between 260 nm and the lowest usable wavelength (dependent on urea concentration). PagP was measured at 6 µM in DMPG at 400:1 LPR or unfolded in 8 M urea. Trp-fluorescence spectra were collected as previously described. PagP refolding into DDM was performed in the same way, but in the presence of 2 M urea.

2.9.2 PagP activity assay

Assay validation: MβCD-pNP (palmitate 4-nitrophenyl) complexes were generated by solubilising pNP in water at 70 °C, and then adding and rapidly mixing room temperature MβCD (200 mM stock) in a 2:1 ratio. DLS and phase transition temperature measurements (by laurdan) of liposomes alone or in the presence of MβCD or MβCD-pNP was performed as in Section 2.4.4. and 2.5.2 respectively. Time resolved phase transition changes were measured on a PTI fluorometer twice, at 440 nm or 490 nm, on different samples. A 30 s baseline was measured and then following addition and mixing of MβCD or MβCD-pNP (5 s lag time) an additional 30 or 60s was added.

For PagP reactions: a stock of 1 mM pNP in a 1:2 ratio with MβCD was prepared. PagP was refolded to a final concentration of 1 µM (3000:1 LPR) overnight as described above, then diluted three times, incubated at room temperature for 10 minutes and centrifuged at 5000 *g* for 5 min to remove any aggregates. Reactions were measured in paired (i.e. baseline and reaction) QS quartz cuvettes on a room temperature spectrophotometer, initialised by adding pNP to a final concentration of 100 µM and rapidly mixing, and recording the A_{410} every 1 s for ~1000 s. For DMPG-PagP 1.5x dilute, PagP was refolded to a final concentration of 0.66 µM. For DDM-PagP reactions all concentrations were 10x less (i.e. final PagP 0.033 µM, initial pNP 10 µM) and conducted in 20 mM Tris-Cl pH 8.5, 100 mM NaCl, 0.05% DDM. Following measurements, the initial activity was determined from the linear region of the activity trace. Folding of tOmpA, OmpX and OmpA (unfolded in 8 M urea) into DMPG at 2 M urea was validated by trp-fluorescence (Section 2.4.5), and each OMP was folded into DMPG-PagP proteoliposomes following urea dilution to 2 M at a 1:1 ratio with PagP (assuming no PagP loss) by incubating at 24 °C for 30 min. Mid-reaction

addition of OmpA was done at ~22 °C in the spectrophotometer, rapidly mixed immediately after OmpA addition.

2.10 LPS purification and characterisation

Cells expressing the desired LPS chemotype (ΔWaaC, ΔWaaF, ΔWaaI, ΔWaaR from the Keio collection) were grown in LB at 37°C (0.5 L) overnight. Following harvesting (5,000 *g*, 10 minutes), the cell pellet was resuspended in 20 ml 0.85% NaCl solution and incubated for 1 h. Three rounds of butanol extraction were performed: 20 ml of butanol were added, and the mixture incubated for 1 h (with shaking). Samples were then phase separated by centrifugation (15,000 *g*, 25 min, 4 °C) and the aqueous phase extracted and taken forward. 1 M Tris-Cl was added to make the sample up to 20 mM Tris-Cl (pH 7.5) and then incubated with 2 µg/ml of proteinase K overnight. The next day, LPS micelles were purified from protein fragments and protease by centrifugal filtration (100 KDa MWCO, LPS micelles are generally too large to pass through). DNase I and RNase were added to the sample each at 2 µg/ml and incubated overnight, and the LPS-micelles then purified and concentrated via centrifugal filtration as before.

LPS purity was assessed by silver stain SDS-PAGE, UV-Vis (190-320 nm) and sugar detection (anthrone assay, see Section 2.5.4). Zinc imidazole staining of LPS was optimised to the following conditions: (1) Rinse gel in water (3 times), incubate with 0.2 M zinc-sulphate for 15 minutes, rinse gel briefly in water, add 0.2 M imidazole for about 30 s and once contrast has developed pour off stain and rinse with water (3 times). (2) Rinse gel in hot water (3 times), incubate in 10 mM zinc-sulphate for 15 minutes and add 0.2 M imidazole for 3 minutes and then rinse in water. Gels were imaged via a Q9 Alliance Advanced gel dock (Uvitec).

Commercial O111:B5 LPS was contaminated with DNA and prior to use treated with 10 µg/ml of DNase I (10 mg/ml LPS stock) for 90 minutes, and the LPS then purified via centrifugal filtration (100 KDa MWCO). Removal of DNA was confirmed via UV-Vis spectra.

2.11 OMP Crosslinking

The desired OMP(s) were folded (from unfolded 8 M urea stock) into DMPG liposomes (with the addition of 1% Ra-LPS when required) at an LPR of 1000:1, incubated at 24 °C for ~30 min and a final urea of 0.5 M (20 mM Tris-Cl pH 8.5, 100

mM NaCl, 5 mM $MgCl_2$ buffer). Proteoliposomes were then subject to 4 rounds of ultracentrifugation (130,000 *g*, 20 minutes, 10 °C) and resuspension in the crosslinking buffer (20 mM bicine-Cl pH 8.5, 100 mM NaCl, 5 mM $MgCl_2$) followed by a final 5000 *g*, 5 min centrifugation to remove any aggregates. The method was validated with tOmpA and OmpX samples which were sized by DLS before and after the ultracentrifugation and their folded-unfolded SDS-PAGE bandshift measured after ultracentrifugation. Crosslinking was performed with a final concentration of 0.1 % (v/v) glutaraldehyde at room temperature for 5 or 15 minutes. Samples were then boiled and crosslinking assessed by SDS-PAGE and densitometry to quantify monomer loss. Due to the similar sizes of OmpA and OmpT, tOmpA was used rather than OmpA in all multi-OMP crosslinking assays. All data for single OMP species in the presence and absence of OMPs at each time-point was combined by taking the difference between each data-point and the average crosslinking (of both DMPG and DMPG:Ra-LPS for that OMP).

2.12 MipA methods

MipA bandshift was tested by analysing 1.5 µM DMPG refolded MipA (with or without boiling) by SDS-PAGE at room temperature. Circular dichroism was conducted in 1 mm QS quartz cuvette using a Chirascan plus CD spectrometer (Applied Photophysics) with bandwidth of 2.5 nm and adaptive sampling. Three scans were averaged between 260 nm and the lowest usable wavelength (dependent on urea concentration). MipA was measured at 6 µM in 0.05% DDM, DMPG at 400:1 LPR or unfolded in 8 M urea.

2.13 Computational methods

2.13.1 Alphafold2

All Alphafold2 predictions were performed using a local install of Alphafold2-multimer (v3) using the reduced sequence databases, without final energy minimisation and generating five predictions. Automated pipelines combined fasta files (multimeric OMPs included enough protein copies to make a single barrel), ran the predictions and carried out initial processing on all predictions for the systematic dimers (including model filtering by pLDDT (predicted local distance difference test) and generation of PAE (predicted alignment error) plots and calculation of intra-/inter-/interacting-PAE scores). Interacting-PAE was defined as average adjacent-residue (Cα-Cα cutoff of 1.2 nm) between residues of different chains. Models were manually

inspected via an automatically generated report containing images of the models and their statistics. For OmpT-OmpA from different species, protein homologs were collected from the relevant species reference genomes at the European nucleotide archive (ENA) and the complex predicted.

Skew and kurtosis values for the residue adjacent inter-chain PAE histograms and the distribution difference independent samples Kolmogorov-Smirnov test was performed using scipy. Angular distributions were determined by aligning predictions against a reference of the primary OMP (which was itself aligned to have the z-axis down its barrel centre) and then determining the centre of mass of the transmembrane regions of each secondary OMP. For the strand vs residue adjacent inter-chain PAE analysis: the number of strands per OMP is known in each case. The control dataset was pair complex predictions of BamA's POTRA 1 with all periplasmic proteins.

2.13.2 OMP extracellular loop bioinformatics

For experimental structure analysis: 394 OM-annotated proteins from the Orientations of Proteins in the Membrane (OPM)[70], of which 198 have transmembrane regions, were sequence clustered to 70% sequence identity using CD-Hit[92] and manually inspected, resulting in 75 structures. Proteins from OPM are aligned in the membrane already, and 3D-space was split into 1 Å slabs parallel to the membrane plane, with residues assigned based on their Cα position. The enrichment/depletion of residues was calculated relative to either the total amino acid content in the protein or in the soluble regions. 2σ/3σ significance was calculated separately for enrichment and depletion by finding the standard deviation of all positive and negative enrichments.

For predicted structure analysis: 2285 OM-annotated proteins were identified in the EBI-Alphafold2 database[72], (December 2021). Signal peptides were predicted (SignalP v5.0[93]) and removed from the structures (proteins with <90% prediction confidence rejected). Proteins were filtered by Alphafold2's pLDDT (>80%) leaving 1765 proteins. Transmembrane regions and membrane orientation was predicted using Immers[70] and 842 proteins identified with >0 transmembrane regions (693 >8 strands, i.e. are full barrels). Sequences were clustered to 70% sequence identity using CD-Hit[92], leaving 343 structures, which were processed as for the OPM dataset.

For sequence data analysis: The ~1.3 x10^6 sequences in the OMPdb (August 2021)[73]

were quality filtered by topology prediction and pHMM coverage score (both >95%) and sequences missing residues removed, leaving 71, 181 sequences. These were sequence clustered to 70% sequence identity using CD-Hit[92], leaving 17, 931 sequences. Residue enrichment was carried out as above using residue count away from the centre of the membrane to split the protein into slabs. A distance calibration for residue count was determined from the OPM structures database combined with sequence topology prediction.

2.13.3 ζ-potential prediction model

A review of the literature ([19, 44-56]) combined with data presented here, yielded 315 data-points that met the following inclusion criteria: (i) T_m of all lipids of each sample must be known (except cholesterol which is handled separately; any liposomes without defined acyl-chain composition are removed), (ii) measurement buffer salt must be NaCl or KCl, and (iii) ethanol must not be present in the buffer. Lipid composition of all the liposomes was parametrised with three values: (A) average overall charge per lipid; (B) average T_m of all lipids; and (C) fraction of lipid composition that is cholesterol.

An Extreme Gradient Boosted model (from XGboost library[79]) was used with a root mean-squared error loss function, a learning rate of 0.05 and an early-stop patience of 25 cycles (as evaluated from the current 25% validation data). The model was trained with the target ζ-potential using eight dataset features: salt concentration (monovalent), salt concentration (divalent), pH, hydrodynamic radii, temperature, overall charge, lipid T_m, cholesterol fraction. The error associated with each measurement (the standard deviation value) was used to weight the features of an individual data-point, with weightings normalised between 0.375 and 0.625. Model hyper-parameters were explicitly optimised to reduce model over-fitting identified in early testing: subsample-per-node: 0.85, subsample-per-tree: 0.85, minimum-child-weight: 2.5 and maximum tree depth: 6. The models were validated with 4-fold cross-validation. Predictions were made by training a 50 model ensemble (all with mean average error (MAE) < 5 mV) on-the-fly and averaging their predictions to obtain a final value. The weight/gain-per-feature was analysed using python package scikit-learn. The code can be accessed at: https://github.com/JonMarks29/zeta-potential-prediction.

2.13.4 OmpA Coarse-grained molecular dynamics (CG-MD)

A structural model of full-length OmpA was predicted using Alphafold2, and the structural accuracy of the transmembrane and soluble domains confirmed by comparison with experimental structures (PDBs: 1G90, 2MQE), and for BamA the crystal structure (PDB 5D0O) was used. Following any *in silico* mutations (using Modeller[82]), structures were coarse-grained using the martinize script with an elastic spring network of 1000 kJ/mol/nm^2 (upper distance cutoff of 0.7 nm). CGMD was conducted using gromacs (v5.0.7)[83] with Martini (v2) force field[84,85]. Bilayers were built around the transmembrane regions of the protein by randomly placing lipids using the insane script with the protein at the centre of the X-Y plane[86]. CG waters were added and then the system neutralised with and 0.1 M of NaCl added. The system was energy minimised (steepest descent algorithm) and equilibrated with the protein backbone particles position-restrained for 3 ns. The equilibrated system was used to generate production systems for 3 μs, with a 20 fs time-step and frames generated at 200 ps intervals, with simulations run at 310 K. The barostat and thermostat were Parinello-Rahman (1 bar) [87] and V-rescale respectively[88]. A compressibility of 3×10^{-4} bar^{-1} was used. The LINCS algorithm constrained bond lengths[89]. Lipid-protein contact analysis used a 0.55 nm distance cutoff to define contacts, performed on merged data from all replicas using gmx mindist. All lipid-protein contacts were normalised to lipid concentrations and simulation time. For lipid density analysis the trajectories of all simulation replicas were concatenated, and the protein orientation centered and fixed (gmx trjconv), gmx densmap was used to calculate densities. Residence time was calculated using pyLIPID[69], with short and long distance cutoffs of 0.475 and 0.8 nm respectively. The simulations were validated by determining the average area-per-lipid (using fatslim[90]) and surface tension (as in ref[91]) over the simulation time-course, the z-axis average density of the membrane components (gmx density), the protein RMSF (gmx rmsf) and the convergence of the lipid-protein contacts between repeats.

Chapter 3: Membrane charge asymmetry and OMP folding

Biological membranes consist of two leaflets of phospholipid molecules that form a bilayer, and typically the composition of lipids in each leaflet is distinct. This asymmetry is created and maintained *in vivo* by dedicated biochemical pathways, but difficulties in creating stable asymmetric membranes *in vitro* have restricted our understanding of how bilayer asymmetry modulates the folding, stability and function of membrane proteins. Here we employ cyclodextrin-mediated lipid exchange to generate asymmetric liposomes and use these to characterize the stability and folding kinetics of two bacterial outer membrane proteins (OMPs). We show that excess negative charge in the outer leaflet of a liposome impedes the insertion and folding of OmpA and BamA into a membrane, while excess negative charge in the inner leaflet accelerates their folding, relative to symmetric liposomes with the same membrane composition. Using molecular dynamics simulations and mutational analyses, we identify three positively charged residues in the extracellular loops of OmpA that play a critical role in folding. Bioinformatics was then used to identify a conserved patch of positive residues in the extracellular loops of OMPs, that lies 6-8 Å from the membrane surface. Together, these results rationalise the well-known 'positive outside' rule for OMP sequences and suggest new insights into the mechanisms that drive OMP folding and assembly both *in vitro* and *in vivo*.

Work described in this chapter is published in Machin, J.M., Kalli, A.C., Ranson, N.A., Radford, S.E., Protein–lipid charge interactions control the folding of outer membrane proteins into asymmetric membranes. *Nature Chemistry* (2023). https://doi.org/10.1038/s41557-023-01319-6). The text and figures presented here have been adapted from this publication. Dr Bob Schiffrin purified the wild-type OmpA and BamA proteins used in this study.

3.1 Background

Membrane asymmetry in biology is ubiquitous, including across the membranes of diderm bacteria. The degree of asymmetry found depends on membrane-type and cell-status, and often induces a charge dipole[558,559] (the outer membrane (OM) of diderms is extracellularly negative). The plethora of enzymes involved in creating and maintaining this asymmetry, often in energetically demanding processes, highlights its importance[560,561], and this is further emphasised by characterised disease states featuring mis-regulated asymmetry[67]. However, its implications on protein folding, stability, structure, and function are poorly understood, in large part due to challenges in generating sufficient quantities of stable asymmetric membranes for these studies.

One approach to generating asymmetric liposomes is using cyclodextrin (CD)-mediated lipid exchange[79,562,563], which allows lipids solubilised in bulk solution to exchange over the available outer surface of a liposome, and thus by exchanging differently charged lipids it is possible to generate charge asymmetric liposomes[80]. These have been used to study membrane protein folding, with the rate of folding of perfringolysin O[85] and the 'pH low insertion peptide' (pHLIP)[86] each being modulated by lipid asymmetries across the bilayer (for pHLIP having less negative charge on the accessible bilayer leaflet increased insertion propensity). However, both of these proteins exist in stable, water-soluble forms that only insert into membranes under specific conditions[564,565], which are then shown to be manipulatable by membrane charge asymmetry. It is thus difficult to generalise these finding to integral membrane proteins, which require a membrane to adopt their native fold.

Outer membrane proteins (OMPs) from diderm bacteria[31,221] are ideal substrates to study the implications of membrane asymmetry on membrane protein partition and folding into the bilayer. In particular, the ability of OMPs to intrinsically fold (that is, to partition into the membrane and adopt their native fold from a denaturant-stabilised unfolded state when mixed with synthetic membranes) allows simple folding systems to be set up to systematically study the role of membrane asymmetry[262]. This is further helped by a wealth of data in the literature describing how different features of symmetric membranes can modulate OMP folding (reviewed in [31]). For example, folding is faster when bilayers contain lipids with short acyl-chains[225], less saturated lipids[265], or more membrane defects[266]. Lipid head groups

also modulate folding, with phosphoethanolamine (PE) and phosphoglycerol (PG) introducing a kinetic barrier for folding into $C_{10:0}$ lipid bilayers[267,268]. However, recent work with $C_{14:0}$ lipids did not show this effect, perhaps because the additional kinetic barrier of a thicker membrane dominates folding[269]. The primary sequence of an OMP is also critical, perhaps even more than the properties of the membrane[269] - a concept supported by mutational analysis of folding efficiency for OmpA, EspP and OmpC variants *in vivo*[566]. While OMP folding into membranes of different lipid composition has been studied for decades[141], the role of membrane asymmetry has not yet been studied in detail.

In this chapter the first systematic study of the effects of membrane asymmetry on OMPs is presented. Using CD-mediated lipid exchange, charge-asymmetric liposomes are generated with DMPC, DMPG, DMPE and DMPS lipids, and their asymmetry validated using measurements and predictions of their ζ-potential. We show for two model OMPs, 8-stranded OmpA and 16-stranded BamA, that folding rate and stability are modulated by a leaflet-specific distribution of negatively charged lipid head groups, irrespective of acyl chain length. Using molecular dynamics (MD) simulations, we identify specific, positively charged residues in the extracellular loops of OmpA that interact with lipids, and show that they are critical for OmpA folding *in vitro*. Bioinformatics analysis of >300 structures and >19,000 sequences of OMPs revealed a highly-conserved enrichment of positively-charged residues in the extracellular loops close to the membrane surface. Collectively using this integrative approach of experiment, bioinformatics and simulation (**Fig 3.1**), the results reveal that efficient OMP folding requires a previously uncharted synergy between the lipid charge in each leaflet of the bilayer and a signature region (the ‘patch of external positive residues’, or *PEP*) of Lys/Arg in the extracellular loops of the folding OMP. This finding is particularly important given the high charge asymmetry in the lipopolysaccharide-containing OM. The results provide new insights into how lipid organisation modulates OMP folding and stability *in vitro,* have implications for understanding OMP folding *in vivo*, and suggest new strategies to control OMP folding and stability for biotechnological applications.

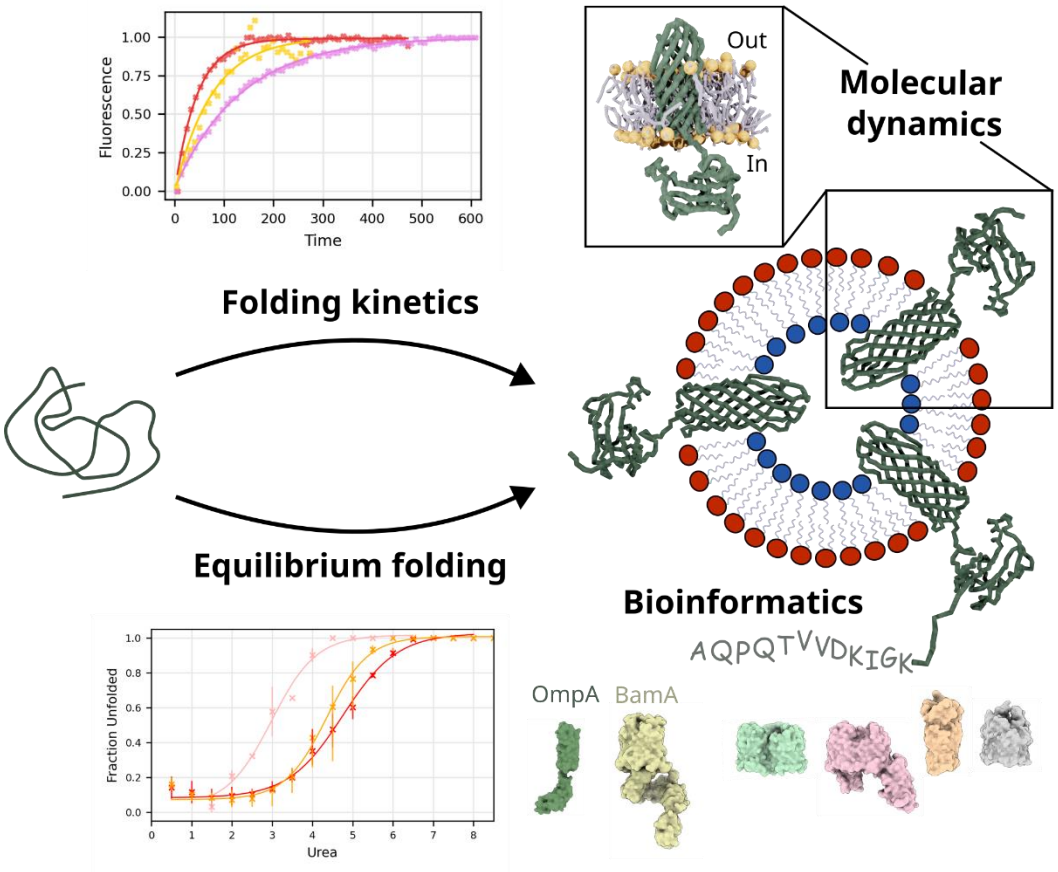


Figure 3.1: Combining experiments, simulations and bioinformatics to reveal how charge patterning in OMP loops and membrane asymmetry synergise for productive folding and stability. Following generation of charge asymmetric liposomes (depicted here by red and blue headgroups), OMP folding kinetics (top left, measured by Trp fluorescence) and stability (bottom left, measured by cold SDS PAGE) of two model OMPs, OmpA and BamA (green and yellow space fill structures) were measured and compared with the results for the same lipids in symmetric membranes (not shown). Molecular dynamics of OMPs pre-folded into different lipid systems (top right), as well as structural and sequence bioinformatics (bottom right) for 300 and 19000 OMPs, respectively (six are depicted) were then used to identify residues involved in the modulation of folding rates and stabilities upon interaction with the lipid head group.

3.2 Results

3.2.1 Cyclodextrin-mediated exchange can generate stable, asymmetric liposomes

To determine the effects of lipid charge asymmetry on OMP folding and membrane stability, dimyristoyl (DM)-phosphatidylcholine and -phosphoglycerol (DMPC and DMPG, **Fig. 3.2a**) lipids were predominantly used. These lipids have similar head group sizes and lipid phase transition temperatures (T_m, 24.0 °C and 23.5 °C respectively[567]), and their identical $C_{14:0}$ acyl-chains, generate a bilayer with a similar hydrophobic thickness to the native OM[31]. Further DM-lipids, especially DMPC, has been extensively used to characterise OMP intrinsic folding into symmetric membranes (e.g. [225,269]) providing the ideal framework within which to begin to determine the role of bilayer asymmetry in OMP folding and stability. DMPC is a net neutrally charged zwitterion, while DMPG has a negative charge, allowing asymmetric lipid systems to be generated with a charge gradient (the expected membranes and their dipoles are shown in **Fig. 3.2a**, right).

DMPC-DMPG asymmetric liposomes were generated by methyl β-cyclodextrin (MβCD)-mediated exchange, which is an efficient method to generate relatively large quantities of controllably asymmetric liposomes (**Fig. 3.2b**). (Symmetric and asymmetric lipid membranes are henceforth indicated by s- and a-prefixes, respectively, and asymmetric liposomes are denoted as donor-lipid/acceptor-liposome. Thus a-DMPG/PC indicates DMPG lipids exchanged into the outer leaflet of DMPC liposomes (all lipid ratios are *mol/mol* unless otherwise indicated)). Following lipid exchange, the integrity and size of the final liposome preparation was confirmed using dynamic light scattering (DLS) and cryoEM (imaged liposomes are smaller than expected due to preferential incorporation into the ice of smaller objects during sample preparation) (**Fig. 3.2c-d**). MβCD was removed from the preparation via two rounds of liposome pelleting and resuspension via ultracentrifugation, determined by detection of residual sugar using a modified anthrone assay (typically >1000x reduction, with <75 μM remaining, **Fig. A1**). Together, this confirmed that gross liposome structure remained intact throughout the preparative process.

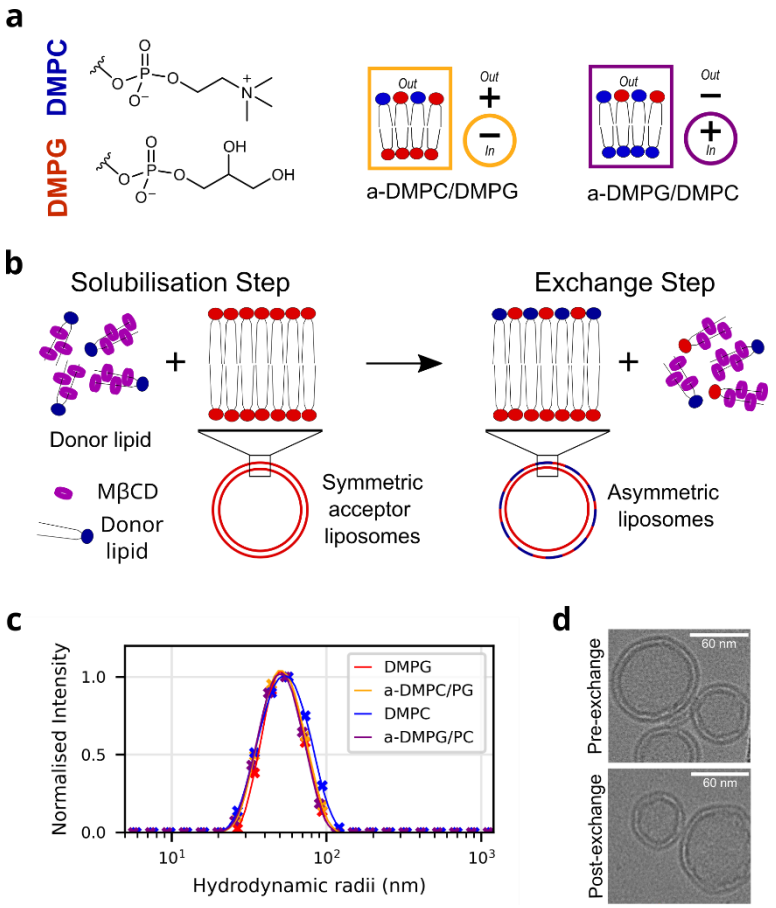


Figure 3.2: Generating and validating gross structure of DMPC-DMPG asymmetric LUVs. (a) Headgroup structure of DMPC and DMPG lipids (blue and red, respectively). The same colours are used throughout. **(b)** Overview of asymmetric liposome generation by MβCD-mediated exchange. **(c)** Pre-/post-exchange liposome size by DLS. **(d)** Pre-/post-exchange liposomes imaged by cryo-EM. The liposomes are smaller than observed using DLS as small liposomes preferentially go into the ice.

Next, the competency of liposomes composed of DMPC and DMPG to undergo MβCD-mediated exchange was demonstrated with a fluorescence-based assay. Lipid exchange was attempted using DMPG liposomes doped with 0.5% DPPE-Rhodamine (DMPG-Rho) and DMPC liposomes, in both directions (i.e. generating a-DMPC/DMPG-Rho and a-DMPG-Rho/DMPC). Measuring the 300-600 nm absorbance spectra (**Fig. 3.3a)** and subsequently applying a custom iterative-based deconvolution method allowed the separation of the liposome and fluorophore absorbance peaks (**Fig 3.3b-c**, section 2.5.5). Normalising the fluorophore fluorescence component to the liposome concentration yields directly comparable fluorophore absorbance traces. This analysis clearly show that DMPG-Rho liposomes

accept non-fluorescently labelled lipids from DMPC-MβCD complexes (a reduction in fluorescence) and DMPG-Rho, when solubilised with MβCD, could donate fluorescently labelled lipids into DMPC liposomes (a gain in fluorescence) **(Fig. 3.3d)**.

Following this, to validate exchange in a label-free manner, thin layer chromatography (TLC) quantified by densitometry, was then used. Controls confirmed that DMPG and DMPC staining intensity was directly proportional to the amount of lipid loaded and that samples run on the same plate stained to the same intensity with the same amount of lipid (<3% different, no comparisons are made between different plates) **(Fig. A2)**, allowing the ratio of DMPC:DMPG in a particular liposome solution to be determined and directly visualising lipid exchange **(Fig. 3.4a)**. Label-free DMPC-PG liposome asymmetry was then confirmed by determining ζ-potential, a measure of the charge associated with the slipping plane (i.e. the particle and associated solvent that diffuse together in solution) of a particle in solution, allowing quantitation of the amount of neutral DMPC and negatively-charged DMPG in the solvent-exposed outer leaflet of liposomes (**Fig. 3.4b**). Coupled with the total lipid ratio, ζ-potential thus allows the inner leaflet composition of a liposome to be estimated and provides a direct readout of lipid asymmetry (**Fig. 3.4c**). For example, the green circled a-DMPG/PC exchange sample has a total DMPG fraction of ~25 % (lower x-axis) and a ζ-potential -26 mV. These data fall onto the dashed, theoretical “asymmetry” line, confirming this liposome is asymmetric, and allowing the outer leaflet DMPG content to be read from the upper x-axis (50%).

Asymmetric liposomes of up to ~30% a-DMPC/PG and ~50% a-DMPG/PC in their outer leaflets could be generated from a single round of exchange. (Greater amounts of asymmetry were possible by subjecting liposomes to a second round of exchange and purification (~55% a-DMPC/PG and ~75% a-DMPG/PC), but led to a significantly lower yield). Asymmetric liposome stability and asymmetry were validated over 72 hours in the presence and absence of 8 M urea using DLS and ζ-potential (**Fig. 3.4d-e**) – thus confirming that the generated charge asymmetry was adequately maintained for the intended folding and stability assays.

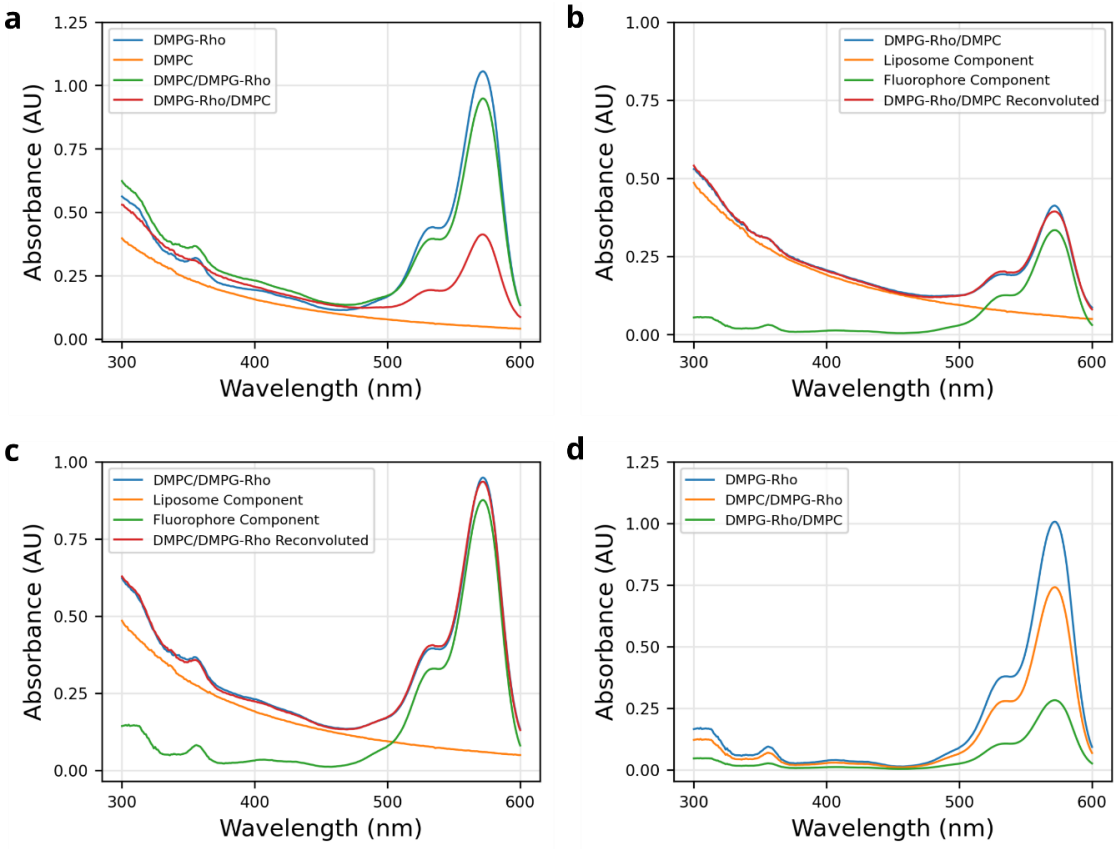


Figure 3.3: DMPC and DMPG lipids are competent to exchange through MβCD-mediated exchange. **(a)** Raw absorbance spectra of unexchanged DMPG-Rho liposomes (DMPG + 1% (mol/mol) DPPE-rhodamine) and DMPC liposomes, and exchanged liposomes, as indicated. **(b)** Deconvoluted absorbance spectra for a-DMPG-Rho/DMPC exchange and **(c)** a-DMPC/DMPG-Rho exchange, separating the liposome and fluorophore absorbance components. **(d)** Liposome concentration normalised fluorophore components of DMPG-Rho and the exchanged samples, showing loss of fluorescence from the a-DMPC/DMPG-Rho and gain of fluorescence in the a-DMPG-Rho/DMPC samples, indicating successful lipid exchange.

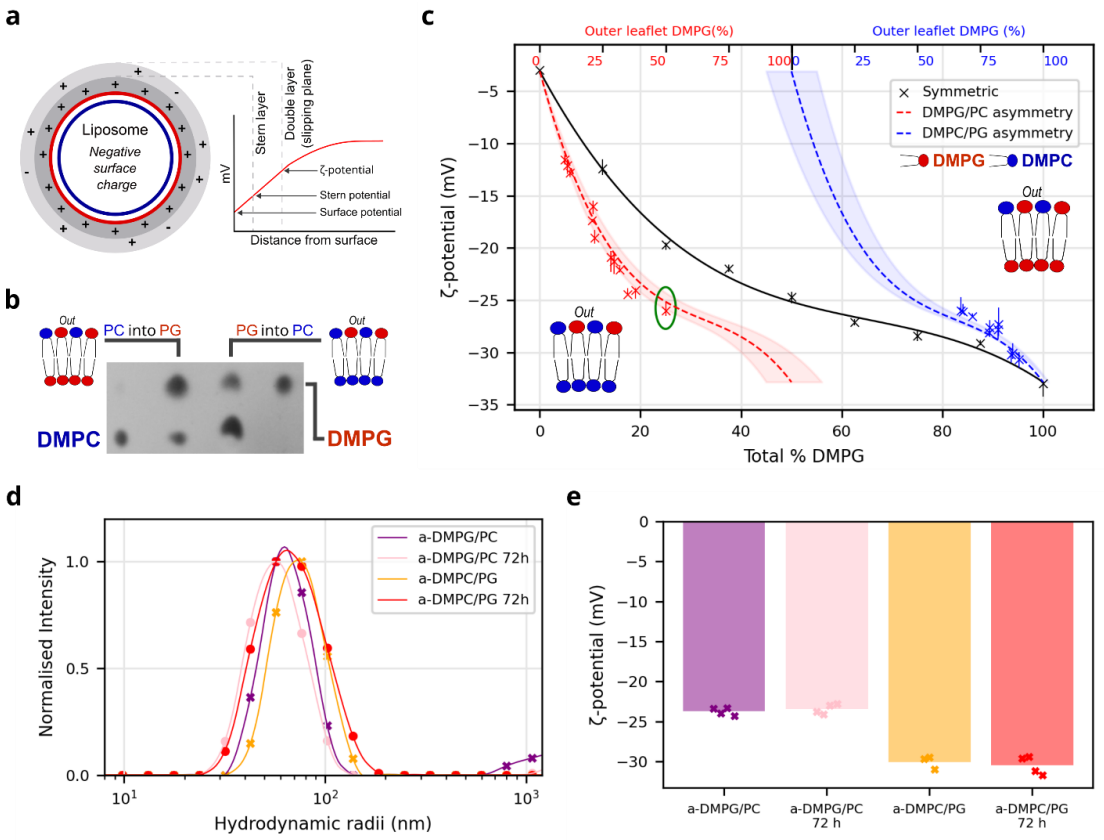


Figure 3.4: Validating liposome asymmetry using ζ-potential shows stable asymmetry in 8 M urea and up to 72 hours. (a) Physical basis of the ζ-potential as the charge on the liposome slipping plane through solution (double layer). Liposome hydrodynamic radius and ζ-potential, and hence asymmetry, is stable over at least 72 hours in the absence. **(b)** Sample TLC plate showing the introduction of DMPC lipid into DMPG liposomes by CD-mediated exchange, and *vice versa,* as indicated. Outer two lanes, DMPC (left) or DMPG (right) liposomes before exchange; inner two lanes, exchanging DMPC into DMPG liposomes (left) or DMPG into DMPC liposomes (right). **(c)** ζ-potential by lipid content for symmetric (black line) and asymmetric liposomes (blue curve: DMPC/PG, red curve: DMPG/PC). Theoretical asymmetry lines are shown with 10% error margin. Generated asymmetric liposome samples (DMPC/PG: red, DMPG/PC: blue) are shown with range bars from repeat ζ-potential measurements (centre is mean average, n ≥3). Green circled measurement discussed in text. **(d)** Liposome hydrodynamic radius and **(e)** ζ-potential, and hence asymmetry, is stable over at least 72 hours.

3.2.2 ζ-potential can be accurately predicted by a machine-learning model

To improve our ability to define the asymmetry of different lipid compositions using experimental measurement of the ζ-potential, a machine learning model was constructed to predict liposome ζ-potential. A gradient boosted, decision-tree architecture was adopted, using the XGBoost library[568]. Briefly, gradient boosted models work by iteratively improving an initially naive model by generating a weak-learning model ensemble that minimises the loss function, and then adaptively adjusting the weighting for each model in the ensemble based on misclassification. A simple example of this principle for classifying two different shapes is shown **Fig. 3.5a**. Using 315 data points (from this study and the literature[80,81,569–580]), the liposomes lipid composition was first parametrised to: (1) average overall charge per lipid, (2) average T_m of all lipids, and (3) fraction of cholesterol present. Combined with five additional liposome/buffer features (monovalent and divalent salt concentrations, pH, liposome size and measurement temperature) models were trained with the ζ-potential target. Following hyper-parameter optimisation, this yielded a model with an average mean absolute error (MAE) of ~3.3 mV. Further improvements to the model were gained by scaling the target weighting based on the associated measurement error of the underlying data, with an optimised scaling interval of 0.18 (centred around 0.5 – i.e., all target values are weighted between 0.41 and 0.59 according to their error) resulting in an improved model with an average MAE of ~3.1 mV (**Fig. 3.5b**). Due to the small amount of data used to train the model, to reduce the effects of noise in the output, the average prediction from a 50-model ensemble is used.

Lipid charge dominates the model, as expected given the physical basis of the ζ-potential (**Fig. 3.5c**), with monovalent and divalent salt concentrations and liposome size being the next most important parameters. Parameter ablation indicated that lipid charge, T_m and salt concentration are the most predictive features (**Fig. 3.5d**). A comprehensive parameter search demonstrated that retaining all eight model features resulted in the best performing model, although inclusion of the cholesterol fraction and the divalent salt concentration led to only very small gains in predictive power (**Fig. 3.5e**).

The prediction for DMPG and DMPC lipid mixtures in the buffer used for experimental measurements (100 mM NaCl, 20 mM Tris-Cl, pH 8.5), excluding the equivalent measured data from the training set, is consistent with the experimental data (MAE=0.86 mV, average experimental measurement range=0.88 mV) (**Fig. 3.6a-b**). DMPC and DMPG lipids are well represented in the training data, but the

model also performs reasonably well on less well-represented lipids (DMPE and DMPS), able to predict the shape of the ζ-potential trends, although with larger error than for the DMPG-DMPC mix (MAE of DMPS-DMPC: 3.5 mV, MAE of DMPE-DMPG: 1.9 mV) (**Fig. 3.6c-d**).

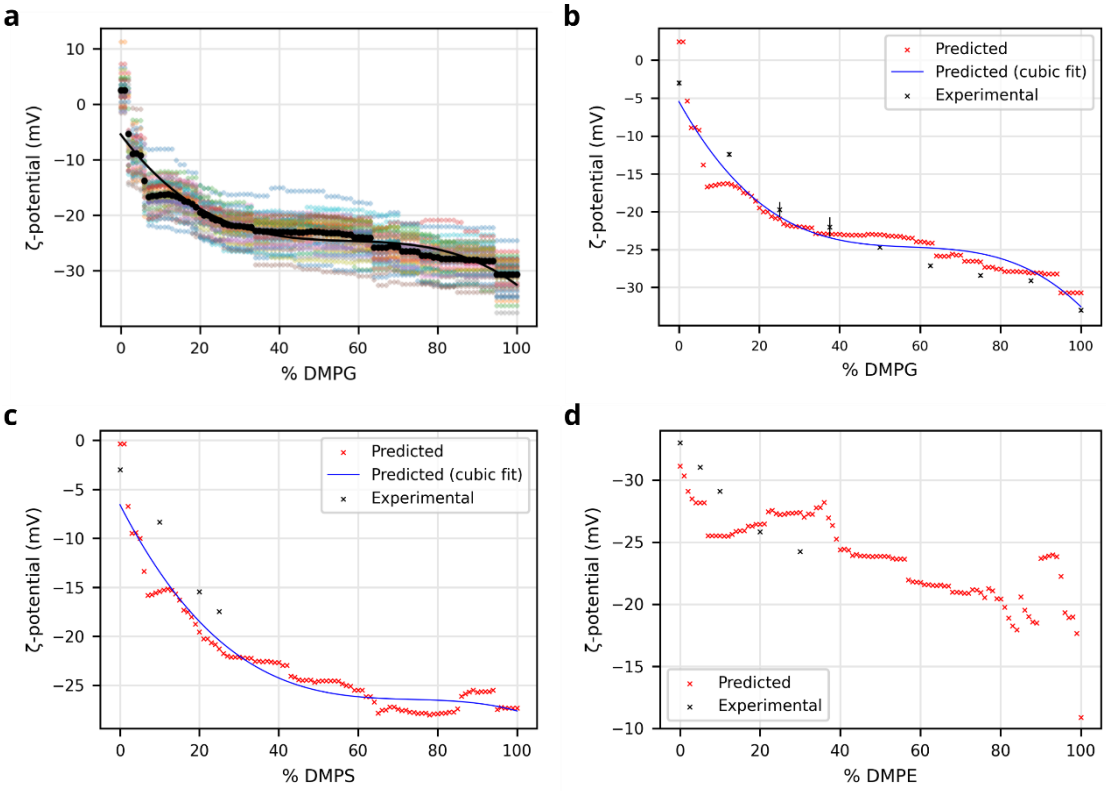


Figure 3.5: Using the ζ-potential prediction model. (a) Individual predictions (coloured crosses) and their average (black circles) from the model ensemble for DMPC-PG, and **(b)** comparison of the experimental (black) and average predicted (red) data for the same curve. The fitted cubic is also shown in both (a) and (b). **(c)** Average prediction for DMPS-PC and **(d)** DMPE-PG lipid mixes compared with the experimental data. Note that high-DMPE content LUVs (> 30-40% mol/mol) cannot be synthesised due to the strong negative curvature of DMPE.

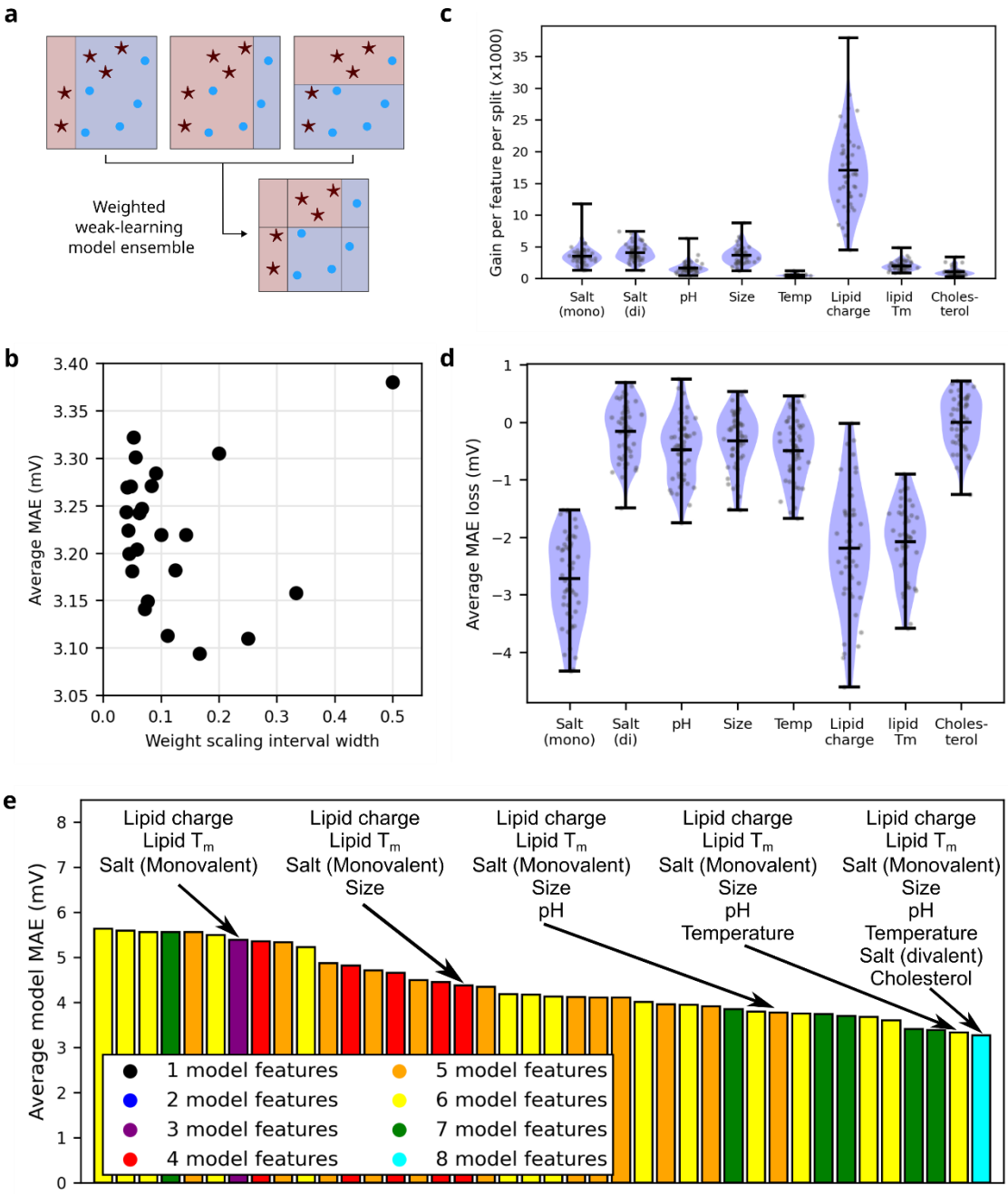


Figure 3.6: understanding the ζ-potential prediction model. (a) A simple example of gradient boosting a three-part weak-learning model ensemble combines to generate a better model weighted to minimise miscalculation. **(b)** Effect of weighting the ζ-potential target values by their associated measurement error. Scaling intervals are centred on 0.5. An interval of 0.18 yields the best model improvement – i.e. all target values are weighted between 0.41 and 0.59 according to their error. **(c)** Feature importance (gain per feature per split) in the liposome ζ-potential model, central bars are the data median (n=50). **(d)** Average MAE loss from single parameter ablation from the final model, central bars are the data median (n=50). **(e)** The average MAE of the best models generated using all combinations of the eight dataset features, the best models generated with a reduced feature count (labelled) indicate relative feature importance.

3.2.3 Lipid asymmetry modulates OMP folding and stability

Next the folding of model OMP OmpA into symmetric and asymmetric bilayers was systematically studied. OmpA is a well-studied model for OMP folding *in vitro*[496,581,582]. It contains two domains: an eight-stranded transmembrane β-barrel and a C-terminal (natively periplasmic) water soluble domain (**Fig 3.7,** upper left), that cannot cross the bilayer (thus ensuring unidirectional membrane insertion[325], confirmed via trypsin cleavage, **Fig. 3.7**, lower left), but has minimal effect on the folding kinetics of the transmembrane region[496]. This allows the effects of lipid asymmetry on the observed rate of OmpA folding and stability to be untangled from directionality.

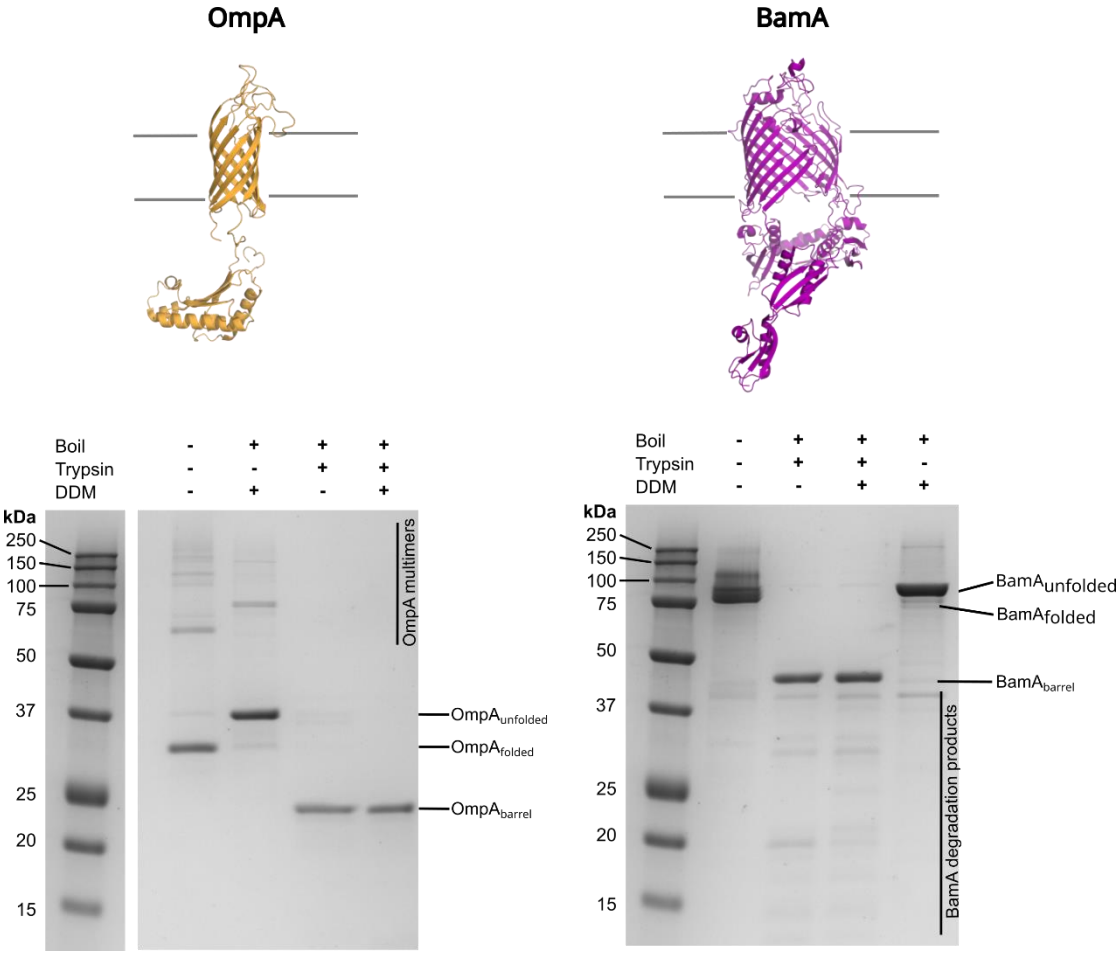


Figure 3.7: OmpA and BamA fold unidirectionally. OmpA and BamA were folded into DMPC liposomes and each were then incubated with trypsin (1000:1 molar ratio substrate:trypsin) overnight and compared to DDM-solubilised and trypsin cleaved samples treated identically. Both OmpA and BamA each show complete cleavage of their periplasmic domains in DMPC liposomes, indicating that they have folded unidirectionally into the bilayer with their water-soluble domains exposed to the bulk solvent. (PDBs: OmpA: 1G90[290], BamA: 5DO0[229]).

The transition temperature of all symmetric lipid mixes studied was measured using luardan fluorescence[255]. Briefly, this works by exploiting the differential fluorescence profile of the fluorophore when in different lipid phase (fluorescent peak shifting from 440 nm to 490 nm as liposomes move from gel to liquid phase), allowing the transition temperature to be determined – easily observed by taking the differential of the transition curves. This demonstrated that all lipid mixes were in their fluid phase at 30 °C (**Fig. 3.8**), the temperature used for all following kinetic and stability measurements. When using lipids like DMPS or DMPE which have significantly higher

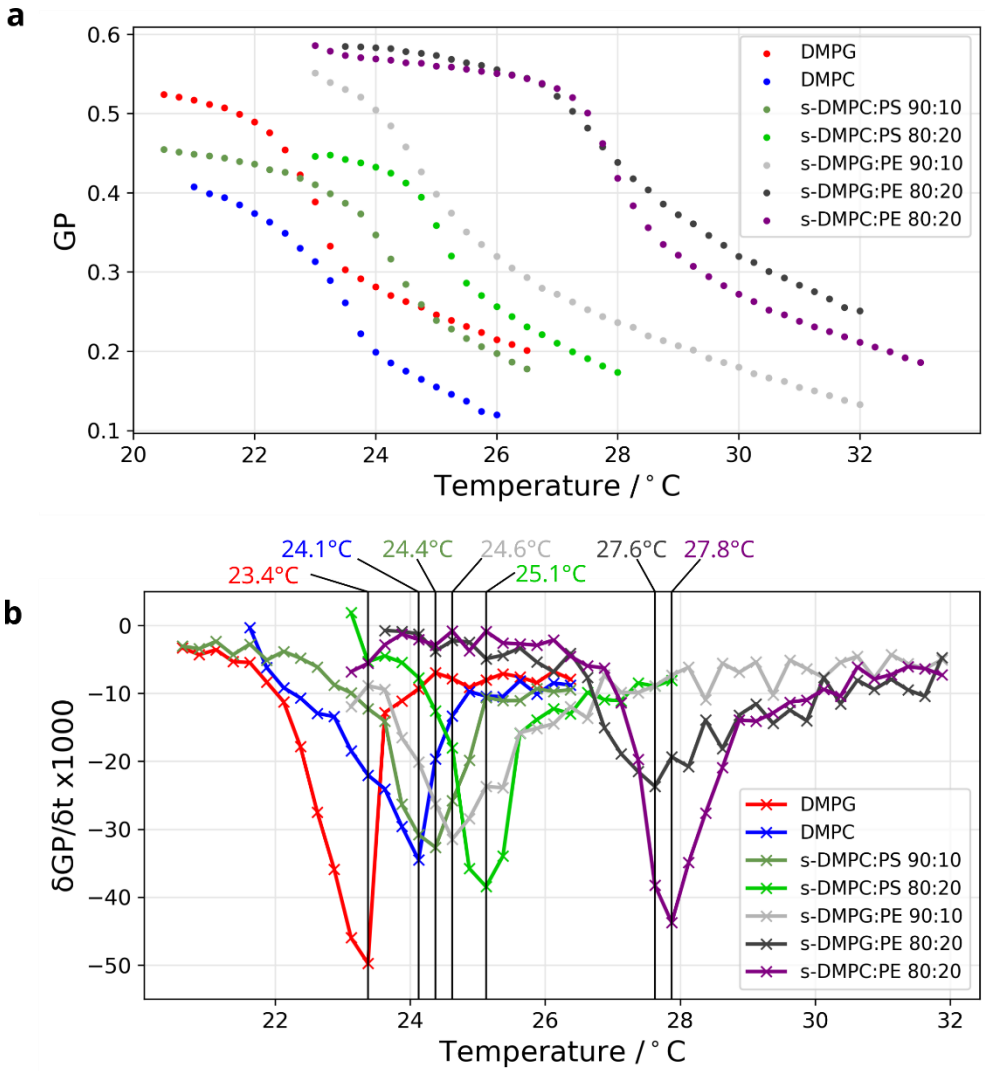


Figure 3.8: Global lipid phase transition behaviour for liposomes used in this study, measured using laurdan fluorescence. (a) The GP (generalised polarisation) ratio of fluorescence at 440 and 490 nm (see Methods) against temperature for pure DMPC and pure DMPG liposomes, and symmetric DMPS-DMPC, DMPE-DMPG and DMPE-DMPC lipid mixes, as indicated, measured using 0.25 °C intervals. **(b)** The first derivative of the GP, with the implied T_ms.

phase transition temperatures (35 °C and 50 °C respectively), no more than 20 % was combined with DMPC or DMPG in liposomes. Thus, regardless of composition, all liposome systems described below were in the fluid lipid phase during the assays described. While the differential presence of charged lipids is expected to cause small changes to the phase between leaflets, and the associated change in mechanical properties, evidence from the literature suggests these will be minimal, and significant trans-leaflet coupling is retained, with properties of the whole membrane remaining consistent rather than each leaflet diverging[583,584].

Folding kinetics were measured by monitoring the change in tryptophan fluorescence that occurs as the protein partitions from its unfolded state (solubilised in urea) to its folded, membrane embedded state, an established method for measuring OMP intrinsic folding (e.g. [225]). Folding was initialised by rapid dilution of unfolded OMP in 8 M urea to the desired urea concentration (0.48 M for folding assays) in the presence of liposomes. The data was fitted to a single exponential curve (except membranes containing DMPS for which a double exponential fit gave much reduced error residuals) to derive the observed rate constant of folding (e.g. **Fig. 3.9a-b**). As expected[33], OmpA folds efficiently (folding yield ~80%) into symmetric DMPC liposomes with a rate constant (k_{obs}) of ~0.5x10^{-3} s^{-1} (**Fig. 3.9c**, blue). Addition of 10% DMPG into both leaflets (i.e. maintaining leaflet symmetry) slows folding slightly (40% lower k_{obs}), while higher (symmetric) concentrations of DMPG accelerate folding (~5-fold higher k_{obs} at 40 % DMPG). Asymmetric membranes produce strikingly different results. In liposomes containing ≥ 20 % DMPG in their outer leaflets and pure DMPC in their inner leaflet, OmpA failed to fold within 15 hrs (0.48 M urea) (**Fig. 3.9c,** purple).

By contrast, while OmpA folds more than 40-times more rapidly into symmetric membranes of pure DMPG compared with pure DMPC (comparing **Figs. 3.9c, d**), titrating DMPC into the outer leaflet of DMPG liposomes increases the rate constant for folding ~ 2-fold relative to symmetric liposomes with equivalent outer leaflet lipid composition, at all compositions measured (**Fig. 3.8d**). The lipid composition of each leaflet of the bilayer thus affects the rate of OmpA folding, with the effect acting in opposite directions depending on the asymmetric directionality.

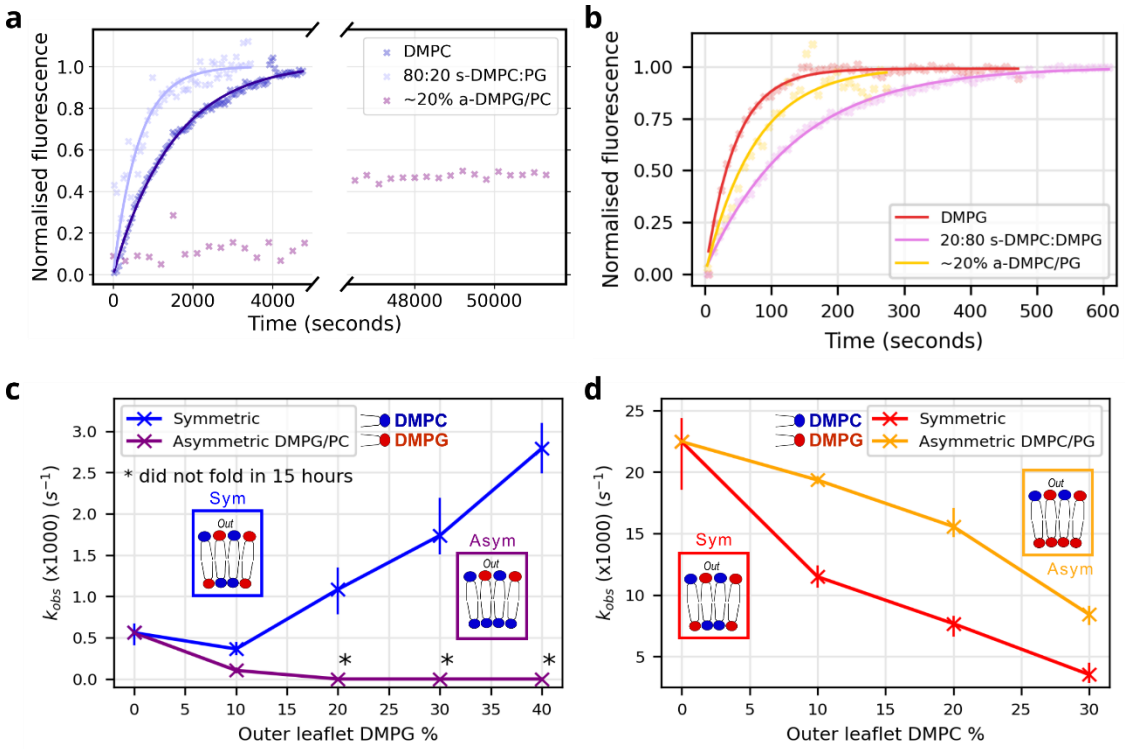


Figure 3.9: DMPC-DMPG lipid asymmetry significantly affects OmpA folding rates. (a-b) Example kinetic data and fits with representative lipid environments, data shown for DMPC, DMPG and 20% symmetric and asymmetric liposomes. Data are normalised for comparison, folding into ~20% a-DMPG/PC did not reach completion and these data were normalised to the DMPC traces. Folding rate constants (s^{-1}) of OmpA into **(c)** a-DMPG/PC asymmetric liposomes compared to symmetric liposomes with the same outer leaflet composition. Bars represent data ranges (n ≥3), * indicates the folding had not reached completion in 15 hours (< 75% folded). **(d)** Folding rate constants (s^{-1}) of OmpA into a-DMPC/PG asymmetric liposomes compared to symmetric liposomes with the same outer leaflet composition. Bars represent data ranges (n ≥3).

The stability of OmpA in symmetric and asymmetric bilayers was also assessed using cold SDS-PAGE, where the differential band-shift of the folded and unfolded forms of OmpA allow quantitation of folding at each urea concentration by gel densitometric analysis (**Fig. 3.10a**). OmpA is more stable in DMPG liposomes compared with DMPC liposomes (urea concentrations at the mid-point (P_m) of 4.5 M and 2.3 M urea, respectively) (**Fig. 3.10b**). Like other OMPs[585], membrane embedded, native OmpA is resilient to unfolding in 8 M urea in the liposomes studied here (**Fig. 3.10c, A3**) (hence equilibrium $\Delta G°_{(eq)}$ values could not be determined). While addition of small amounts (20%) DMPC into the outer leaflet of DMPG LUVs has little effect on P_m (4.5 M urea), adding 20% DMPC symmetrically into both leaflets destabilised the protein (P_m 3 M urea). Adding 10 % DMPG asymmetrically into the outer leaflet of DMPC liposomes also destabilised OmpA relative to pure DMPC liposomes (P_m 1.3 M urea), with a symmetric organisation of the same lipid composition having an even greater effect (P_m ~0.8 M urea) (**Fig. 3.9b**). These data were confirmed by assessing folding using tryptophan fluorescence, which yielded the same P_m's to within 0.5 M urea (**Fig 3.9d-e, A3-4**). Thus membrane asymmetry modulates both the rate of folding and apparent stability of OmpA. An excess of DMPG (and negative charge) in the outer leaflet slows folding and decreases P_m, while excess DMPG in the inner leaflet accelerates folding and increases P_m, compared with symmetric liposomes with the same outer leaflet lipid composition.

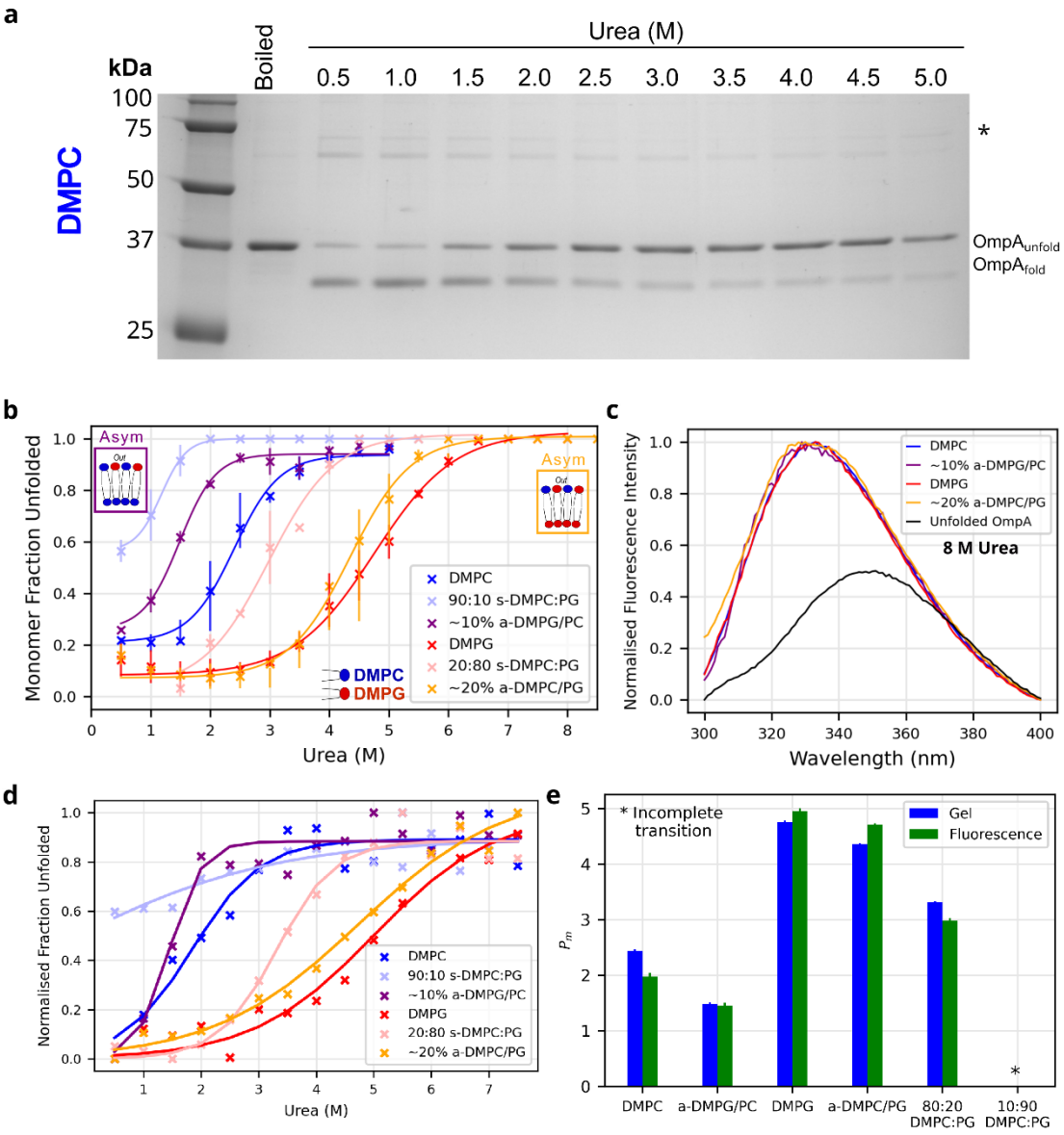


Figure 3.10: DMPC-DMPG lipid asymmetry significantly affects OmpA folding stability. (a) Example SDS-PAGE gel showing OmpA folding yields over 0.5 – 5 M urea in DMPC liposomes following incubation overnight at 30 °C. **(b)** Urea dependence of OmpA folding into DMPC-DMPG symmetric and asymmetric liposomes, as indicated. The lines are the fit to the average of at least two repeats, bars are data range. **(c)** Tryptophan fluorescence emission spectra of OmpA folded into LUVs of different composition show that the protein does not unfold after overnight incubation at 30 °C in 8 M urea in any liposome. The spectrum of OmpA unfolded in 8 M urea in the absence of lipid is shown for comparison. **(d)** Extracting the 335/350 nm ratio of each spectrum and normalising to each liposome condition shows urea transition curves that (e) agree well with the midpoints of folding (Pm) determined using cold SDS-PAGE (the differences are < 0.5 M urea, the increment size). Error bars indicate +/- goodness of fit (average difference between observed and fitted data) (n = 2 (gel) or n = 1 (fluorescence)).

3.2.4 Asymmetric modulation is dominated by charge-mediated effects

Given the similarity in T_m, area-per-lipid, and acyl-chain length of DMPC and DMPG, the asymmetric effects observed presumably arise from the different charge of the lipid head groups. To confirm this (as well as its general applicability), additional measurements were made to ensure that the observed effect was: (1) not unique to OmpA; (2) was not unique to lipids with DM-acyl chains; (3) the apparent charge effect was consistent across similarly charged lipids other than PC and PG; and (4) minimal effects were observed in lipid-asymmetric but charge-symmetric liposomes.

Firstly, to determine whether these effects are unique to OmpA, the 16-stranded OMP BamA was also studied. Like OmpA, BamA has a large (47 kDa) water soluble domain that ensures the unidirectional folding of its 43 kDa transmembrane β-barrel (**Fig. 3.7**, right). BamA folding into symmetric and asymmetric liposomes showed similar trends to OmpA, although differing in magnitude: a-DMPG/PC asymmetry slows (or abrogates) folding while a-DMPC/PG asymmetry accelerates folding, relative to symmetric systems with the same outer leaflet composition (**Fig 3.11**). Next, to assess whether the effects of lipid asymmetry on OMP folding and stability are unique to DM-lipids, we generated stable asymmetric POPC-POPG liposomes. PO-lipids have 16:0-18:1 (palmitoyl and oleoyl) acyl chains (**Fig. 3.12a**) and are thus longer than DM-lipids, as well as having much lower T_m's (-4 °C, caused by the oleoyl unsaturation), which confer additional fluidity to the membranes[586,587]. Asymmetric liposomes were generated using the same protocol as for DMPC-DMPG liposomes, and validated by TLC, ζ-potential and DLS, as described above (**Fig. 3.12b-d**).

The urea stability of OmpA showed generally the same trends in PO-lipids as in DM-lipids, but with differing extents of P_m modulation. Due to the longer acyl chains OmpA folding yield is lower than into DM-lipids (~80% compared to ~30% for DMPC and POPC, **Fig. 3.12e** c.f. **Fig. 3.10b**), consistent with the literature[588]. However, folding kinetics can be measured for the folding fraction, and rates determined. These again showed similar trends, as observed for DMPC-PG asymmetric systems: a-POPG/PC asymmetry slows (or abrogates) folding while a-POPC/PG asymmetry accelerates folding, relative to symmetric systems with the same outer leaflet composition (**Fig. 3.12f-g**), although the magnitude of the effects vary. Thus, the effects of asymmetry on folding kinetics and stability are independent of acyl-chain length, and the changes in mechanical/physical membrane properties associated with changing from DM to PO acyl chains.

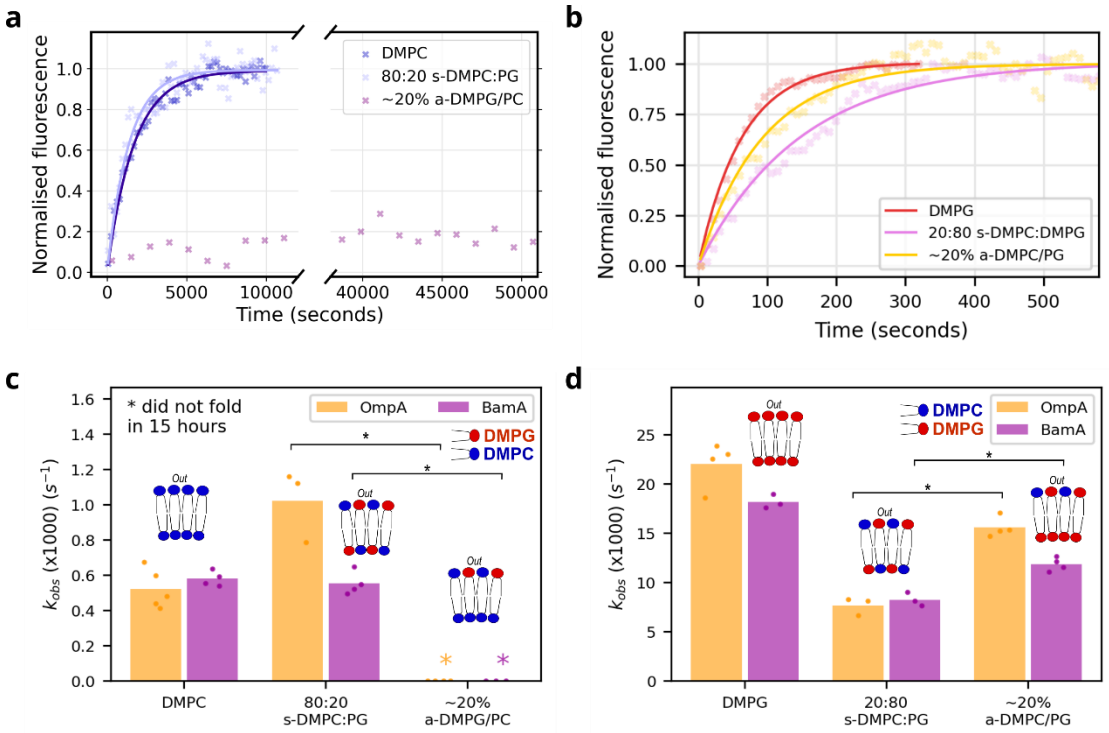


Figure 3.11: DMPC-DMPG lipid asymmetry effects BamA folding rates similarly to OmpA. (a-b) Example kinetic data and fits with representative lipid environments, data shown for DMPC, DMPG and 20% symmetric and asymmetric liposomes. Data are normalised for comparison, folding into ~20% a-DMPG/PC did not reach completion and these data were normalised to the DMPC traces. **(c-d)** Observed folding rate constant (s^{-1}) of OmpA/BamA into asymmetric and symmetric liposomes, demonstrating similar trends for the two proteins in each liposome type. * Indicates the folding was not complete (< 75% folded) in 15 hours and hence a rate constant could not be determined. Significance labels (*) p-values (left to right) for (e) 0.029, 0.015, and (f) 0.029, 0.029, determined by permutation testing.

Figure 3.12: Generating and folding OmpA into POPG-POPC symmetric and asymmetric liposomes. (a) Acyl-chain structure of DM- and PO-lipids. **(b)** Sample TLC plate showing the introduction of POPC lipid into POPG liposomes and *vice versa,* as indicated. Outer two lanes, POPC (left) or POPG (right) liposomes before exchange; inner two lanes, exchanging POPG into POPC liposomes (left) or POPC into POPG liposomes (right). **(c)** Experimentally measured ζ-potential calibration curve for asymmetric and symmetric POPC-PG lipid mixes, showing symmetric (black crosses) and asymmetric liposomes (red and blue crosses). Error bars are data range (n ≥3). **(d)** Liposome size measured by DLS. **(e)** Urea dependence of OmpA folding into POPC-PG symmetric and asymmetric liposomes. POPG and a-POPC/POPG are fitted to the average of two repeats, all other lines are to guide the eye only as there is insufficient amplitude to enable a fit (bars show the data range of two repeats). **(f, g)** Observed folding rate constant (s^{-1}) of OmpA into asymmetric and symmetric liposomes made of DM- or PO-acyl chained lipids, as indicated, demonstrating similar trends to those using DM-lipids for all membrane types (compare with **Fig. 2E, F**). * Indicates the folding was not complete (< 75% folded) in 15 hours and hence a rate constant could not be determined. Significance labels (*) p-values = 0.029, determined by permutation testing.

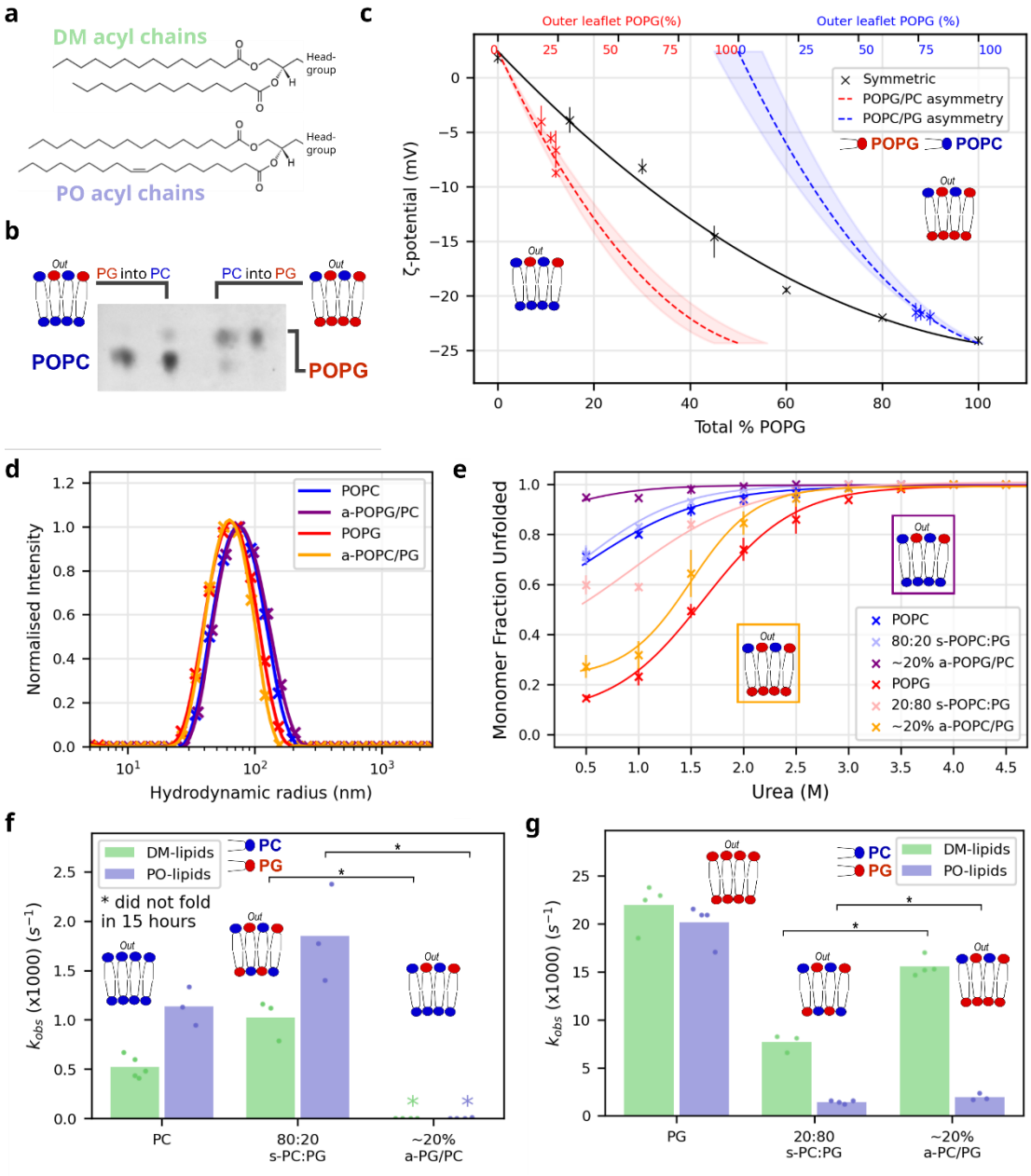


Figure 3.12: (Legend overleaf)

Thirdly, the specificity of the effect to PC/PG lipids was studied, by generating membranes with DMPS and DMPE and measuring OmpA folding kinetics and stability in these systems. Like DMPG, DMPS has a net negative charge, while DMPE, like DMPC, is net neutral. Thus liposomes can be created with different lipid compositions but the same charge dipoles across their bilayers (a-DMPS/PC is equivalent to a-DMPG/PC and a-DMPE/PG to a-DMPC/PG). DMPS and DMPE were used at low concentrations (< 20%) with DMPG or DMPC, to ensure membranes were in a fluid lipid phase (confirmed using laurdan fluorescence[255], **Fig. 3.8**, T_m of DMPS and DMPE is ~35 °C and 50 °C respectively[567]). The a-DMPS/PC and a-DMPE/PG LUVs were prepared and validated as described previously (**Fig. 3.13** and **3.14**).

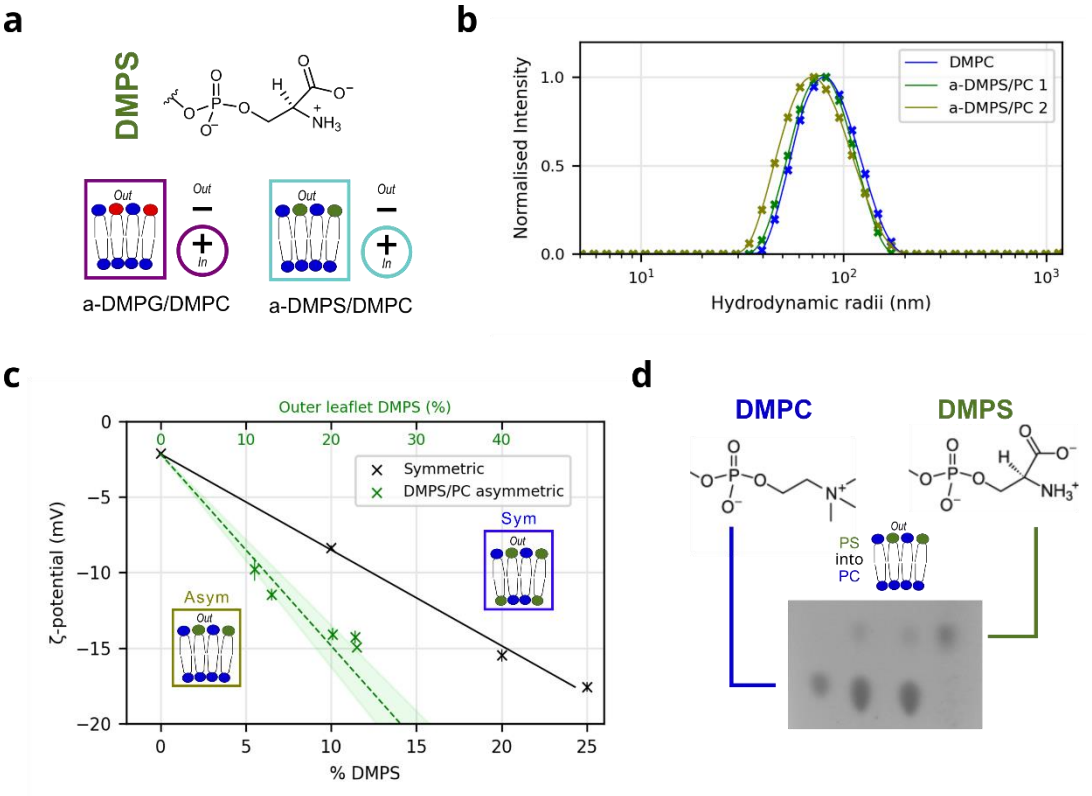


Figure 3.13: Generating asymmetric DMPS/DMPC liposomes. (a) Headgroup structure of DMPS, and the charge dipole formed by a-DMPS/DMPC liposomes – equivalent to a-DMPG/DMPC liposomes. **(b)** DLS of pre-exchange DMPC and duplicate post-exchange a-DMPS/DMPC liposomes. **(c)** Experimentally measured ζ-potential calibration curve for symmetric (black) and asymmetric DMPS-DMPG lipid mixes (green) for DMPS concentrations in the outer leaflet of 0-25%. Error bars are data range (n ≥3). **(d)** TLC of duplicate ~20% a-DMPS/DMPC exchanged (central two samples) and DMPC (left) or DMPS (right) liposomes.

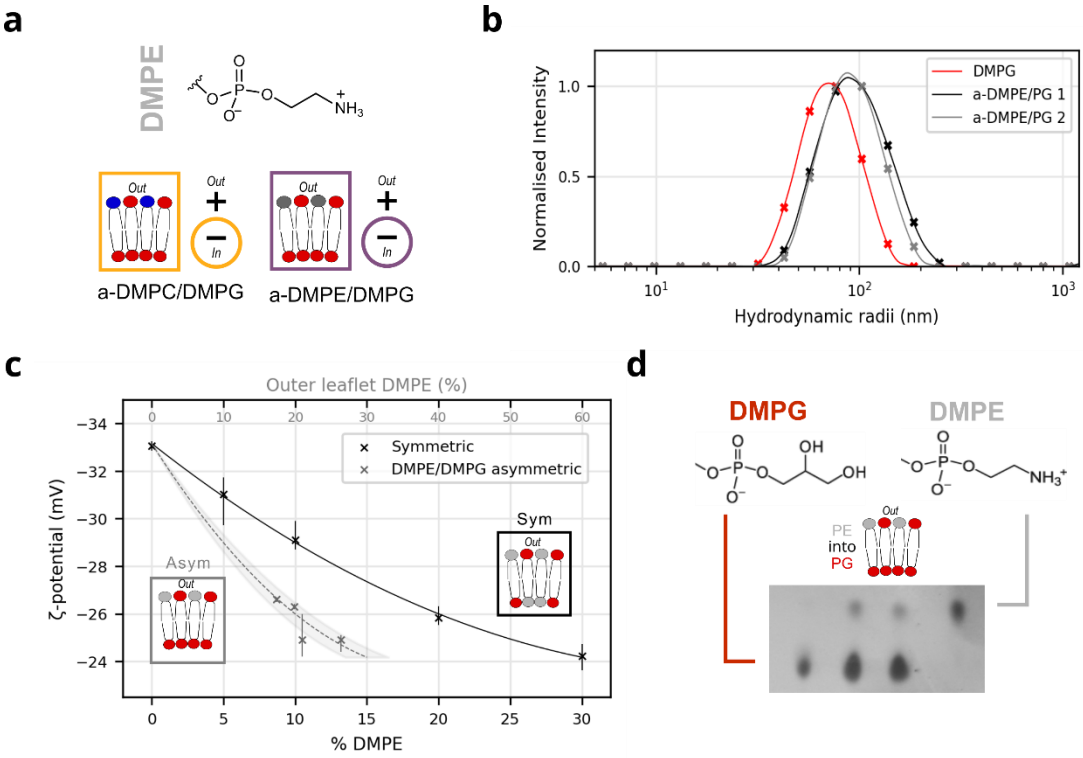


Figure 3.14: Generating asymmetric DMPE/DMPG liposomes. (a) Headgroup structure of DMPE, and the charge dipole formed by a-DMPE/DMPG liposomes – equivalent to a-DMPC/DMPG liposomes. **(b)** DLS of pre-exchange DMPG and duplicate post-exchange a-DMPE/DMPG liposomes. **(c)** Experimentally measured ζ-potential calibration curve for asymmetric DMPE-DMPG lipid mixes (grey) for DMPE fractions 0-30% in the outer leaflet. Data for symmetric liposomes are shown in black. Error bars are data range (n ≥3). **(d)** TLC of duplicate ~20% a-DMPE/DMPG exchanged liposomes (central two lanes), with DMPG (left) and DMPE (right) unexchanged liposomes.

The kinetics of OmpA folding into a-DMPS/PC and a-DMPE/PG (**Fig. 3.15a-b**) showed the same patterns as for their charge equivalent liposomes a-DMPG/PC and a-DMPC/PG (**Fig. 3.15c**). Asymmetric liposomes containing DMPS and DMPE showed a greater magnitude of the observed asymmetric effects – a-DMPS/PC liposomes retarded folding more efficiently at lower DMPS concentrations than DMPG and a-DMPE/PG enhanced folding more efficiently than DMPC – presumably due to the greater accessibility of the charge on the headgroups of these lipids compared to DMPC and DMPG. Unlike DMPG-PC lipid mixes, the stability of inserted OmpA is similar in the DMPS-PC symmetric and asymmetric membranes while the a-DMPE/PG showed the same stability trend as a-DMPC/PG, all compared to liposomes with an equivalent outer leaflet lipid composition (**Fig. 3.15d**).

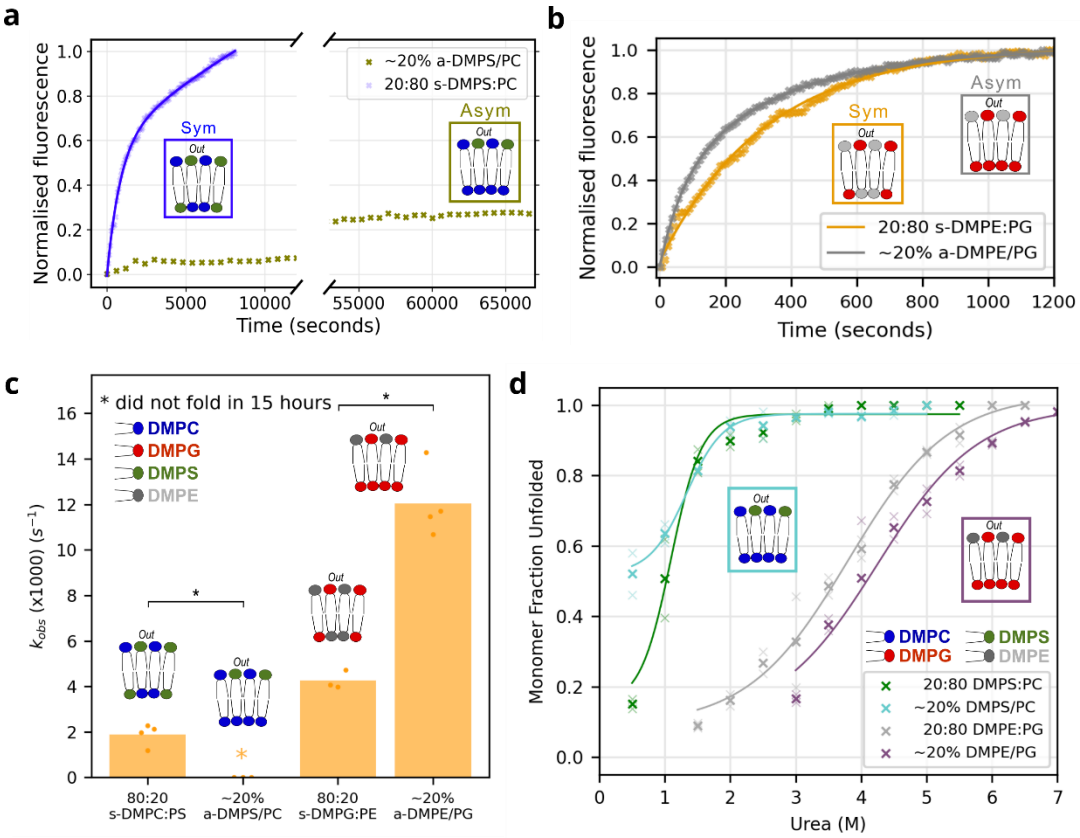


Figure 3.15: OmpA folding rates and stability in a-DMPE/DMPG and a-DMPS/DMPC liposomes are generally similar to DMPC-PG asymmetric liposomes of the same charge dipole. (a) Example OmpA folding kinetic traces, measured by tryptophan fluorescence into 20% s-DMPS:DMPC (double exponential kinetic fit, blue line) and ~20% a-DMPS/DMPC (note that the latter sample did not complete folding (< 30% folded over >15 hours, not fitted). **(b)** Example OmpA folding kinetic traces, measured by tryptophan fluorescence into 20:80 s-DMPE:DMPG (kinetic fit: yellow line) and ~20% a-DMPE/DMPG liposomes (kinetic fit: grey line). **(c)** OmpA folding rate constants (s^{-1}) into a-DMPS/PC or a-DMPE/PG LUVs and the equivalent symmetric liposomes (with the same outer leaflet content). * Indicates the folding had not reached completion in 15 hours (< 75% folded). Significance labels (*) p-values = 0.029, determined by permutation testing. **(d)** Urea dependence of OmpA folding into DMPC/PS or DMPE/PG symmetric and asymmetric liposomes. The data are fitted to the average of two repeats (line and crosses, 20% a-DMPS/PC line added to guide the eye but amplitude change too low to accurately fit). Bars are the data range of two repeats.

Finally, charge-similar but headgroup dissimilar a-DMPE/PC liposomes were generated. Owing to the same charge on DMPE and DMPC lipids, ζ-potential measurements will not inform on outer leaflet composition with sufficient resolution to accurately determine asymmetry of generated asymmetric liposomes. Instead, a FRET-based assay was used, exploiting FRET between the fluorescent lipid NBD-DPPE and solution BSA-ANS. The magnitude of the FRET between excited BSA-ANS and NBD-DPPE provides a readout of the amount of NBD-DPPE in the outer leaflet of the membrane (**Fig. 3.16a**). DMPE for exchange was doped with 1 % NBD-DPPE, and following exchange the liposomes FRET from BSA-ANS measured (**Fig. 3.16b**). Combining the FRET with the magnitude of the NBD-DPPE fluorescent signal (i.e. the total fluorophore concentration) and comparing to the calibration curves demonstrated a much stronger FRET for the exchanged samples than expected symmetric values, indicating a retention of asymmetry (**Fig. 3.16c**). While not completely quantitative, this assay clearly demonstrates that the generated liposomes are asymmetric. DMPE exchange was further confirmed by TLC, and liposome integrity by DLS (**Fig. 3.16d-e**). Despite attempts, stable a-DMPS/PG liposomes could not be generated – the lipids consistently aggregated into micron-sized aggregates, perhaps owing to the difficulties of inserting a negatively charged lipid into a negatively charged membrane.

No difference in OmpA folding kinetics or urea-stability was observed using symmetric/asymmetric DMPE/PC liposomes at 20:80 DMPE:PC (**Fig. 3.16f-g**), indicating that lipid asymmetry alone does not modulate OMP folding, whilst asymmetry in charge has a dramatic effect. Therefore, the observed effects of DMPC-PG asymmetry on OmpA folding are consistent across an other OMP (BamA), in membranes made with lipids of different acyl-chains (PO), in asymmetric membranes consisting of different lipids but with the same charge dipole, and it is not observed in asymmetric membranes without that charge dipole. Collectively, these data show that the observed asymmetric effects are mediated by the distribution of lipid charge across the bilayer, and the induced global membrane charge dipoles.

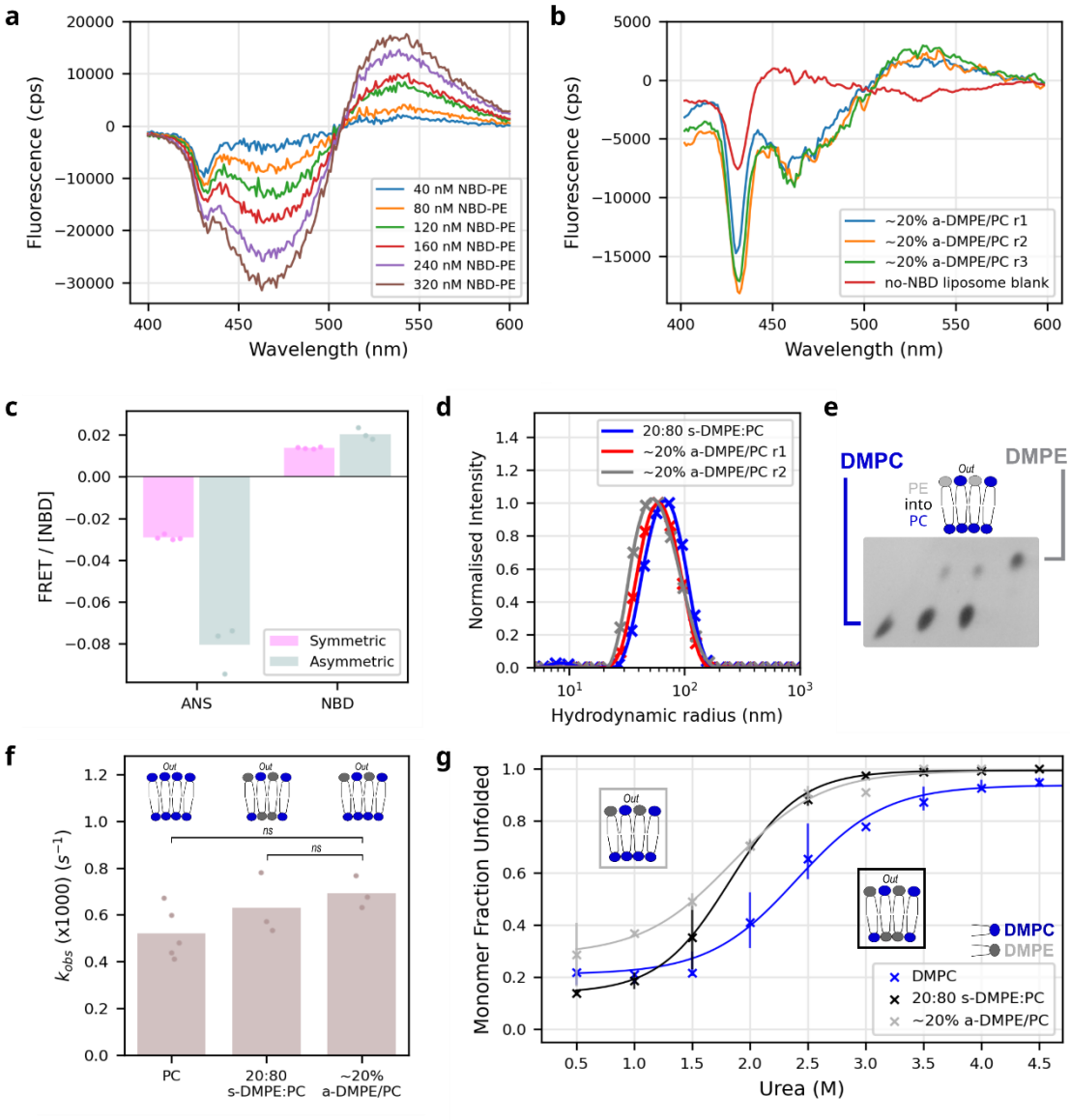


Figure 3.16: **(Legend overleaf)**

Figure 3.16: Folding OmpA into symmetric and asymmetric DMPE/DMPC liposomes. (a) Background subtracted FRET signal spectrum between BSA-ANS (bovine serum albumin (BSA) bound to the fluorescence donor aniline-naphthalene sulphonate (ANS)) and increasing concentrations of NBD-DPPE lipid (NBD: 7-nitro-2-1,3-benzoxadiazol-4-yl amino, a fluorescent acceptor) in DMPC liposomes (377 nm excitation). **(b)** Background subtracted FRET signal spectrum between BSA-ANS and NBD-DPPE asymmetrically incorporated into the outer leaflet of liposomes and control liposomes (no NBD-DPPE included), indicating clear FRET in the asymmetric liposomes. **(c)** Absolute differences between the fluorophore concentration normalised FRET signals of symmetric and asymmetric liposomes. The larger NBD signal in the asymmetric liposomes is consistent with more of the fluorophore being in the external leaflet and thus closer to the FRET donor ANS bound to BSA in solution. **(d)** DLS of symmetric and two repeats of asymmetric DMPE-PC liposomes with 20% DMPE incorporation. **(e)** TLC of duplicate ~20% a-DMPE/PC exchanged liposomes (central two lanes) with unexchanged DMPC (left) and DMPE (right) liposomes. **(f)** Observed rate constants (s^{-1}) of OmpA folding into symmetric and asymmetric DMPC-PE membranes, showing there are no significant differences in symmetric and asymmetric membranes. **(g)** Urea dependence of OmpA folding yield into DMPE-PC symmetric and asymmetric liposomes. The data are fitted to the average of two repeats (bars are the data range).

3.2.5 OmpA's extracellular loops interact with negatively charged residues

Charges on lipid headgroups are not the only charges relevant to OMP folding. OMPs typically contain charged residues in their extracellular loops, which could interact with lipid head groups to promote/stabilise the native state, and it has been shown previously that there is an enrichment of positively charged residues in these loops[224,589]. However, these charges must cross the bilayer for this to occur. To identify residues that might engage in stabilising lipid-protein interactions, coarse-grained molecular dynamics (CG-MD) was used to explore the interplay between membrane asymmetry, lipid head groups of different charge, and charged residues in the extracellular loops of natively-folded, membrane-embedded, OmpA (**Table 3.1**). Natively folded OmpA was placed into different membranes, the system minimised and equilibrated (e.g. **Fig 3.1**, inset) and simulated in five replicas of 3 µs each in all systems. Membrane and protein properties (lipid-protein contact convergence, area-per-lipid over the simulation time, average density for lipid headgroup, phosphate and acyl-chain beads and protein RMSF) were assessed to ensure simulation equilibration and stability.

For each system the number of contacts between the different lipids and each residue of the protein were calculated and normalised by the lipid concentration and simulation time to facilitate comparison. This analysis identified consistent,specific

interactions between the negatively-charged head groups of DMPG and DMPS, and three positively charged residues (R81, K94 and R124) in the extracellular loops of OmpA (**Fig. 3.17a-b**). No such interactions were found with the zwitterionic-neutral lipids DMPC or DMPE (**Fig. 3.17c**), further evidenced by calculating the average lipid density around the protein (**Fig. 3.17e**) – which shows clear enrichment of lipids localised to OmpA for DMPG and DMPS, while no such enrichment occurs with DMPC or DMPE. Calculation of the average time of occupancy of lipid at each site[556], showed that the interaction time of DMPG with R81, K94 and R124 also depends on the DMPG concentration, with R81 and K94 apparently binding the same lipid molecule (**Fig. 3.17d**, as shown in the simulation end frame in **Fig. 3.17a**). Simulations with asymmetric membranes of DMPC-PG found similar lipid-protein interaction profiles (**Fig. A5**). The necessity of these three residues for lipid-protein interactions was confirmed by *in silico* mutation of them to serine, which entirely removed the interaction to a background level (**Fig. 3.18a**). Simulations of natively folded BamA in s-DMPC:PG membranes also showed specific interactions between DMPG and residues K507, H533, K566, S764 and K793 in its extracellular loops (and K580 in an intracellular turn) (**Fig. 3.18b**). These results suggest that charge-mediated lipid-protein interactions involving the extracellular loops of OMPs could play a role in stabilising membrane embedded OMPs in their natively folded states and, thereby, contribute to the favourable driving force for OMP folding.

Protein	Membrane Composition	Box size /nm	Time /μs	Replicas
OmpA-WT	10:90 DMPC:DMPG	16x16x20	3	5
	25:75 DMPC:DMPG			
	50:50 DMPC:DMPG			
	30:70 DMPC:DMPG			
	85:15 DMPC:DMPG			
	95:5 DMPC:DMPG			
	97.5:2.5 DMPC:DMPG			
	95:5 DMPC:DMPS			
	95:5 DMPE:DMPG			
	95:5 DMPE:DMPS			
	Inner leaflet: DMPC Outer leaflet: 90:10 DMPC:DMPG			
	Inner leaflet: 90:10 DMPC:DMPG Outer leaflet: DMPC			
	Inner leaflet: DMPG Outer leaflet: 10:90 DMPC:DMPG			
	Inner leaflet: 10:90 DMPC:DMPG Outer leaflet: DMPG			
OmpA-M3	95:5 DMPC:DMPG			
BamA-WT	95:5 DMPC:DMPG	20x20x20	3	5

Table 3.1: Summary of all simulations run in this chapter. Membranes are lipid symmetric unless inner and outer leaflets are specified. Membrane composition fraction is by lipid number. OmpA-M3 is OmpA with R81S, K94S and R124S mutations.

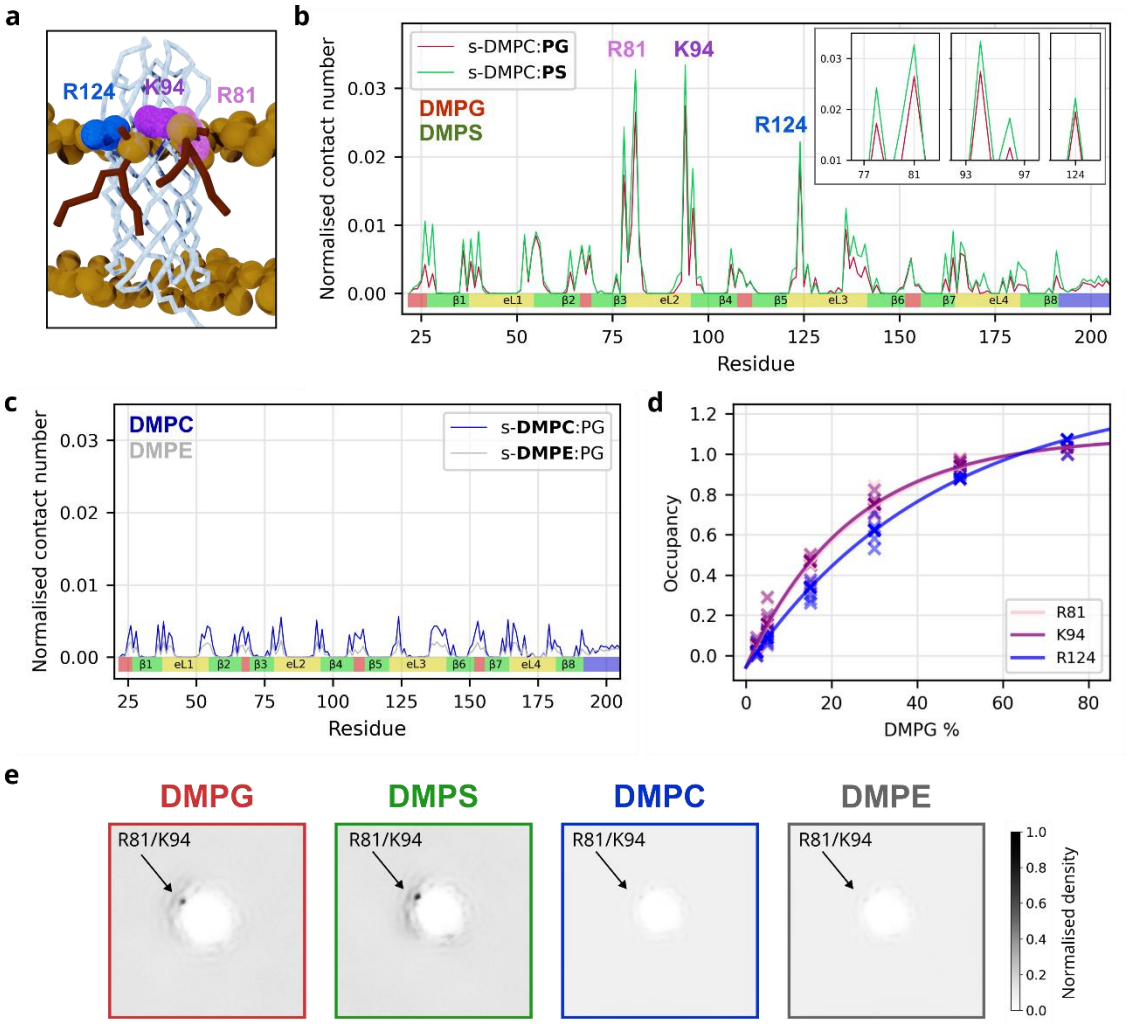


Figure 3.17: The extracellular loops of OmpA specifically interact with negatively charged lipids. (a) Final frame from a CG-MD simulation of native OmpA in s-DMPG:DMPC membranes, showing two DMPG molecules (red) in the outer leaflet interacting with OmpA at R81, K94 and R124 (space fill in different colours). **(b)** Normalised contact count (number of interactions between each type of lipid and each protein residue normalised by lipid concentration and simulation frame number) between residues in the transmembrane region of OmpA and the negatively charged lipids DMPG (red) or DMPS (green). **Inset:** Expanded views of the peaks around the three lipid interacting residues R81, K94 and R124. **(c)** Normalised contact count between the transmembrane region of OmpA and the zwitterionic lipids DMPC (blue) or DMPE (grey). Contacts are calculated over the average of five replicates. The secondary structure of the OmpA β-barrel is shown below (strands (green), extracellular loops (yellow), intracellular turns (red) and 14 residues of the periplasmic soluble domain (blue). **(d)** DMPG occupancy (fraction of time DMPG interacting with R81, K94 or R124) at different ratios of DMPC:DMPG determined from the lipid residence time. Data for five replicates are shown. **(e)** Average lipid density plots (values normalised by lipid concentration) for the phosphates of DMPE, DMPG, DMPS and DMPC over the trajectory, following centring and fitting of the transmembrane region of OmpA. DMPC and DMPG calculated from 95:5 s-DMPC:DMPG, DMPS from 95:5 s-DMPC:DMPS and DMPE from 95:5 s-DMPE:DMPG systems.

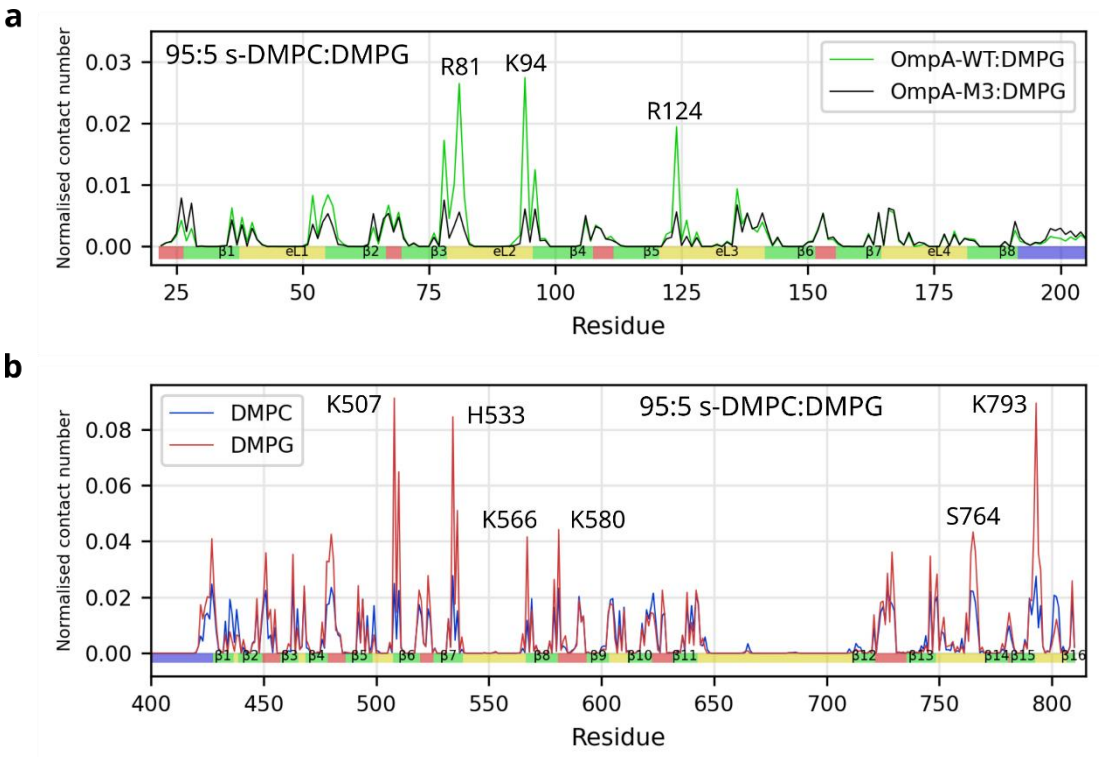


Figure 3.18: OmpA-M3 residues are required for negative lipid binding and BamA also specifically interacts with negative lipids. (a) Normalised lipid-protein contact counts (number of interactions between each type of lipid and each residue in the transmembrane domain of OmpA normalised by lipid concentration and simulation frame number) for the transmembrane region of OmpA-WT and OmpA-M3 in a 95:5 s-DMPC:PG membrane. Substitution of these three positive residues with Ser eliminates specific DMPG binding. **(b)** Full length BamA was simulated in a 95:5 s-DMPC:PG system. Only the transmembrane region is shown for clarity. Interactions with a normalised contact number >3σ are labelled, and indicated in main text. Structural features are shown at base of plot (strands (green), extracellular loops (yellow) intracellular turns (red) and 24 residues from POTRA5 (blue)).

3.2.6 Matching of lipid-OMP charge is critical for efficient folding

The extracellular loops of native OmpA contain seven positively charged residues (R81, K85, K94, R124, K128, K134, R177), and seven negatively charged residues (D41, E53, E89, D126, D137, D170, D179), many of which are highly conserved (**Fig. 3.19**) including the three lipid-interacting residues (R81, K94 and R124) identified by CG-MD above. To investigate the role of OMP loop-lipid charge interactions experimentally, four variants of OmpA that differ in their extracellular loop charge were created: OmpA-NP (No Positives, loop charge -7); OmpA-NN (No

Negatives, loop charge +7); OmpA-NC (No Charges), and OmpA-M3 (R81S, K94S & R124S) (see Methods for sequences). The folding rate and apparent stability of these variants folding into symmetric and asymmetric DM-liposomes was then determined (**Fig. 3.20** and **3.21**).

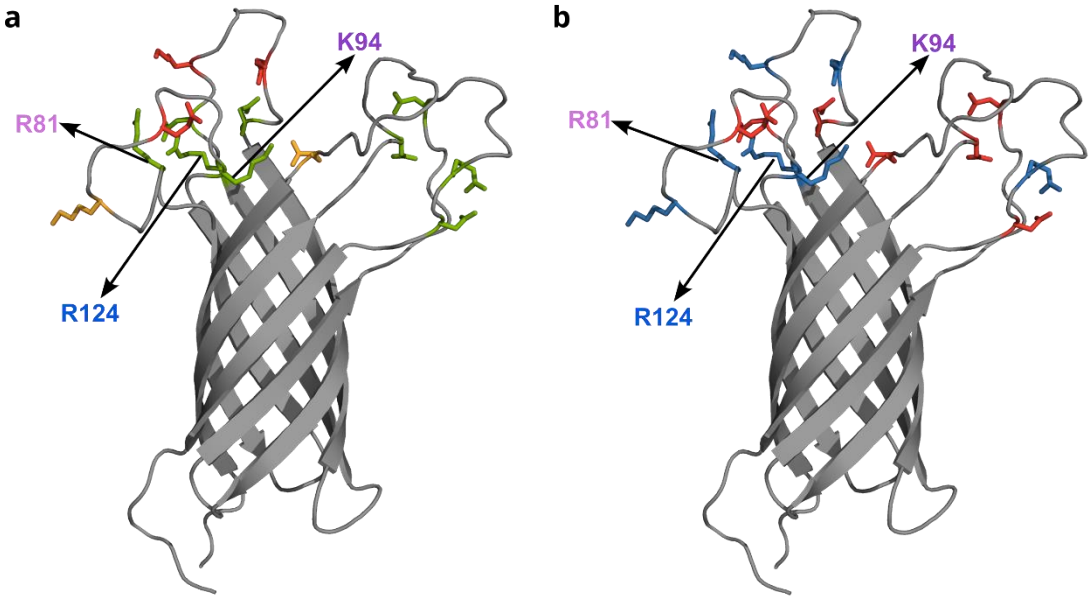


Figure 3.19: Conservation and location of Lys/Arg positively and Asp/Glu negatively charged residues in the extracellular loops of OmpA. (a) Relative conservation of charged residues (green: well conserved (>99%), yellow: partially conserved (>90%), red: less conserved (<90%)) over 2750 OmpA β-barrel sequence homologs. Conservation is the retention of a K/R or D/E at a given position. Note that R81, K94 and R124 are highly conserved (but are not the only highly conserved residues in the loops). **(b)** Spatial distribution of positively (Lys/Arg) and negatively charged (Glu/Asp) residues in the extracellular loops of the NMR structure of OmpA (blue: positive, red: negative). R81, K94 and R124 that specifically interact with negatively charged lipids are labelled. (PDB: 1G90[290], note that the z-axis locations of R81, K94 and R124 are highly consistent across all solved structures, including those solved by NMR and X-ray crystallography (**Fig 3.24**).

These experiments revealed that translocating either positively or negatively charged loops of OmpA across a bilayer constitutes a major barrier to folding irrespective of the charge orientation of the bilayer. Hence, OmpA-NC folds more rapidly than OmpA-WT, OmpA-NP and OmpA-NN in the majority of bilayers tested. However, and importantly given that OMP extracellular loops typically contain charged residues, the presence of positive charge consistently favours rapid folding compared to its absence (i.e. OmpA-WT and OmpA-NN fold more rapidly than OmpA-NP in all lipid types) (**Fig. 3.20a-c** and **3.21a-c**, left). Indeed, OmpA-NN folds > 9x more rapidly than OmpA-NP in all bilayer types. Neutralisation of the three lipid-interacting M3 positive residues retards folding to a similar extent as neutralising all seven positively charged residues (OmpA-NP), demonstrating the key importance of these three residues in folding kinetics.

Similar trends were observed on protein stability. An excess of loop negative charge destabilises OmpA-NP compared with OmpA-WT, and an excess of positive charge stabilises OmpA-NN compared with OmpA-WT in all lipid environments (**Fig. 3.20a-c** and **3.21a-c**, right). Again, OmpA-M3 mirrors the behaviour of OmpA-NP. Switching loop charge can also have different effects on the folding rate and apparent stability. For example, OmpA-NN and OmpA-WT tend to fold at similar rates in DMPG-rich membranes (**Fig. 3.21**, left), but OmpA-NN is significantly more stable (**Fig. 3.21,** right), likely due to favourable electrostatic interactions with the negatively charged lipid. Collectively, these results highlight the importance of the positively-charged loop residues in facilitating OmpA's translocation across the bilayer, and then stabilising the native protein once folded into the membrane. For OmpA, this effect is dominated by the three, highly conserved, M3 residues.

Directly comparing the folding kinetics and urea stability of OmpA-NN and OmpA-NP in symmetric and asymmetric DMPC-DMPG bilayers informs about how the differently charged proteins interact with the various charge distributions and dipoles across the bilayer (**Fig. 3.22**). OmpA-NN folds more rapidly into symmetric membranes of 90:10 DMPC:PG compared with ~10% a-DMPG/PC in which only the outer leaflet of the bilayer contains the negatively charged lipid (**Fig. 3.22a**, left). This suggests a rate enhancing interaction between the protein positive loops and DMPG negative charge in the inner leaflet of the bilayer. By contrast, OmpA-NP folds very slowly into both of these membrane types (**Fig. 3.22a**, right). For both OmpA-NN and OmpA-NP folding is faster into a-DMPC/PG than its symmetric counterpart (s-DMPC:PG 20:80), although folding is more rapid overall for OmpA-NN since it

contains positively-charged loops. Bilayer lipid charge asymmetry also affects stability. For example, OmpA-NP is more

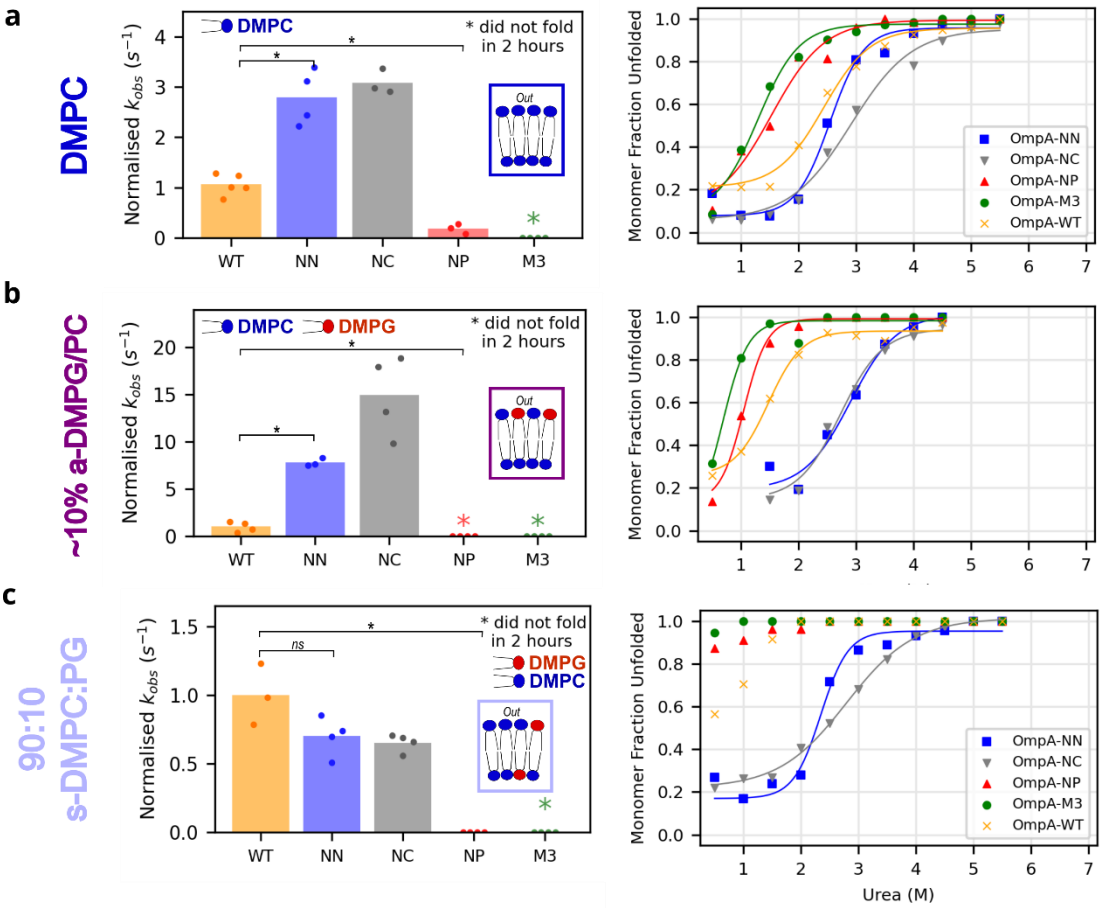


Figure 3.20: Folding kinetics and stability of OmpA charge variants compared to WT-OmpA for DMPC based symmetric and asymmetric lipid environments. The relative folding rate constant (normalised to WT) and urea-titration stability curves for OmpA variants measured using cold SDS PAGE in **(a)** DMPC (p-values: WT-NN: 0.008, WT-NP: 0.018), **(b)** a-DMPG/PC (the fit for OmpA-M3 and -NP in urea are included to guide the eye, but the stability is too low to accurately fit the data) (p-values: WT-NN: 0.029, WT-NP: 0.014), and **(c)** 90:10 s-DMPC:PG (p-values: WT-NN: 0.086, WT-NP: 0.029). All p-values determined by permutation testing.

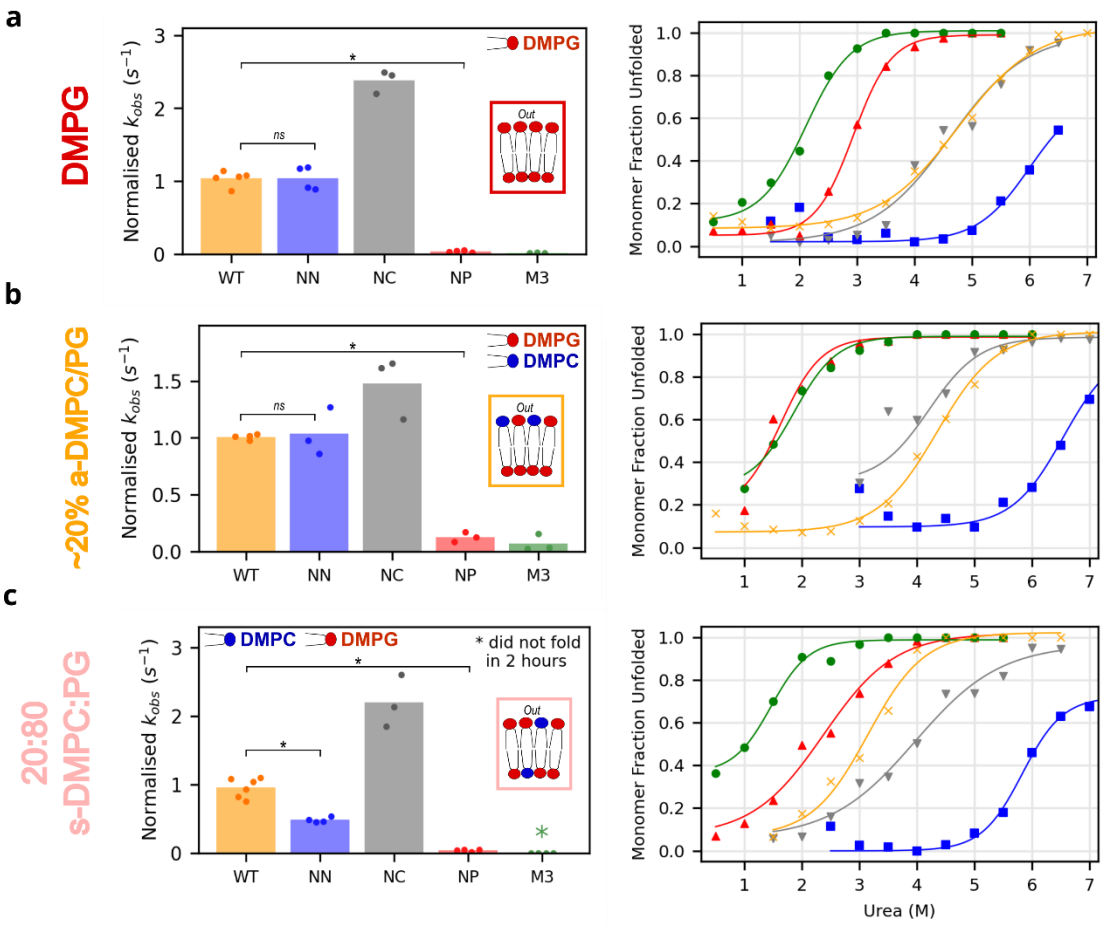


Figure 3.21: Folding kinetics and stability of OmpA charge variants compared to WT-OmpA for DMPC based symmetric and asymmetric lipid environments. The relative folding rate constant (normalised to WT) and urea-titration stability curves for OmpA variants measured using cold SDS PAGE in **(a)** DMPG (p-values: WT-NN: 1.0, WT-NP: 0.008), **(b)** a-DMPC/PG (p-values: WT-NN: 0.829, WT-NP: 0.029), OmpA-NC fraction folded at 3.5 M urea was excluded from the fit in (b), and **(c)** 20:80 s-DMPC:PG (p-values: WT-NN: 0.005, WT-NP: 0.005), as indicated in the key. All p-values determined by permutation testing.

stable in s-DMPC:PG (20:80) than in its symmetric bilayer counterpart, while the protein is less stable in s-DMPC:PG (90:10) than a-DMPC/PG (**Fig. 3.22b**, left). Similarly, OmpA-NN is more stable in ~20% a-DMPC/PG compared to the equivalent symmetric membranes, while OmpA-NP shows the opposite effect (**Fig. 3.22b**, **r**ight). While many details of the complex interplay between lipid charge asymmetry and OMP loop charge remain to be determined, these data unambiguously show that the efficient folding and stability of OmpA depend on positive charges in its extracellular protein loops, and charge asymmetry between the two leaflets of the target bilayer.

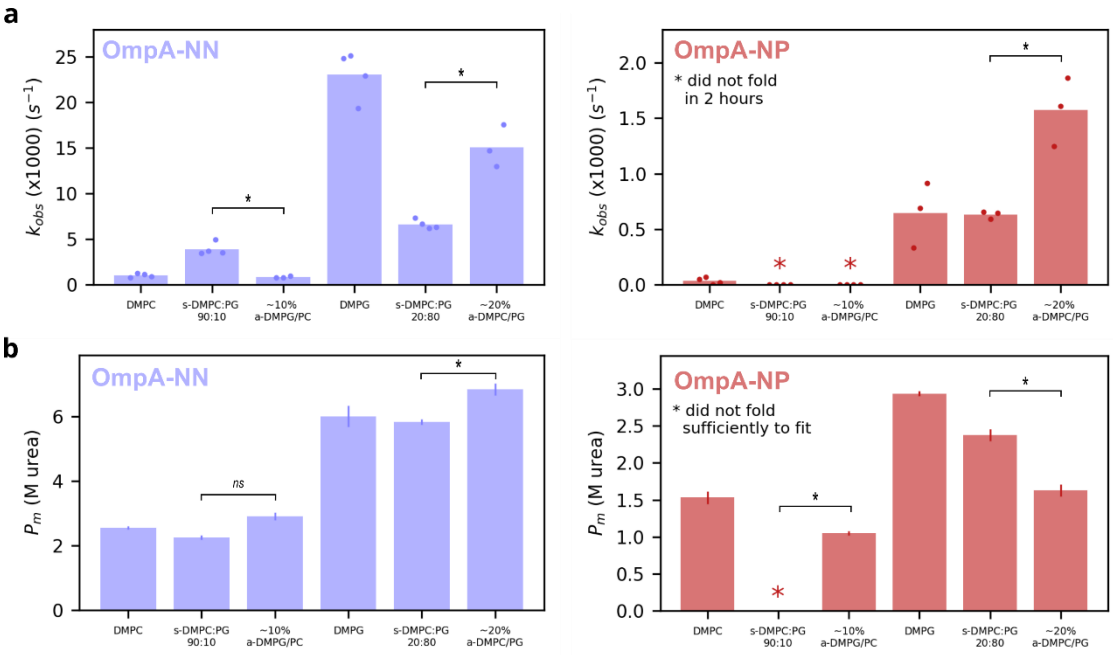


Figure 3.22: Folding kinetics and stability of the OmpA-NN and OmpA-NP. (a) Un-normalised folding rate constants of OmpA-NN and OmpA-NP into symmetric or asymmetric liposomes, as indicated, demonstrating the different patterns of folding rate observed for the different OmpA charge variants. Note the difference y-axis scale in the two plots. Significance labels (*) p-values = 0.029, determined by permutation testing. **(b)** P_m values for OmpA-NN and OmpA-NP into symmetric and asymmetric liposomes, as indicated. Error bars represent the goodness of fit to the data shown in **Fig. 3.21** and **3.22** (standard deviation of the P_m's estimated from the covariance of fitted parameters), centre is the fitted parameter value. Significance labels (*) p-values from left to right: 0.020, 0.031, 0.016; (ns indicates no significant difference, p-value = 0.423), determined by a two-tailed paired t-test.

3.2.7 OMPs have a conserved, positive extracellular charged region

As noted above, the enrichment of positive charges in the extracellular loops of OMPs has been described previously[224,589] and given the importance of positive charge for OmpA folding and stability in the data presented above, the distribution of charged residues in OMPs extracellular loops was further analysed. The Orientations of Proteins in Membranes (OPM[590]) database, contains all experimentally solved OMPs, embedded in a membrane and with the barrel and z-axis aligned. Following sequence clustering (70 %), space was divided into 1 Å slabs parallel to the membrane plane and the spatial enrichment of different residues calculated. However, while this identified the well-characterised OMP aromatic girdle[224] that flanks the acyl chains on each side of the membrane (**Fig. 3.23a**), patterns of charged residue distributions were not obvious from this analysis, presumably because different residue probabilities in the transmembrane and water-soluble regions of the protein skew the enrichment statistics. The analysis was therefore repeated, determining residue enrichments in the extracellular loops relative to soluble regions of the protein only (**Fig. 3.23b**), revealing a patch of (>2σ significant) positive residues 6-10 Å above the plane of the membrane's outer leaflet.

While the OPM database is information rich with experimental structural information, it contains relatively few OMPs (75, non-redundant). We therefore interrogated OMP structures predicted by Alphafold2[591,592] (database version 3) and sequences from the OMPdb[593] database, a repository of 1.3×10^6 putative OMP sequences, identified using low confidence threshold cutoffs. Quality filtering and sequence clustering yielded 343 Alphafold2 structures and 19, 055 OMPdb sequences of OMPs (residues per bin are shown in **Fig. A6**). Sequence data from the OMPdb lacks explicit structural information, but approximate distances can be estimated using the residue count from the membrane centre combined with a calibration curve calculated from proteins in the OPM dataset (**Fig. A7**). These analyses also show a peak (> 2σ) for enrichment for positive residues ~8 Å from the membrane surface (**Fig. 3.23c-d**). No consistent pattern emerged for negatively charged residues (**Fig. A8**).

Collectively, these analyses identify an enrichment of positive residues in the extracellular loops ~6-10 Å from the membrane surface, which we term the *Positive Extracellular Patch* (PEP). Comparing the locations of charged residues in the extracellular loops of OmpA and BamA with the positive residue enrichment curve derived from the OPM data shows that both these OMPs fit the pattern with 4/7 and

5/10 of their Lys/Arg residues respectively falling with this region (**Fig. 3.24**). Given how strongly the M3 set of positive charges, all of which are in the PEP, impact the folding of OmpA, it seems possible that these charges are conserved to facilitate efficient OMP folding into the OM.

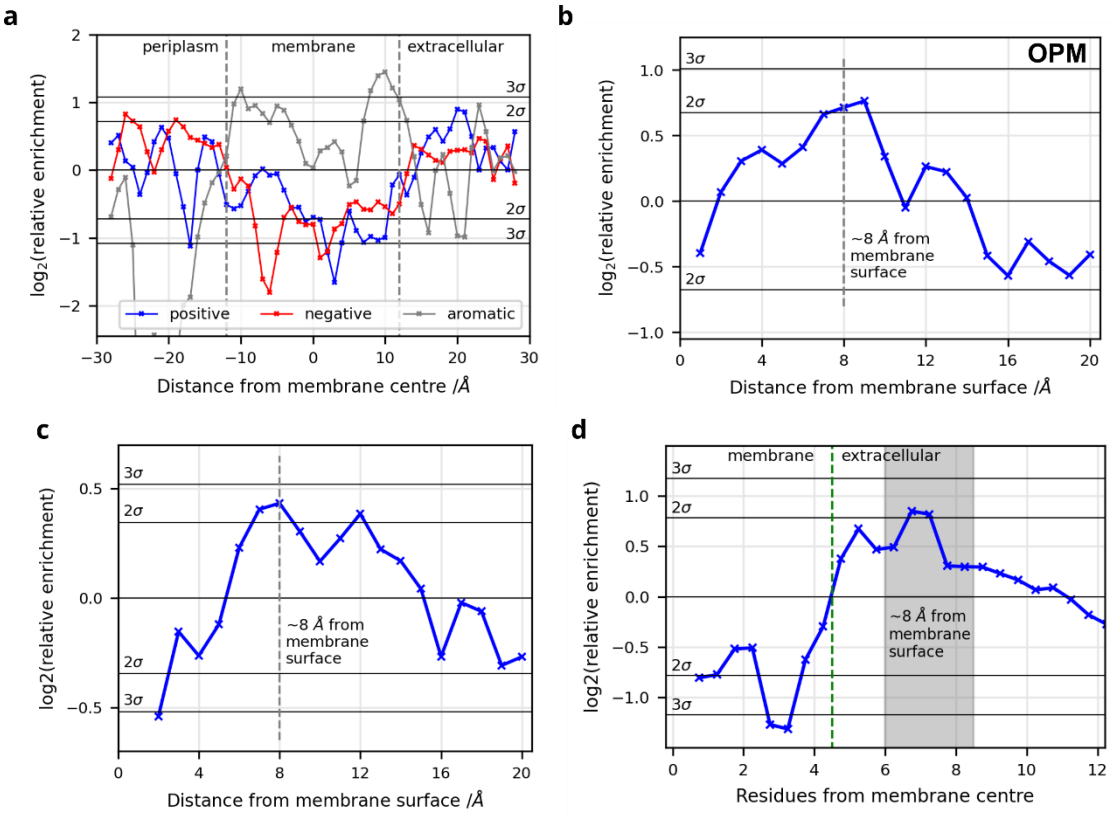


Figure 3.23: OMP residue enrichments perpendicular to the membrane plane show a conserved enrichment of positively charged residues in the extracellular loops ~ 8 Å from the membrane surface. (a) Residue enrichments of aligned OMPs from the OPM database (experimentally solved structures) relative to the probability of finding an amino acid randomly, calculated over the whole protein sequence. Membrane thickness is the average of all OPM structures. **(b)** Residue enrichments of Lys/Arg in the extracellular loops of OMPs relative to the probability of finding an amino acid from the soluble regions of the protein sequence calculated from proteins in the OPM database (i.e. transmembrane residues are omitted from this analysis). **(c)** Residue enrichments for structures from the Alphafold2 database, and **(d)** sequence data from the OMPdb (see **Fig. A7** for the residue count from the membrane centre to approximate distance calibration). The dashed green line in (**d**) indicates the approximate membrane hydrophobic-hydrophilic boundary.

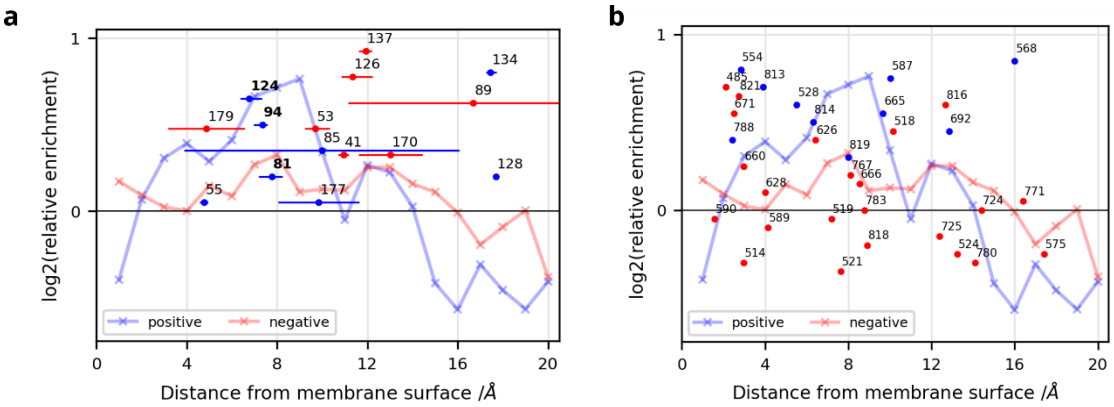


Figure 3.24: Distribution of charged residues in OmpA and BamA show a positive region ~8 Å from the membrane. (a) OmpA and **(b)** BamA charge distribution matches the bioinformatic profile. Residues R81, K94 and R124 in OmpA, identified as lipid interacting by CG-MD (OmpA-M3 cluster), are shown in bold. OmpA distances are calculated as the average of the solved *E.* coli OmpA structures (PDB 1G90, 1QJP, 1BXW). Error bars indicate maximum and minimum values. Positive residues are shown as blue circles, negative residues as red circles and are labelled with residue number above-right of the marker.

3.2.8 Protein-induced membrane charge asymmetry

Given the large effects observed on OMP intrinsic folding into membranes with charge dipoles induced by lipids, and the well-conserved PEP, it was next considered whether a protein induced charge dipole was sufficient to alter the folding kinetics. OmpA-NN, OmpA-NC or OmpA-NP were folded into DMPC liposomes at an LPR of 1:320 to generate membranes with charge dipoles, induced by the extracellular loop charges (OmpA has one positively and three negatively charged residues in its intracellular turns, the soluble domain is assumed to be negligibly contributing). The magnitude of the protein induced dipole is challenging to determine, but the direction is clear: DMPC:OmpA-NN is more positive while DMPC:OmpA:NP is more negative inside the liposomes. (**Fig. 3.25a**). Despite the better folding and stability of OmpA-NN in a large excess of lipid (see **Fig. 3.20**, **3.21**), with a low LPR it folds to only ~50% while OmpA-NC/NP fold ~80%, equivalent to OmpA-WT (**Fig. 3.25b**, **c**), suggesting that the presence of high concentrations of OmpA-NN is inhibiting folding more than the other OmpA variants. Proteoliposomes were separated from unfolded protein by low and high speed centrifugation and the liposome integrity confirmed by DLS (**Fig. 3.25d**).

The folding kinetics of OmpA-WT into each of these membranes was then measured (**Fig. 3.25e**). Folding into DMPC:OmpA-NN proceeded significantly slower

than into DMPC:OmpA-NP/NC, and included a lag-phase. Data are thus compared by calculating the T_{50} (time to half maximum amplitude). DMPC:OmpA-NN facilitates folding ~2x slower than DMPC:OmpA-NC, while membranes with OmpA-NP are ~2x faster than those with OmpA-NC. These trends recapitulate those seen with the lipid induced charge dipoles: Excess negative charge on the outer liposomal leaflet inhibits folding (a-DMPG/PC and DMPC:OmpA-NN) and excess negative charge on the inner liposomal leaflet enhances folding (a-DMPC/PG and DMPC-OmpA-NP) compared to controls (symmetric liposomes or DMPC:OmpA-NC). While this effect requires additional exploration, it is clear that charge asymmetry effects on OmpA-WT folding can be induced by both lipids and proteins – significant due to the high concentration of proteins in the OM.

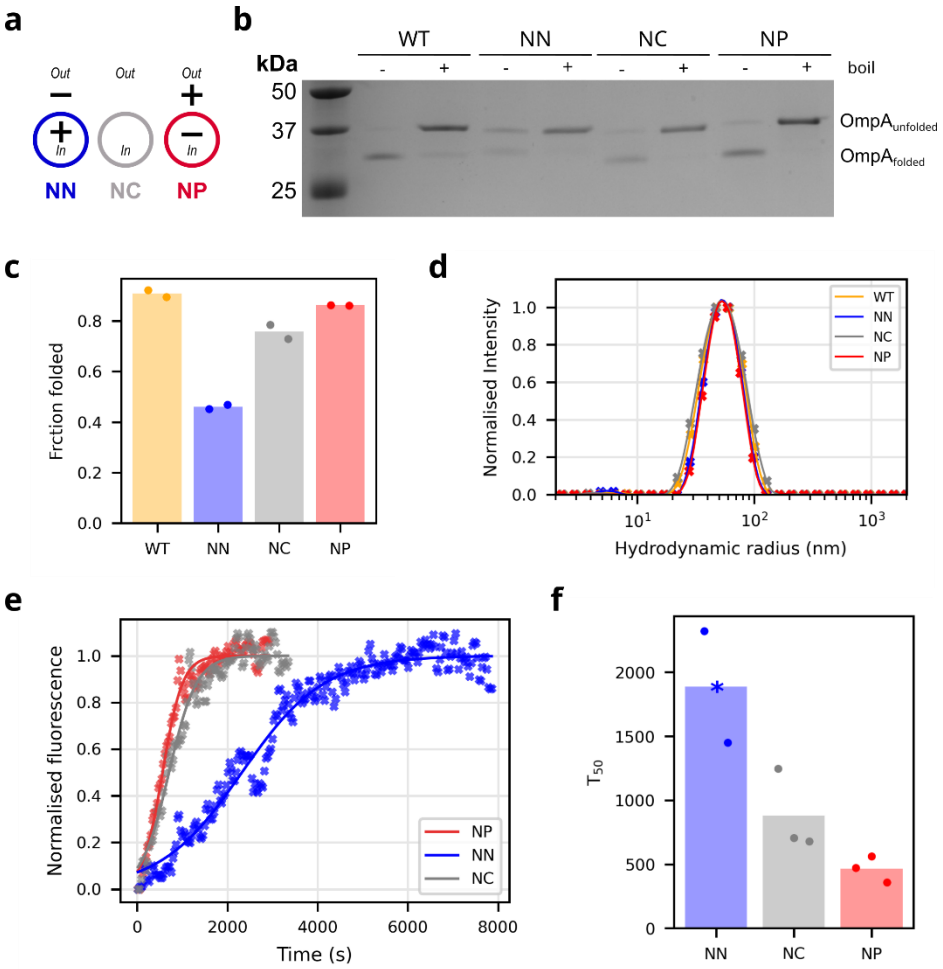


Figure 3.25: Protein-induced membrane charge dipoles modulate OmpA folding. (a) Membrane dipoles induced by the presence of OmpA extracellular loop mutants. **(b)** SDS-PAGE folded-unfolded band-shift of OmpA variants folding into DMPC liposomes at an LPR of 1:320. **(c)** Quantification (by gel densitometry) of fraction folded for each of the OMPs (n=2). **(d)** DLS of the cleaned proteoliposomes, prior to OmpA-WT folding. **(e)** Example kinetic traces of folding OmpA-WT into DMPC:OmpA-NN/NC/NP. **(f)** Comparing the folding rates into each proteoliposome by the T_{50} (time to half maximum amplitude) (n=3). (* indicates that sample did not complete folding in 2 hours).

3.3 Discussion

More than 35 studies on the folding kinetics of 15 different OMPs have been published over the last 30 years (reviewed in [255,262]). Despite this extensive literature, and the fact that lipid asymmetry between the two leaflets of a bilayer is the norm for biological membranes, very little is known about its implications for OMP folding or stability. Here we present a systematic study of OMP folding into asymmetric bilayers, and the results are striking. This work shows that lipid asymmetry has a profound effect on both the observed rates of folding and the apparent stability of the protein in the bilayer. This effect is mediated by charge distribution. Increasing the number of negatively-charged lipid head groups (from either DMPG or DMPS) in the inner leaflet of the liposome (functionally equivalent to the outer leaflet of the OM), progressively reduces the kinetic barrier for folding and thus increases its rate. However, when negatively charged lipids are only present in a liposome's outer leaflet (equivalent to the inner leaflet of the OM), stabilising lipid head group interactions with the natively folded protein loops in the inner leaflet cannot occur, with the result that both OmpA and BamA fold poorly. In addition to altered protein-lipid interactions, folding may be modulated by changing membrane mechanical properties[588]. In symmetric membranes charged lipids decrease global membrane stiffness due to electrostatic repulsion[584,594,595], offering a possible explanation why higher fractions of DMPG facilitated easier protein insertion, but the effects of asymmetric charge distribution are unclear[595,596]. While this merits further study, it has been shown here that the observed effects of asymmetry are consistent with both the shorter DM-acyl chains and the longer, mono-unsaturated PO-acyl chains, indicating that the effects of membrane mechanical properties are unlikely to provide a whole explanation. Regardless of the exact underlying physical phenomena, the interplay between protein and lipid charge distribution has major consequences for OMP folding and stability.

The results from liposomes composed of DMPE-PG are particularly interesting, as these are the dominant lipids in the inner leaflet of the bacterial OM[597]. The symmetric incorporation of DMPE into DMPG liposomes slows folding, consistent with symmetric DMPC-PE mixes[268]. However, we show this inhibition can be partially overcome by introducing DMPE asymmetrically into the liposome outer leaflet only. Thus, the reduction in folding rate is mediated by the type of lipids in *both* leaflets, with the balance of negative charge across the two leaflets forming a rheostat that

tunes the folding rate by more than three orders of magnitude in the conditions sampled here. In a biological context, our results suggest that the asymmetric presence of the negatively charged LPS in the bacterial OM could directly facilitate folding, with the rate being modulated by lipids in the inner leaflet of the OM, as well as the presence of BAM and other folding factors. Interestingly, the inner membrane (IM) of canonical rod-shaped *E. coli* is also asymmetric, with an ~three-fold excess of PE in the leaflet facing the cytoplasm compared to the periplasm-facing leaflet[71]. This excess of a zwitterionic lipid head group in the inner leaflet would create an effective excess of negative charge on the periplasmic leaflet, which would disfavour aberrant folding of OMPs into the IM and hence could play a role in determining the flux of OMPs to the OM. Thus, the native charge gradients across the cell envelope *in vivo* match that seen *in vitro* to allow efficient folding into the OM and limit folding into the

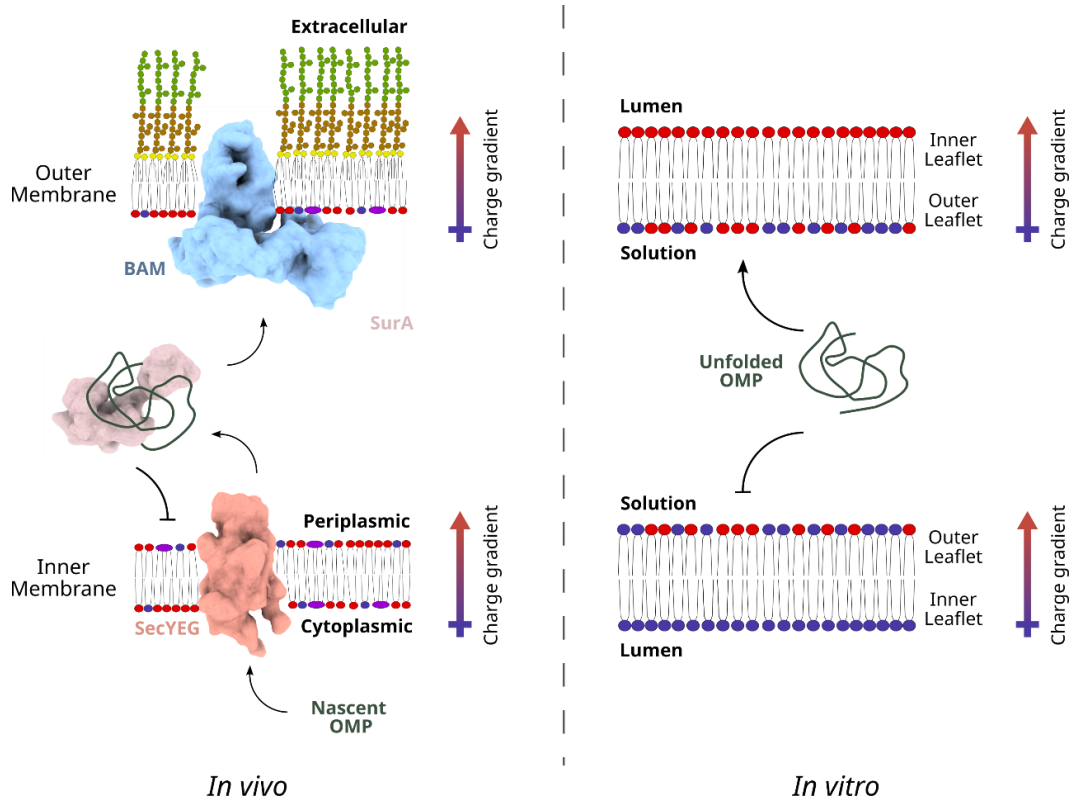


Figure 3.26: Comparison of membrane charge dipoles and folding pathways for nascent (*in vivo*) or unfolded (*in vitro*) OMPs. The charge gradients found to that support efficient folding *in vitro* is in the same direction as in the OM that OMPs are folded into natively, via the BAM complex. Membrane charge gradients found to retard OMP folding *in vitro* is thought to be in the same direction as the inner membrane, which natively does not incorporate OMPs. SEC-SurA-BAM is thought to be the dominant folding pathway i*n vivo*, but many other factors are involved that are not shown for clarity.

IM, helping to maximise flux through the folding pathway, and minimise off-pathway folding errors (**Fig 3.26**).

While the LPS in the OM is strongly negatively charged, inducing the global membrane charge dipole, OMPs and LPS tend to phase separate *in vivo* to generate OMP islands with only low concentrations of LPS – weakening the lipid induced charge asymmetry. However, while BamAs positive charges are organised into the PEP (**Fig. 3.24b**), it is an unusual OMP due to the large excess of negative charge in its extracellular loops (23 negatives, 10 positives), making its extracellular loops significantly electronegative (**Fig 3.27**), especially at the top of the lumen and the back of the external face of the barrel. Given the charge dipole preference for OMP intrinsic folding, that can be induced by either lipids or proteins, it is suggestive that BAM, responsible for natively folding OMPs, presents a predominantly negative external surface to the folding OMP – perhaps helping to facilitate and drive folding, especially if folding is occurring into protein islands depleted in negatively charged lipids. Indeed, this may provide a mechanism for BamA catalysis of OMP folding[225].

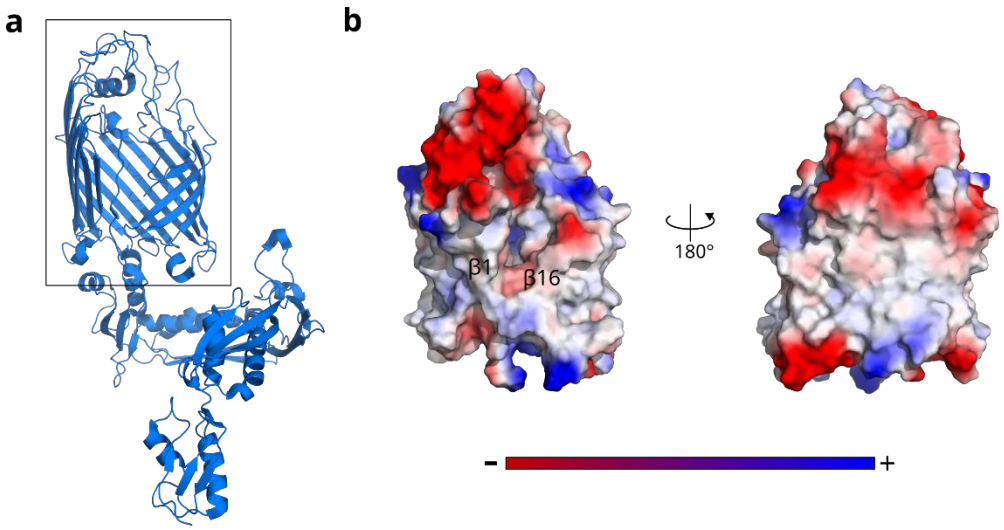


Figure 3.27: The extracellular loops of BamA are predominantly electronegative. (a) Cartoon model of lateral-open BamA (PDB: 5LJO[56]). **(b)** Electrostatic potential surface of the barrel region of BamA.

Protein charge interactions more generally are known to play a role in folding[598], and the ‘positive outside’ rule, first described in 2005[589], is a well-recognised feature of OMP sequence/structure. Here we begin to reveal the molecular detail that underpins this phenomenon. Using MD simulation and mutational analysis we identify that a patch of external positive (PEP) residues in the loops of OmpA are critical for productive folding, rather than a general requirement for positive charge. These PEP residues lie ~6-10 Å from the membrane surface and mediate OMP folding via interactions with the excess negative charge that we have identified above as being a key driver for efficient folding. Using bioinformatics, we show that the PEP is a generic feature of OMP sequences, suggesting that the PEP may be a conserved determinant of efficient OMP folding. For the studies in liposomes, this excess negative charge is the inner leaflet, but in the native OM the protein would approach the membrane from the periplasm, and excess negative charge particularly on LPS molecules, would be in the outer leaflet of the OM. Thus natural selection on OMP sequences and the machinery for the generation and maintenance of lipid asymmetry might plausibly operate synergistically to maximise the efficiency of OMP folding.

In summary, these results provide new insights into how bilayer charge asymmetry affects the folding and stability of OMPs. Specifically, we reveal charge-mediated features in both the lipid environment and protein sequences that reduce the kinetic barrier to OMP folding and stabilise the final, membrane inserted state. Although the exact nature of the modulation, and its interplay with other parameters that might modulate folding, such as membrane mechanical properties, will require further studies of a broad range of OMPs, including lipid mixes incorporating LPS, the results suggest routes to manipulate OMP behaviour for biotechnology applications, and how bacteria might exploit lipid asymmetry to modulate the efficiency of OMP folding into the highly asymmetric OM, and more broadly how cells might exploit lipid asymmetry to modulate the efficiency of folding of their membrane proteins.

Chapter 4: Lipopolysaccharide mediates functional and structural modulation of OMPs

Lipopolysaccharide (LPS) is a glycolipid found uniquely in the outer membrane of diderm bacteria, formed of 4-7 acyl chains covalently linked to an oligosaccharide consisting of a well conserved core-region of 6-7 mainchain sugars and a long O-antigen that extends into the extracellular milieux. The importance of LPS for OMP function has been elucidated in a few examples: notably the protease OmpT, which depends on LPS for optimal *in vivo* activity, and TonB-dependent Transporter BtuB, for which LPS is thought to order the loops to allow for function. Here the LPS dependence of OmpT activity is further dissected biochemically to discover a dual activation mode, including the requirement for specific regions of the LPS oligosaccharide, while activity of the phospholipase OmpLA is shown not to be modulated by LPS. Using molecular dynamics (MD) and via resolving a ~6 Å cryoEM structure, a model of LPS-mediated enhancement for OmpT activity is proposed. It is also shown that the OMP FusA strictly requires LPS to enable its structural solution by cryoEM, indicating LPS induced protein stabilisation, and a 2.4 Å structure, combined with MD, reveals FusA:LPS binding sites. Together, these data demonstrate the critical importance of LPS in maintaining OMP structure and function and thus a functional OM and highlights the matching between specific regions of LPS and proteins for optimal function.

Cloning, expression and purification of FusA was carried out by Khedidja Mosbahi (University of Glasgow). *E. coli* Keio mutants were kindly provided by Tracy Palmer (Newcastle University). OmpLA was expressed and purified by Dr. Bob Schiffrin.

4.1 Background

Interactions between membrane proteins and their lipid microenvironments are known to be critical in modulating protein structure[70], dynamics[504] and function[26,599], but much remains unclear about the implications of lipid organisation and diversity[26,30]. The outer membrane of LPS-diderm bacteria consists of a unique, highly asymmetric membrane, with phospholipids on the inner (periplasmic facing) leaflet and lipopolysaccharide (LPS) in the outer leaflet[31]. LPS is composed of 4-7 variably-long acyl chains, and a long, extracellular sugar chain consisting of a well-conserved 6-7 sugar chain close to the cell (the core sugars) and a highly diverse repeating unit extending away from the membrane (the O-antigen)[164,600]. LPS biosynthetic mutants produce LPS with various core oligosaccharide lengths (**Fig. 4.1a**, see **Fig. 1.9**)[145]. The composition and asymmetry of this bilayer is strictly controlled[601], and LPS composition itself is further modulated by the cell for optimal fitness, including during virulence[602], in response to temperature[170] or environmental stressors like antibiotics[603], and in control of OMV formation[604]. Under stress conditions, LPS is critical in maintaining the membrane barrier[145]. LPS is known to have unique binding fingerprints to a range of outer membrane proteins (OMPs)[509], nearly all of which adopt a β-barrel architecture, with transmembrane β-strands linked by shorter intracellular turns and longer extracellular loops[605]. LPS is important for the oligomerisation[489] and function of specific OMPs[504,606], although much remains unclear about the nature of LPS-OMP interactions and modulations, including how the presence of LPS helps enable antibiotic resistance, and OMP sequence/structural determinants and motifs of binding. While OMPs represent an attractive drug target due to their cell-surface accessibility and range of essential and virulence related functions[173,607,608], general lack of understanding of how LPS modulates their function, structure and accessibility hampers the design and targeting of drugs to these proteins.

The omptin family of proteases are a pathogenically important group of OMPs implicated in the virulence of a range of pathogens, including *Escherichia*[609], *Shigella*[359] and *Yersinia*[360], infecting humans[390], all major livestock species (e.g. [394–396]) and plants[610]. Omptin proteases are 10-stranded β-barrels with large, structured extracellular loops, the upper cleft of which house the active site (**Fig. 4.1b**)[363]. OmpT, the archetypal protein from *E. coli*, has been extensively studied (e.g. [105,355,611,612]) and a facile fluorescence-based activity assay developed where auto-quenching of a substrate is released by its cleavage (thus increasing fluorescence, **Fig. 4.1c**)[105].

Omptins have a well-characterised LPS activity dependence, but the mechanism underpinning this is uncertain, with most proposals suggesting a subtle conformational change to the active site geometry, in particular to the accessibility and presence of the active site water[369,612,613]. A single crystal structure of the *Yersinia* omptin Pla reportedly binds LPS, but only density for its acyl chain is visible, making confident assignment distinguishing LPS from co-crystallising detergent challenging[614]. The dependence of OmpT activity on LPS has been partially studied using shorter forms of LPS (Lipid A component to the Rc-oligosaccharide, **Fig. 4.1a**), showing that at least hexa-acylated Lipid A and a phosphorylated heptose in the inner core are required for activity[612] (typical *E. coli* LPS when grown in rich media has these modifications)[145]. Another enzyme of the OM, OmpLA (**Fig. 4.1d**, also known as PldA) is a phospholipase that has an absolute requirement for divalent cations, which mediate dimerization, to be functional[615], and has also been implicated in the virulence of a broad range of bacterial genera (notably, *Yersinia*, *Heliobacter*, *Neisseria*, *Pseudomonas*, *Escherichia*)[616]. In the OM, OmpLA is responsible for degrading phospholipids mislocalised to the outer leaflet[420]. While its LPS dependence is unknown, given its role in OM lipid homeostasis, LPS activity modulation is plausible.

CryoEM has transformed the study of membrane proteins[617] and OMPs are no exception, for example revealing key details about protein insertion into the membrane via the beta-barrel assembly machinery (BAM)[228,259] and the mechanisms of oligomeric transporters[57]. Despite these successes, the majority of OMPs remain challenging targets for cryoEM due to their small size, β-strand architecture and the rotational pseudo-symmetry of their barrel domains, thus also making structural characterisation of LPS-OMP interactions challenging. Recently solved structures of smaller, monomeric OMPs have been achieved by exploiting native or synthetic protein binding partners, including macrobodies[66], lipoproteins[56] and phage components[507]. Many cryoEM resolved OMP structures include TonB-dependent Transporters (TBDTs), including both dimeric[57,618] and monomeric[619] examples. TBDTs are a broad, essential class of active nutrient transporters that couple the inner membrane's proton motive force with active substrate transport across the OM, via the periplasm spanning TonB protein and inner membrane ExbBD motor complex[142]. Importantly, the TBDTs BtuB and FhuA are known to have structurally and functionally important interactions with LPS, and both have been experimentally demonstrated with EPR (BtuB[504,506]) or structurally resolved by crystallography and EM (FhuA[381,507]).

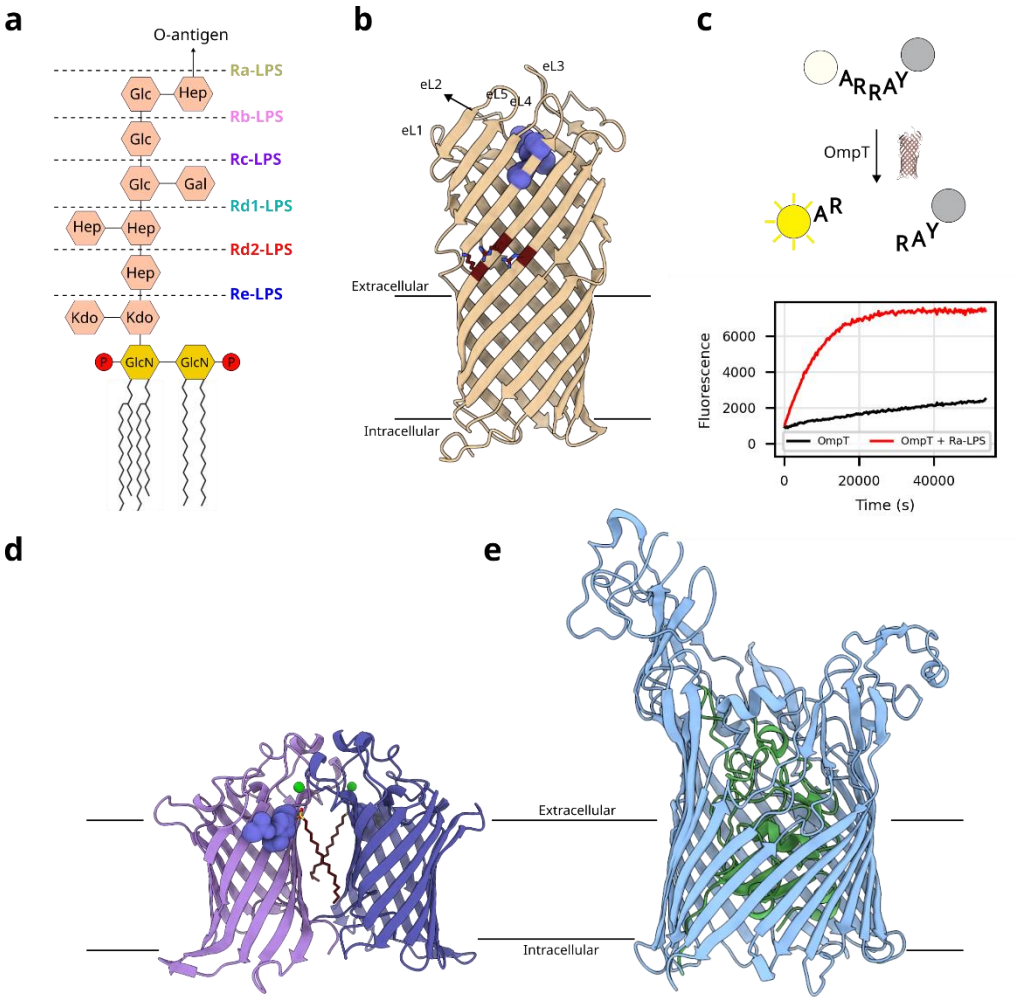


Figure 4.1: LPS chemotypes and OMPs. (a) Model of K12 LPS showing sugar identity and truncation names of the full core (GlcN: D-glucosamine, Kdo: 2-keto-3-deoxyoctulosonate, Hep: L-glycero-D-manno-heptose, Gal: galactose, Glc: glucose). **(b)** Structure of OmpT, indicating the active site at the top of the barrel cleft (blue space-fill) and the key residues in the proposed LPS binding site (red residues) (PDB: 1I78[363]). **(c)** OmpT activity can be measured by a fluorescence assay where an auto-quenching peptide is cleaved, releasing quenching, and increasing fluorescence. Typical data and LPS-dependence is shown in the lower plot. **(d)** OmpLA forms a divalent cation (green spheres) mediated functional dimer, that cleaves phospholipids in the OM (an acyl chain in the each of two active sites is shown in red) (PDB: 7EZZ[381]). **(e)** Structure of FusA including the plug (green) and barrel (blue) domains. The TonB box (residues 23-58) was not resolved in the crystal structure and is not shown (PDB: 4ZGV[351]).

A recently discovered family of TBDTs imports iron-containing host proteins to the periplasm, where they are proteolytically degraded, and their iron released for use by the bacterium[347]. These protease-associated import systems are wide-spread in diderm bacteria and include the *E. coli* transporter/protease YddB/PqqL which is known to be important for uropathogenic bacterial fitness[348,349], and has been further linked to both optimal growth in rich media and import of the antibiotic novobiocin[288]. The most well-studied member of this class is the plant pathogen *Pectobacterium spp.* transporter FusA (**Fig. 4.1e**), coupled with its protease FusC and specific TonB-like protein FusB. FusA imports the 12 kDa ferredoxin host protein, which contains an iron-sulpher cluster[350]. A crystal structure of FusA showed that the extracellular loops of FusA are particularly extensive, at least in part to help form a binding site for the large ferredoxin substrate[351]. These loops increase the size of FusA compared to many other OMPs (MW = 100 KDa), making it more amenable to cryoEM, as well as providing a large surface for possible LPS interactions. *Pectobacterium* are plant pathogens, and represent a significant threat to food security, in particular by causing soft rot in potato and crops such as cucumbers[620]. Bacteriocins, antibacterial proteins, are able to exploit the FusA import pathway by adopting a ferredoxin like homology that is then transported into the cell, representing possible pathways to develop highly specific antibiotics[621]. A better understanding of how this TBDT family behaves in its native membrane context, and in particular how it interacts with LPS, would thus offer clearer routes for both understanding virulence and drugging pathogenic bacteria.

In this chapter OmpT's activity dependence on LPS core-oligosaccharide length is explored biochemically via purification of different LPS truncations and activity characterisation in detergent and membrane environments. It is shown that OmpT has a two-stage activation by LPS, with a smaller rate enhancement mediated by LPS with short oligosaccharides (Re to Rc) and a larger rate enhancement with longer oligosaccharides (Rb, Ra and smooth LPS). Coupled with protein-lipid interaction data from molecular dynamics (MD) and an ~6 Å cryoEM structure of OmpT-LPS, an activation model is proposed. However, OmpLA is shown not to depend on LPS for activity. Further it is shown that the cryoEM structure determination of the OMP FusA was only possible upon the addition of LPS, suggesting LPS induced protein stabilisation, and the FusA:LPS binding sites are revealed in a 2.4 Å structure and by MD. Together, these data demonstrate the critical importance of LPS-OMP interactions to maintain a structural and functional OM, and in particular the importance of matching between sugar and protein location and identity.

4.2 Results

4.2.1 Characterising and purifying LPS

In order to study OMP-LPS interactions, techniques to purify and characterise LPS are required. UV-vis absorbance spectra indicate the presence of contaminating protein or nucleic acids (the absence of 280 nm and 260 nm peaks respectively), and the presence of LPS which absorbs ~195-205 nm (**Fig. 4.2a**). (Commercial LPS can be significantly contaminated with DNA, and if required was treated with DNAse before use, **Fig. 4.2b**). Sugar moieties in LPS can be quantitatively detected by the anthrone assay (c.f. Section 3.1.1) (**Fig. 4.2c**). This allows concentration determination owing to the complete conversion of reaction-competent sugars in LPS, causing the colorimetric change. Matching their expected reactivity (**Fig.** A9), Ra-LPS consisting of 4:1 and Re-LPS 1:1 glucose equivalents (gradients of fitted lines in **Fig. 4.2c**, 4.03x and 0.98x of glucose respectively). Silver-stain SDS-PAGE enables the direct visualisation of LPS (**Fig. 4.2d**) (and indicates any protein contamination). A quicker and non-toxic, although less sensitive, alternative for visualisation on SDS-PAGE is zinc-imidazole staining. The results from two different staining methods are shown in **Fig. 4.2e** (centre and right, see also Section 2.10). Following optimisation, clear staining of the protein control (amylase) and oligosaccharide laddering from smooth LPS was obtained (**Fig. 4.2e**, c.f. the silver stain equivalent).

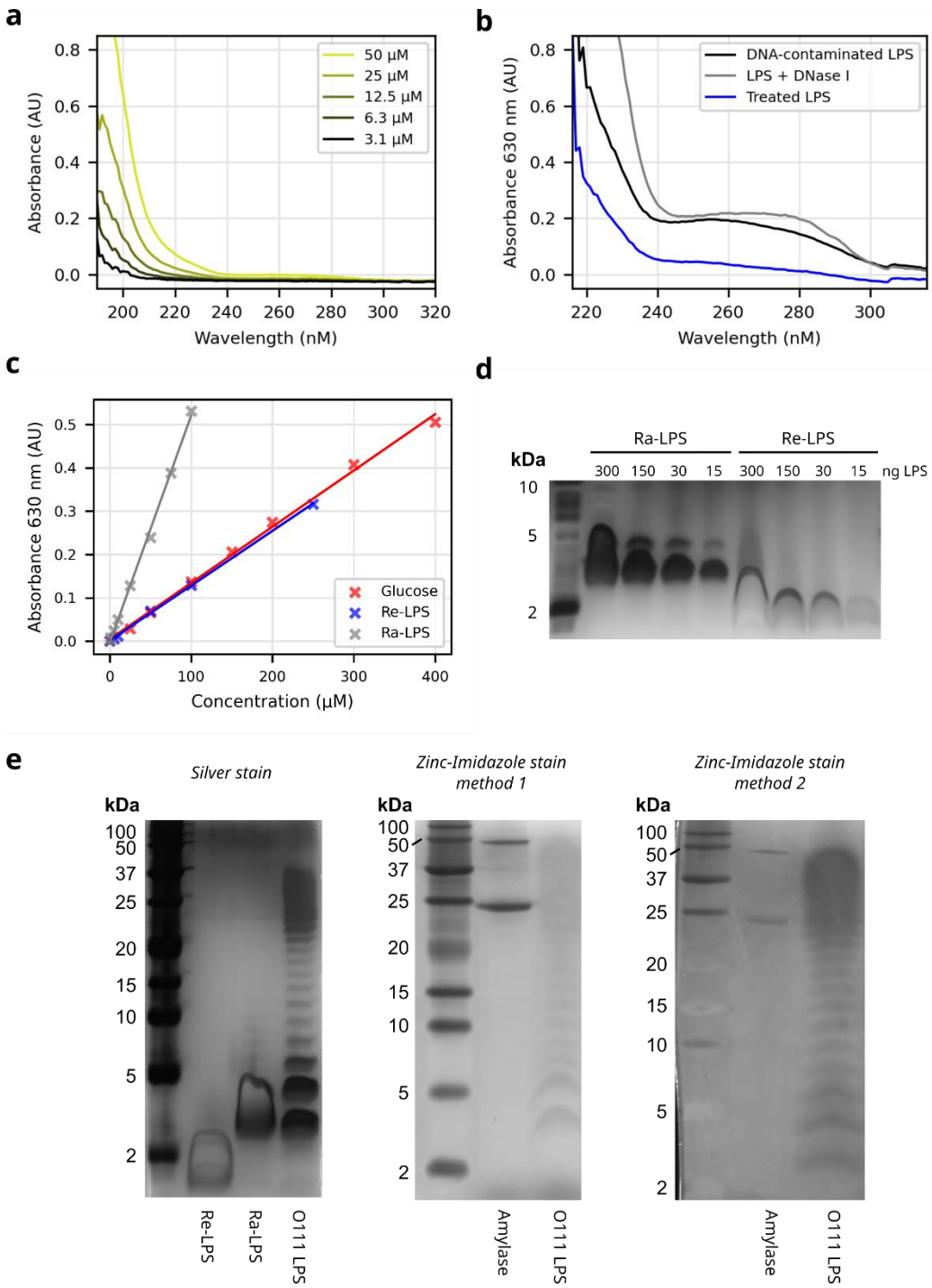


Figure 4.2: LPS characterisation. (a) UV-vis absorbance of Ra-LPS at 3.1-50 µM. **(b)** Treating commercial LPS with DNA contaminants (black) with 10 µg/ml DNase I (gray) leading to clean LPS following DNase I removal (blue). **(c)** Sugar detection calibration curves of glucose control (black), Re-LPS (red) and Ra-LPS (yellow) using the anthrone assay. Re-LPS has a 1:1 equivalence with glucose, Ra-LPS 4:1. **(d)** Ra-/Re-LPS visualisation by silver stain SDS-PAGE. **(e)** LPS visualisation by optimised zinc-imidazole staining (centre, right) compared to silver staining (left) (see Section 2.10).

LPS can also be purified from cells. Here *E. coli* K12 mutants from the Kieo collection are used, advantageous as different purified LPS have an identical cellular background (thus minimising LPS heterogeneity). Exploiting LPS biosynthesis mutants (Δ*WaaF*, Δ*WaaC*, Δ*WaaR*, Δ*WaaI*), different core-length truncations of LPS (Re, Rd2, Rc, Rb) were purified (**Fig. 4.1a**). Briefly, cells in salt solution were butanol extracted three times, keeping the aqueous phase, followed by incubations with proteinase K and RNAse/DNAse to remove contaminating protein and nucleic acid (see also Section 2.10). LPS purity was assessed by silver stain SDS-PAGE gel (**Fig. 4.3b**) and UV-absorbance (**Fig 4.3b**). LPS concentrations were determined based on sugar quantification (**Fig. 4.2c, A9**). Thus, LPS variants have been purified directly from cells.

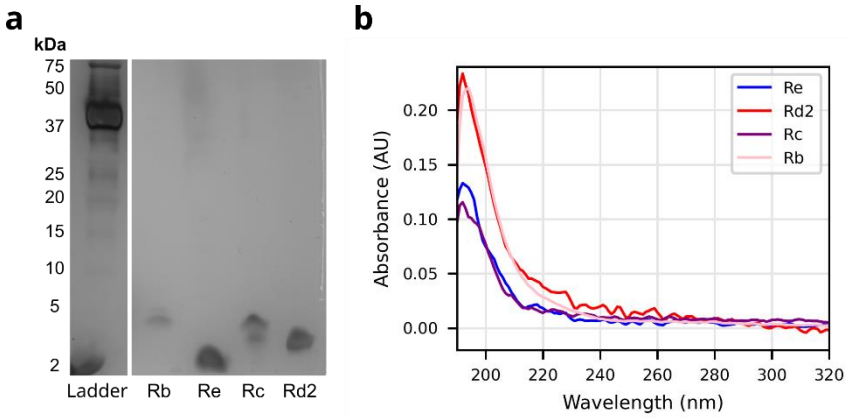


Figure 4.3: LPS core-truncation purification. (a) Silver stain SDS-PAGE gel and **(b)** UV-vis absorbance of purified LPS variants, indicating the presence of LPS and the absence of significant protein/nucleotide contaminants.

4.2.2 The dependence of OmpT activity on LPS

The dependence of OmpT catalytic activity on LPS of different oligosaccharide truncations was next characterised. Here the activity of inclusion-body expressed and purified OmpT refolded into detergent (LDAO, DDM) or reconstituted into liposomes (DMPS, DMPG, DMPC, DLPC, ECPL (*E. coli* polar lipid)) without LPS or with Re-LPS/Ra-LPS is assessed, **Fig. 4.4a** (via the fluorescence assay shown in **Fig. 4.1c**) (see methods in Section 2.7.2-3). OmpT showed a non-LPS dependent basal activity when refolded into DDM but not LDAO micelles (the *Yersinia* homolog Pla is also more basally active in DDM than LDAO[614]). Similarly, in DLPC and DMPC liposomes 25 nM OmpT basal activity was undetectable, while in DMPS, DMPG and ECPL liposomes there was distinct, although low, activity. This indicates that basal OmpT activity in liposomes requires negatively charged lipid headgroups, while in detergent the bulky, hydrophilic sugars of DDM facilitate enzyme activity better than those of the small, zwitterionic LDAO. LPS is characterised by both negative charge and steric bulk at the membrane interface, suggesting that both of these properties may be important for LPS-mediated OmpT activation. However, the basal activities remain substantially below those with Re-LPS and were further enhanced with Ra-LPS, highlighting the specificity of the interaction. All following work in this chapter uses DDM-OmpT or OmpT in ECPL liposomes, unless otherwise stated.

To further characterise LPS-detergent interactions, Re-LPS and Ra-LPS were titrated with DDM-OmpT, and the LPS concentration dependence of OmpT activity determined (**Fig. 4.4b**). Re-LPS more readily incorporates into micelles, with rate-enhancement saturation occurring at lower LPS:protein ratios (assuming equivalent amounts of Re-/Ra-LPS are required). This was expected as transfer of the smaller Re-LPS to DDM from LPS micelles should be more energetically favourable.

OmpT's activity was also assayed against the unfolded OMP BtuB as a substrate (in 2M urea, DDM-OmpT remains folded, **Fig. 4.4c** but BtuB is unable to refold) and a natural substrate, the cationic peptide LL37, by monitoring loss of intact protein by SDS-PAGE (**Fig. 4.4d-e**). In all cases, OmpT cleaves these substrates only in the presence of LPS, with Ra-LPS being more effective than Re-LPS. LL37 could not be used with proteoliposomes as it disrupts membranes[622]. Thus, while LPS is required for large activity enhancements, LPS-like features also facilitate basal activity.

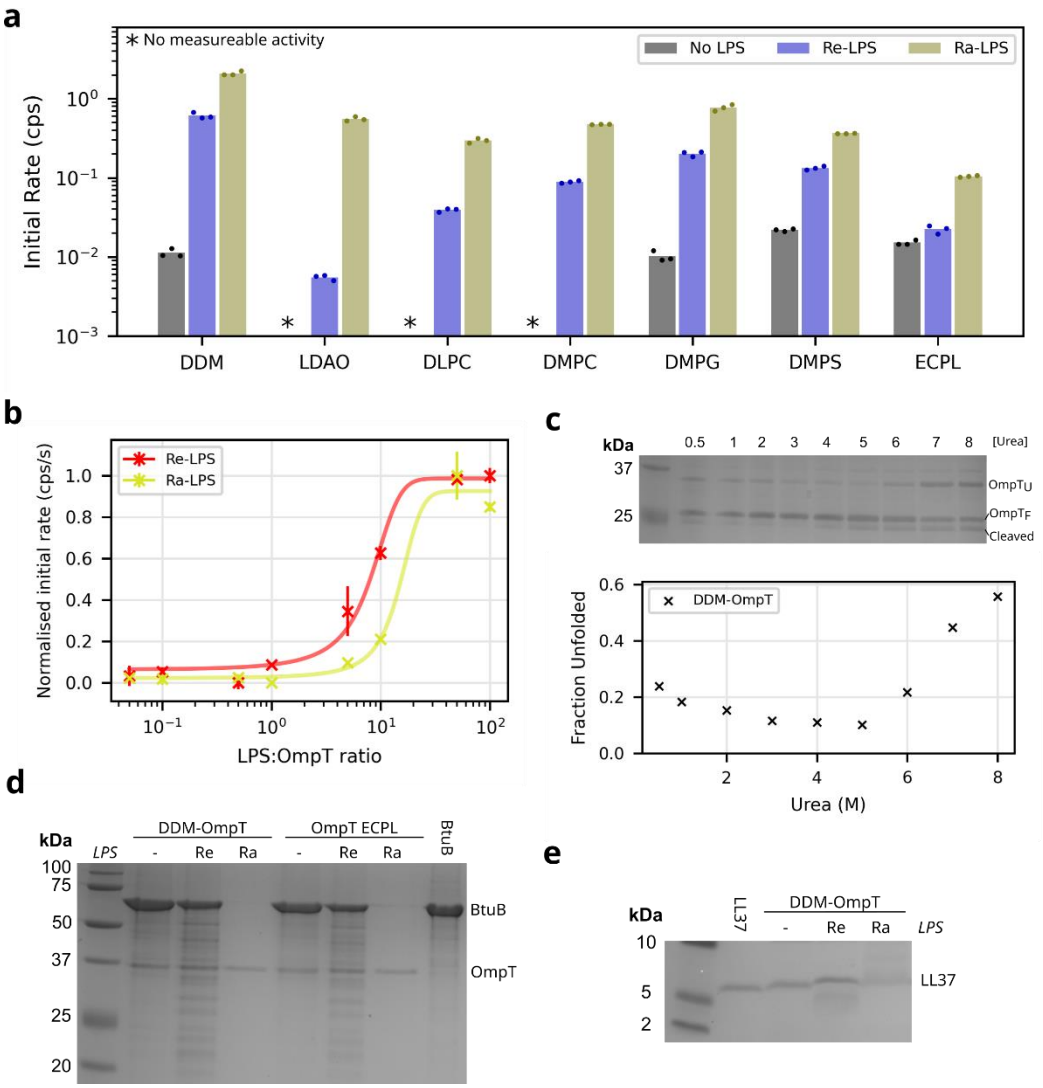


Figure 4.4: OmpT dependence on LPS for activity. (a) OmpT activity (25 nM) in various detergent and liposome contexts without LPS or with Re-LPS or Ra-LPS (n=3, technical replicates). **(b)** Re-LPS/Ra-LPS concentration dependence on activity of DDM-OmpT (25 nM) (n=3, error bars are data range). Fitted lines are to guide the eye only. **(c)** DDM-OmpT remains partially folded even in high urea concentrations, determined by folded/unfolded band-shift quantification by SDS-PAGE (at higher urea concentrations OmpT will cleave itself, indicated by the appearance of the cleaved band). Activity of OmpT against **(d)** unfolded BtuB (n=2) and **(e)** LL37 (n=4) without LPS or with Re-LPS or Ra-LPS. Loss of intact band with Ra-LPS in each case indicates complete degradation.

In addition to the changes in initial rates between OmpT and OmpT:Ra-LPS, the activity curve amplitudes are also different, with apo-OmpT activity plateauing at ~50% of OmpT:Ra-LPS (which consumes all substrate) (**Fig. 4.5a**, shown for DDM-OmpT). While adding fresh DDM-OmpT mid-plateaux does not lead to additional substrate processing (**Fig. 4.5b**, grey line), indicating that it is not due to OmpT degradation or denaturation, adding Ra-LPS leads to a rapid recovery of the enzyme's

activity until the reaction is complete (**Fig. 4.5b**, black line). Together, these data suggest that without Ra-LPS, OmpT is likely inhibited by its product, and that this inhibition is completely overcome by the binding of Ra-LPS (and nearly entirely overcome by Re-LPS (**Fig. 4.5a**, red)), presumably via conformational changes which alter the relative affinity of OmpT for its substrate or product. To confirm the product inhibition, a mixture of the cleaved products was purified. The mixture was then added to DDM-OmpT mid-reaction and observed to prematurely inhibit the activity (**Fig. 4.5c**). Adding the product mixture at the start of the reaction demonstrated the while DDM-OmpT was significantly inhibited, the OmpT:Ra-LPS was not (**Fig. 4.5d**). (Substrate inhibition is unlikely as adding LPS consistently enhances the activity). Therefore, LPS can increase processivity of the reaction at least in part by relieving product inhibition of apo-OmpT.

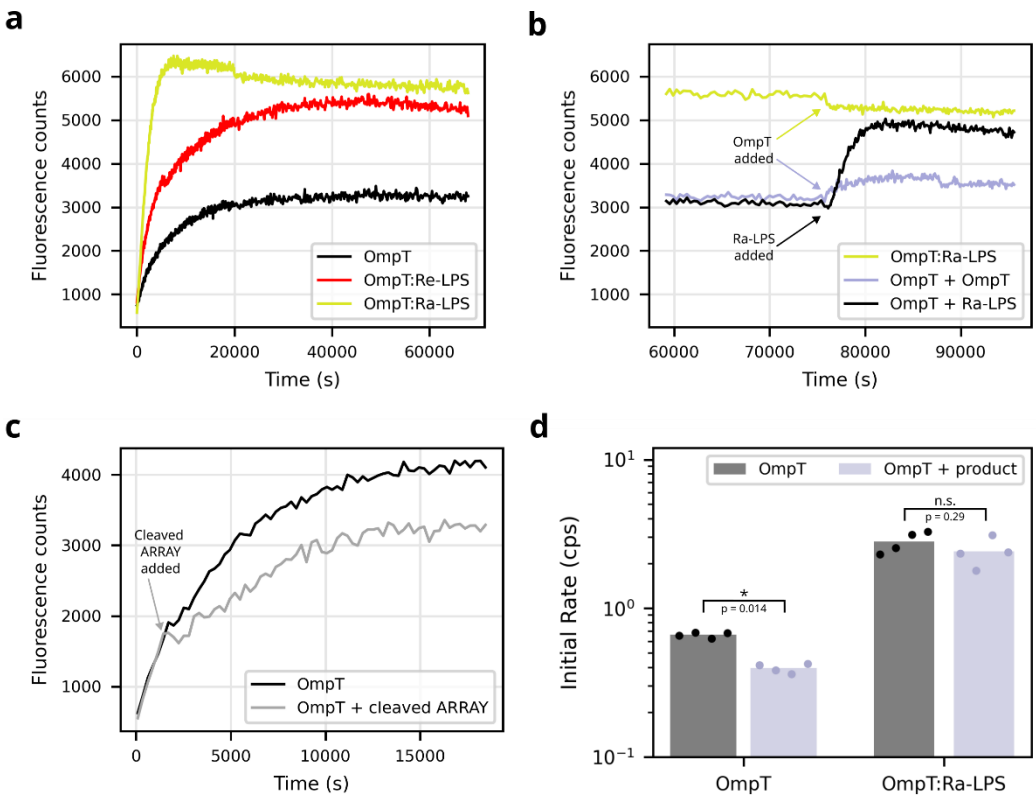


Figure 4.5: Product inhibition of apo-OmpT is overcome in the presence of LPS. (a) OmpT and OmpT:LPS fluorescence activity traces plateaux at different amplitudes for the same substrate concentration (50 µM). **(b)** Adding fresh OmpT facilitates minimal additional substrate processing, but adding Ra-LPS brings the reaction to completion (n=2). OmpT:Ra-LPS (Ra-LPS added at start of reaction, yellow) is shown for comparison. **(c)** Adding a cleaved product mixture mid-reaction prematurely inhibits apo-OmpT. (The fluorescent background from the cleaved product has been subtracted for clarity, n=3). **(d)** Starting reactions in the presence of a cleaved product mixture inhibits apo-OmpT but not OmpT:Ra-LPS (n=4). (In all parts OmpT at 50 nM).

4.2.3 OmpT activation by LPS of different core lengths

Ra-LPS consistently enhances OmpT activity more than Re-LPS (**Fig. 4.4a**), demonstrating the importance of the additional sugars present in the LPS core. Following incubation with the LPS purified in Section 4.2.1 and commercially purchased Ra-LPS and smooth LPS (O55:B5), DDM-OmpT activity was measured (**Fig. 4.6a**). Comparing the initial rates demonstrates a two-stage rate-enhancement (**Fig. 4.6b**). The presence of any LPS, regardless of oligosaccharide, accelerates activity compared to the basal rate by >2x (different conditions mean these experiments are not directly comparable to **Fig. 4.4a**). A larger difference is identified between Rc-LPS and Rb-LPS, with an additional >3x rate enhancement seen for LPS with oligosaccharides longer than Rc-LPS. The plateauing amplitude is also consistent, with the reaction fully completing with Rb-oligosaccharides or longer.

This effect was also assayed for in ECPL proteoliposomes doped with 0.5% LPS (~0.04% OmpT, mol/mol) (**Fig. 4.7a**). The initial activity rates showed a similar but exaggerated trend compared to the DDM-OmpT (**Fig 4.7b** c.f. **Fig. 4.7b**), with a two-phase rate enhancement with the larger step at the Rb-sugar level. It is intriguing that Rb-LPS enhances the rate significantly more than Ra/O55-LPS, perhaps a protein lipid interaction at this sugar is inhibited by further extending the oligosaccharide. Determining the Michaelis-Menten kinetics allows the decomposition of the rates into V_{max} and K_m (**Fig. 4.7c-d**). V_{max} shows broadly the same trends as the initial rates (**Fig. 4.7c** c.f. **Fig. 4.7b**), with LPS always mediating a V_{max} enhancement and the enhancement being greater for longer oligosaccharides (Rb/Ra/O55-LPS).

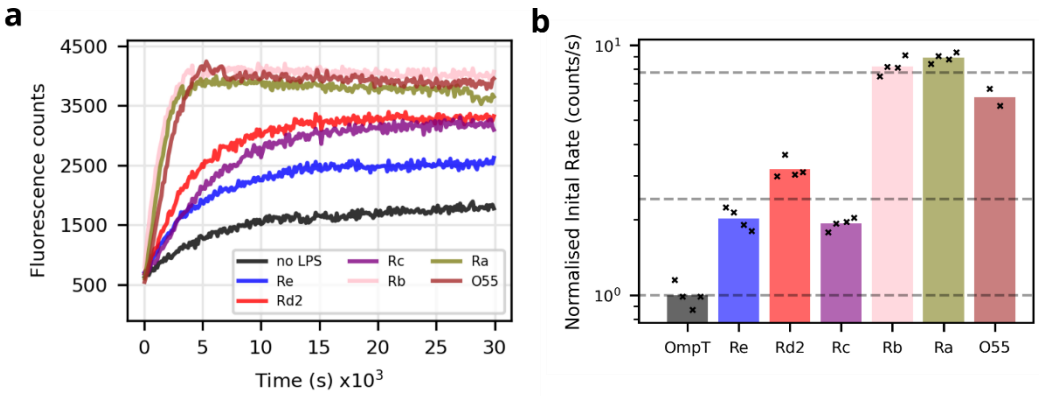


Figure 4.6: DDM-OmpT with different LPS truncations show a two-stage rate enhancement. (a) Example raw activity traces and **(b)** normalised initial rates of OmpT with different LPS oligosaccharide lengths (n=4, 2x LPS preparations). Rates normalised to the average of apo-OmpT. Dashed lines in (b) are the average of apo-OmpT, Re/Rd2/Rc-LPS or Rb/Ra/O55-LPS (bottom to top). (0.01 µM OmpT and 50 µM ARRAY).

The changes in K_m are less consistent and much smaller, with the majority remaining around the value for apo OmpT (~100 μM). Rb/Rd2-LPS have higher K_ms (370 μM and 220 μM, respectively), but in both cases these are offset by a higher V_{max}. In contrast O55 has a reduced K_m (50 μM) as well as an elevated V_{max}, indicative of at least a minor effect by the O-antigen sugars, intriguing as in most wild LPS-diderm bacteria smooth LPS dominates the OM's outer leaflet lipid composition.

While the details of this data merit further exploration, in particular the different trends in Michaelis-Menten kinetics of O55-LPS and the very large V_{max} enhancement mediated by Rb-LPS, it is clear that the upper sugars of the LPS core are important for maximal OmpT activity. The top of the core sugars in LPS is at approximately the same level as the upper region of OmpT's loops and near the active site of the enzyme (**Fig. 4.1b**, c.f. **Fig. 1.17b**, **Fig. 4.11f**), suggesting that the LPS may be inducing specific conformational changes in these regions leading to an increase in enzyme activity.

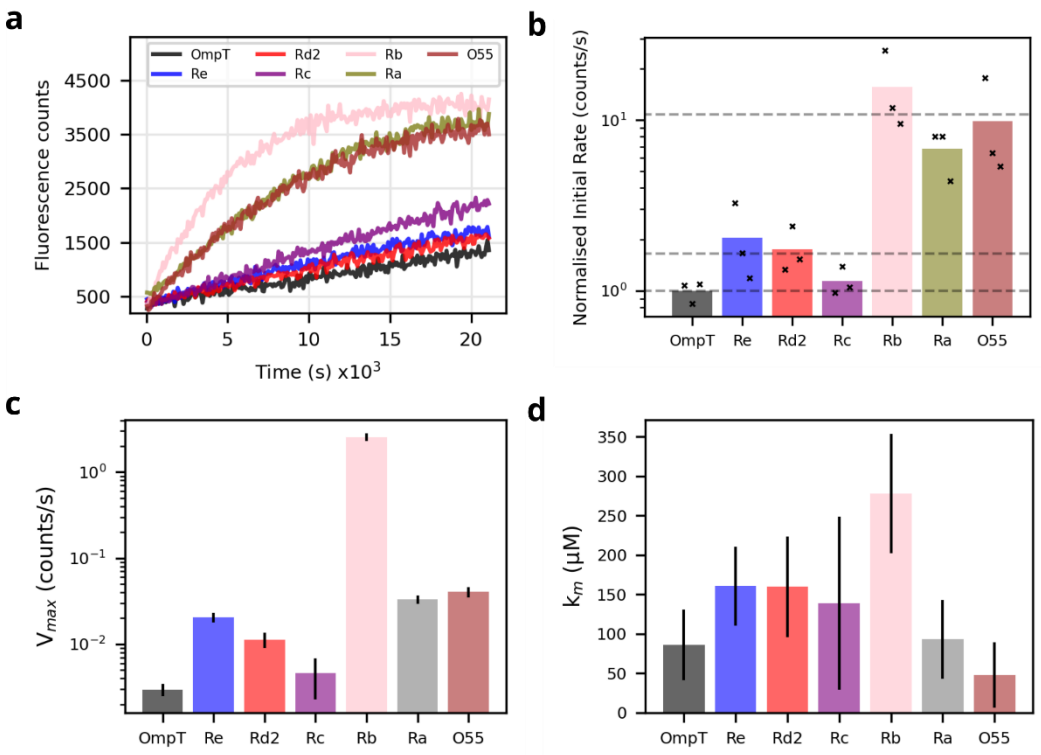


Figure 4.7: OmpT in ECPL proteoliposomes with different LPS truncations also shows a two-stage rate enhancement. (a) Example raw activity traces and **(b)** normalised initial rates of OmpT with different LPS oligosaccharide lengths (n=4, 2x2 LPS and proteoliposome preparations). In **(a)** and **(b)** OmpT at 0.005 μM and ARRAY at 50 μM. Rates normalised to the average of apo-OmpT. Dashed lines in (b) are the average of apo-OmpT, Re/Rd2/Rc-LPS or Rb/Ra/O55-LPS (bottom to top). **(c)** V_{max} and **(d)** K_m for OmpT in proteoliposomes with each LPS. Errors the 95% confidence limits estimated by bootstrapping over 20 data-points. (OmpT at 0.005 μM)

4.2.4 Molecular dynamics reveals specific OmpT-LPS interactions

The OmpT:LPS interaction was next explored with MD. LPS with polysaccharides of different lengths have been previously parameterised for the Martini coarse-grained forcefield, including Re-LPS and Ra-LPS (**Fig. 4.8a**). Lipid only simulations with 1:3 Re/Ra-LPS:DMPE in the outer leaflet and an 80:15:5 DMPE:DMPG:cardiolipin inner leaflet (matching native inner leaflet headgroup composition) were set-up using a customised script to ensure limited initial LPS clustering (**Fig. 4.8b** upper, see Section 2.6.5). These demonstrated that LPS behaved as expected with both slow diffusion (**Fig. 4.8c**) and clustering (**Fig. 4.8b** lower**, d**) in the membrane, with the larger Ra-LPS being more prone to both these features than Re-LPS, confirming both known *in vivo* and *in silico* behaviour.

Due to the slow diffusion of LPS, to allow for sufficient sampling in a computationally reasonable simulation timeframe (<10 μs), simulations with OmpT were set-up with only 8 LPS moieties (~1.5 %) initialised singularly in a distant ring around the central protein (**Fig. 4.9a**). This setup allowed much faster initial diffusion of the LPS in the membrane (**Fig. 4.9b**), which slowed after interaction with OmpT or other LPS molecules. Analysing the contacts between Ra-LPS and each residue of OmpT revealed preferential lipid binding regions (**Fig. 4.9c**), which can be seen clearly by determining lipid density over the simulation time (**Fig. 4.9d**), and by mapping the interaction probability onto the OmpT structure (**Fig. 4.9e**). Splitting the lipid-protein contacts into those formed by each sugar group (CG-bead mapping in **Fig. A10**) reveals that an excess number of contacts is formed between LipidA/Re-sugars and Rb-/Ra-sugars and OmpT compared to the other groups (**Fig. 4.9f**). A phosphorylation in the CG-representation mediates the higher interaction count of the Rd1 sugars. Normalising the contacts to account for the different number of beads in each sugar grouping demonstrates that the upper sugars of the oligosaccharide (Ra and Rb) interact significantly more than the lower sugars (**Fig. 4.9g**). While not conclusive, the strong similarities between the MD-interaction profile (**Fig. 4.9g**) and the OmpT activity enhancement profile (**Fig. 4.6b**, **4.7b**) is suggestive of a functionally important interaction occurring at the Ra/Rb-LPS sugar region.

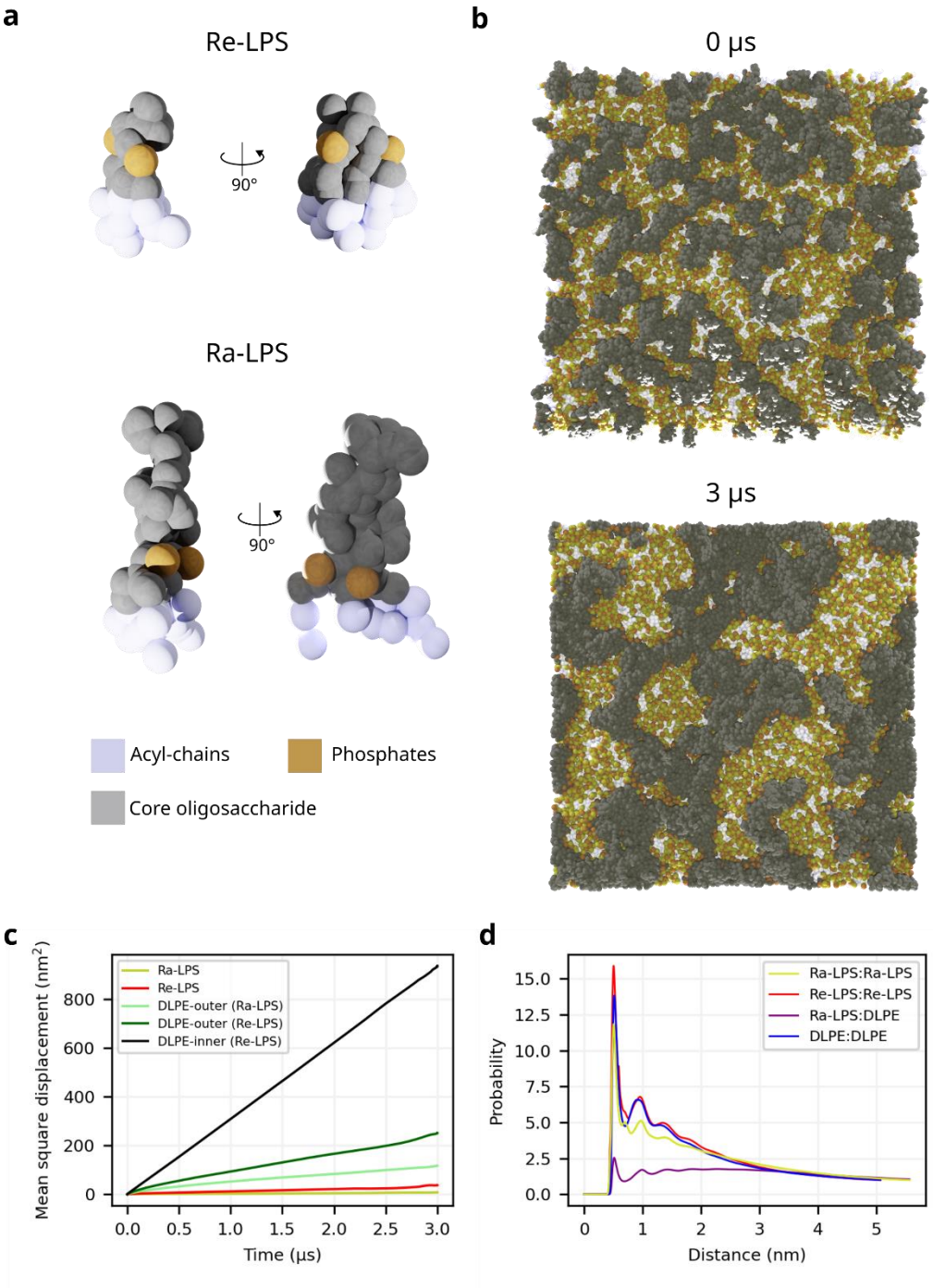


Figure 4.8: CG-MD Re-/Ra-LPS characterisation. (a) CG-bead structure of Re-LPS (upper) and Ra-LPS (lower), highlighting acyl-chains (purple), LipidA phosphates (brown) and core-oligosaccharide beads (grey). **(b)** Start (upper) and end (lower) frame of 3 µs 25% (mol/mol) Ra-LPS simulation clearly shows Ra-LPS cluster formation (Ra-LPS sugars: grey). **(c)** Mean squared diffusion distances of Re-/Ra-LPS and phospholipids in different simulations and leaflets. DLPE is shown for the outer leaflet of both Ra-LPS and Re-LPS containing simulations (light/dark green), and for the inner leaflet of Re-LPS containing simulations (black). **(d)** Radial probability distribution function indicating probability of finding a pair of lipids a given distance apart.

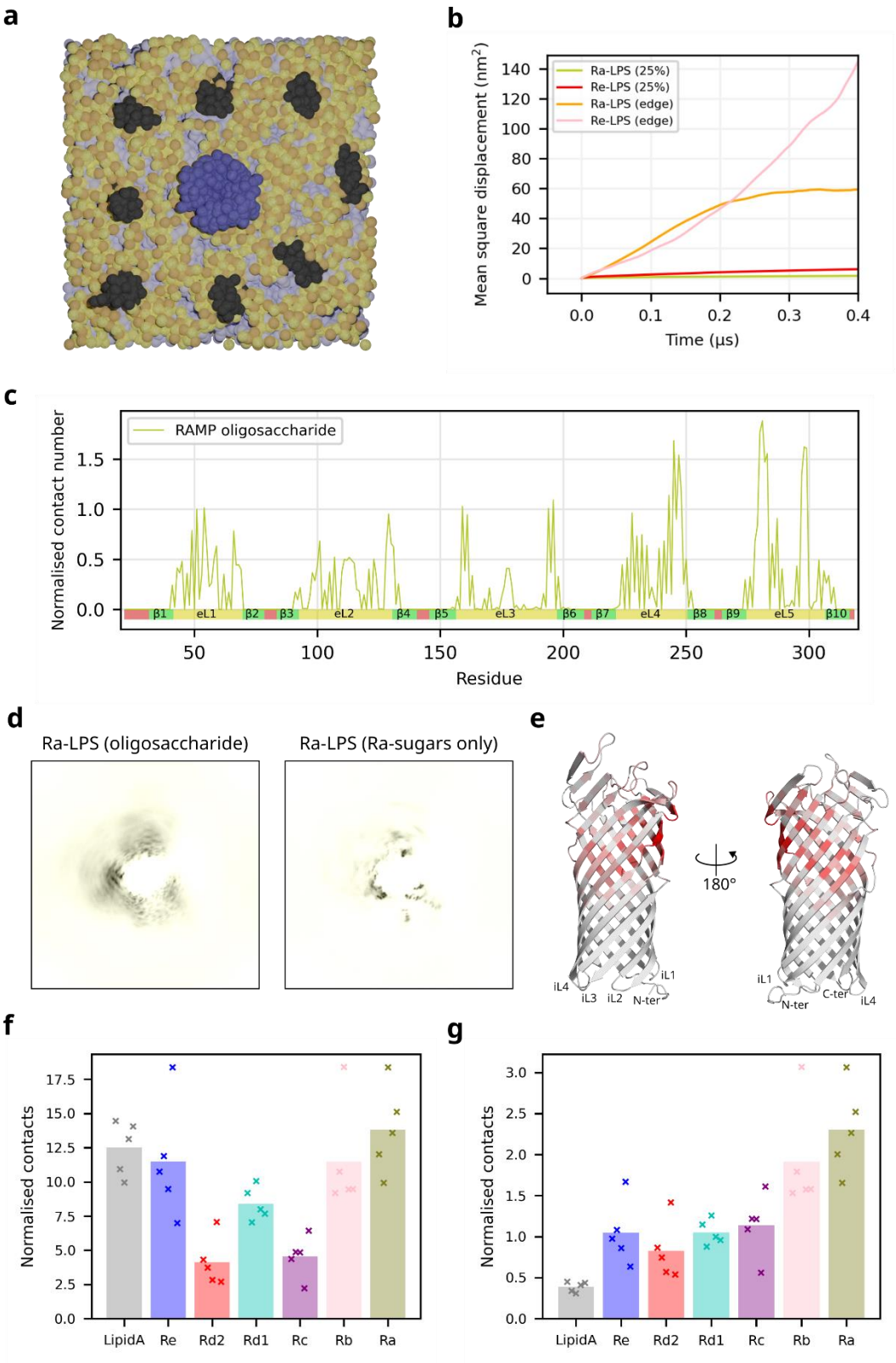


Figure 4.9: (Legend overleaf)

Figure 4.9: Different sugars of Ra-LPS specifically interact with OmpT. (a) Equilibrated pre-production frame of sparse (~1.5 %) Ra-LPS simulations (blue: OmpT, dark gray: LPS, orange: phosphates, yellow: lipid headgroups, purple: acyl chains). **(b)** Mean squared diffusion distances of LPS in the sparse LPS simulations. **(c)** Contact number between the Ra-LPS oligosaccharide and each residue of OmpT, indicating a preference for eL 4-5. **(d)** Lipid density (averaged over simulation time) around OmpT for the Ra-LPS oligosaccharide (left) and the Ra-sugars only). Trajectories centred and fitted to OmpT prior to analysis. **(e)** Probability of each residue of OmpT interacting with Ra-LPS mapped onto the crystal structure (PDB: 1I78[363]). **(f)** Normalised contact count between OmpT and different parts the Ra-LPS moiety, and **(g)** OmpT-Ra-LPS contacts normalised to the number of beads in each group.

4.2.5 CryoEM of OmpT-LPS

CryoEM was then used to attempt to solve the structure of OmpT:Ra-LPS. OmpT is an exceptionally difficult cryoEM target due to its small size (33.5 KDa), its mostly featureless, pseudo-symmetric cylindrical shape, and its β-sheet only composition. Further, disordered signal from the solubilising detergent conflates with the OmpT signal, making it challenging to separate and align to target signal. It is worth noting here that even small LDAO micelles would contribute at least as much signal to each particle image as the protein. Initial attempts to solve OmpT in MSP1D1 nanodiscs (verified by SDS-PAGE gel, DLS (6 nm peak) and OmpT activity, **Fig. 4.10a-b**) or solubilised in DDM revealed clear, although noisy, protein density in 2D classification (**Fig. 4.10c-d**) but reconstructions remained at low resolution (**Fig. 4.10e-f**). The extracellular loops of DDM-OmpT resolved to sufficient resolution (~11 Å) to orientate the OmpT crystal structure into the density (**Fig. 4.10f**, right) with reasonable confidence, but failed to further improve.

To minimise non-target signal (i.e. from nanodisc or micelle), OmpT in LDAO micelles, which are ~50 % smaller than DDM micelles (compare micelle sizes in **Fig. 4.10d** and **Fig. 4.11b**), was imaged. Following multiple rounds of divisive 2D classification, crYOLO/topaz picking model training on specific views of the protein and classification via multi-model initial model generation, good particle stacks were obtained (initial 2D classes **Fig. 4.11a** c.f. final 2D classes **Fig. 4.11b**, image processing detailed in Section 2.7.8). As expected, the best views obtained were side-on, with clear asymmetry in the barrel resolved. Reasonable top-views were also obtained (i.e. down the barrel axis), although they remained mostly featureless. Although tilt-views were seen, they were at a lower resolution, presumably due to the

alignment challenges presented by these viewing angles.

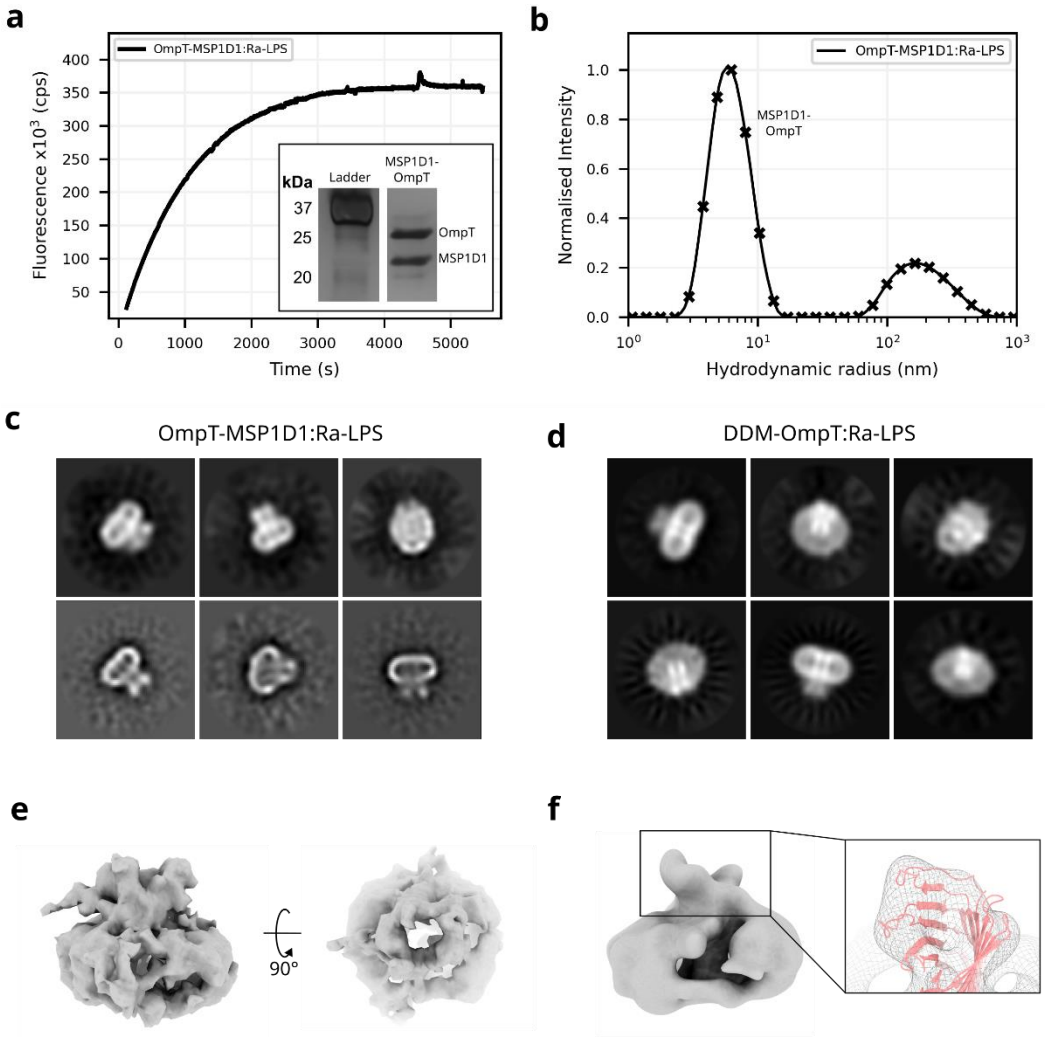


Figure 4.10: DDM-OmpT and nanodisc-OmpT cryoEM yielded very low-resolution reconstructions. (a) OmpT activity when reconstituted into MSP1D1 nanodiscs with ECPL and Ra-LPS. (Inset: silver-stain SDS-PAGE gel showing OmpT the presence of MSP1D1 and OmpT). **(b)** Nanodisc size by DLS indicates the majority of the sample is in monomeric discs (6 nm peak), although some aggregate has formed (200 nm peak). Typical 2D classes obtained from **(c)** OmpT-MSP1D1:Ra-LPS and **(d)** DDM-OmpT:Ra-LPS, revealing clear but generally featureless protein density (boxsize: 18.2 nm). Reconstructions obtained from **(e)** OmpT-MSP1D1:Ra-LPS and **(f)** DDM-OmpT:Ra-LPS, both of which show the expected shape for OmpT, but remain at > 10 Å resolution. Even at this resolution the asymmetry in the loops of OmpT are partially resolved in DDM **(f)**. (Model PDB: 1I78[363]).

Following extensive optimisation, an ~6 Å reconstruction was resolved (**Fig. 4.11c**), with protein density almost entirely emerging from the micelle. It is the expected shape and size of OmpT, given previous crystal structures (**Fig. 4.11d**). Across a wide-range of reconstructions via both different particle stacks and processing parameters, additional density was consistently observed rising out of the micelle and contacting the OmpT-assigned density near the top of the protein (**Fig. 4.11c** red circles, **Fig. 4.11e**), suggestive of the expected shape and form of an LPS molecule, although the resolution is not high enough to certainly assign it. While the model is not high enough resolution to confidently assign either map handedness or the rotation of the model via its pseudo C2 symmetry axis (i.e. across the long axis of the barrel oval) – resulting in four possible positions to rigidly fit the crystal structure into the density – under the assumption that this is LPS, combined with the MD LPS binding sites, this allows a single orientation of the model to be determined (**Fig. 4.11d**). LPS binding near this position has previously been hypothesised[363] (**Fig. 4.1b**).

While the low-resolution limits map interpretation, it is suggestive that the protein-LPS interactions occur at two broad interaction points: lipid-A and the lowest oligosaccharide sugars and a second interaction point in the upper regions of the Ra-LPS, with a visible gap between the protein and proposed Ra-LPS in the central oligosaccharide region (**Fig. 4.11e**). Aligning an extended Ra-LPS molecule along the axis of the proposed LPS density (**Fig. 4.11f**) suggests that the upper interaction is the Rb/Ra sugars and the lower interaction is the lipid-A and Re sugars. Coupling this with previously presented biochemical data, which demonstrated a two-phase rate enhancement by LPS in agreement with these LPS-truncations, supports this assignment, as does the MD data which shows weaker LPS-OmpT interactions for the central sugars of the oligosaccharide compared to the other lipid regions.

Taken together, the biochemical, simulation and EM analysis allows a proposed model for OmpT activation by LPS, where the negative charges and/or steric bulk of Lipid-A and the lower Kdo sugars facilitate a basal level of activity while the upper core-sugars of LPS interact with OmpT's extracellular loops and mediate a much larger rate enhancement via an undetermined conformational change. At least part of this dependence is mediated via a relief from product inhibition by LPS binding. OmpT activity will be further dissected in relation to its protein context in the OM in section 5.2, but it has been shown here that evolution has exploited the different parts and features of the LPS moiety in the OM environment to ensure optimal activity of

OmpT and hence maximising the protection it confers to the cell.

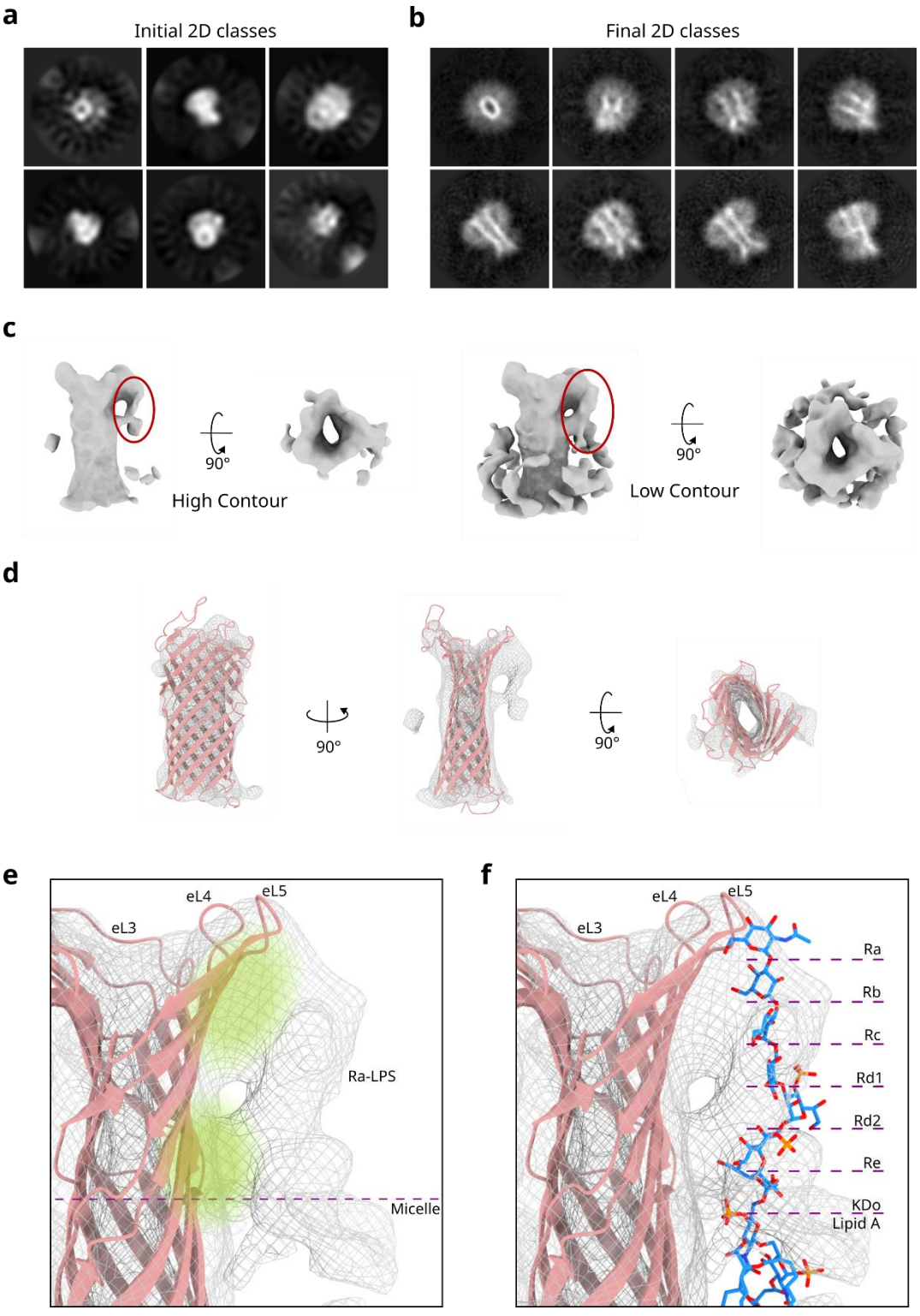


Figure 4.11: (Legend overleaf)

Figure 4.11: OmpT-LDAO:Ra-LPS resolved to ~6 Å and had clear additional density, attributable to Ra-LPS. (a) Initial 2D classes (boxsize 22.3 nm) and **(b)** 2D classes from cleaned particle stacks. Classes in (b) demonstrate data quality as despite limited signal clear asymmetric features can be observed (boxsize: 13.9 nm). **(c)** High and low contour map of the final reconstruction, additional non-protein and extra-micellar density is circled in red. **(d)** The OmpT crystal structure fits well into the obtained density (for determination of model orientation see main text). **(e)** Enlarged view of the non-protein density, green highlights indicate probable OmpT interaction regions. **(f)** Comparing the sugar locations of an extended Ra-LPS molecule (branching sugars not shown for clarity) to the cryoEM density suggests stable OmpT interactions between the upper and lower, but not central, sugars of LPS. (OmpT PDB: 1I78[363]).

4.2.6 OmpLA activity is not altered by LPS

Given the strong effects of LPS on OmpT activity, the activity of the OM phospholipase OmpLA (**Fig. 4.1d**) in the presence and absence of LPS was assayed, to see whether this was a general phenomenon. In Chapter five, two assays are described and validated for OmpLA activity and are used to assess possible consequences of OMP-OMP interactions. Both assays monitor the degradation of the liposomal membrane, either by measuring lipid aggregation by increasing light scattering or detecting increasing concentrations of free fatty acids by fluorescence (full details of experimental validation and rationale are described in Section 5.2.5, **Fig. 5.14**). Here OmpLA activity is measured in DMPG and DMPG-1% Ra-LPS liposomes (OmpLA at ~ 0.1% mol/mol). Initial measurements with the light scattering suggested that there was no difference between the activity of OmpLA in these two liposomal membranes (**Fig. 4.12a**) and this was confirmed using the fluorescence assay which showed no rate changes between the absence and presence of LPS

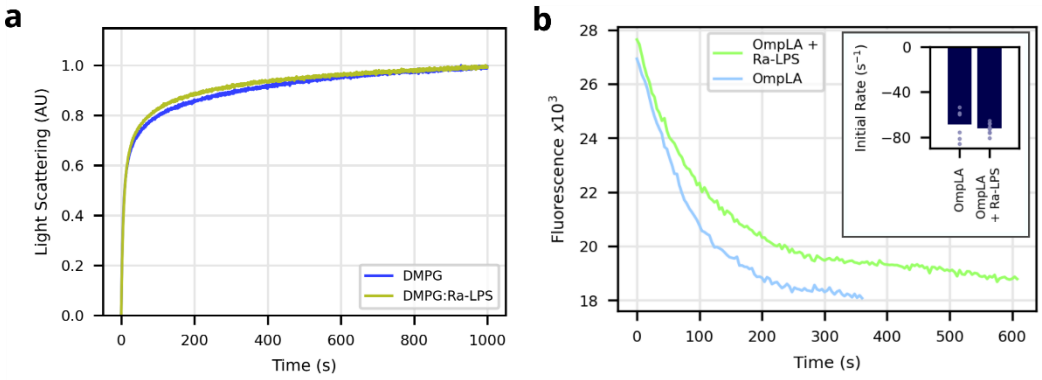


Figure 4.12: OmpLA activity is not modulated by LPS. (a) OmpLA activity by light scattering and **(b)** fluorescence in DMPG and DMPG + 1% Ra-LPS liposomes is not significantly different. (166 nM OmpLA, 1mM DMPG).

(**Fig. 4.12b**). Thus, while the activity of OmpT is strongly affected by LPS, the activity of OmpLA is not.

4.2.7 FusA requires LPS to be resolved by cryoEM

Thus far the nature of LPS-mediated rate enhancement of an enzyme's activity (OmpT) and the lack of activity change of another enzyme (OmpLA) have been biochemically analysed in detail, and possible conformational changes induced by LPS have been hypothesised for OmpT. To understand how LPS binding might affect the structure of another OMP, a larger target that is more amenable to solution structural studies by cryoEM was sought. The protein FusA was chosen due to its particularly extensive extracellular loops (which are more likely to be LPS interaction surfaces) and its relatively large (for an OMP) monomeric size of 100 kDa (**Fig. 4.1e**). FusA from *Pectobacterium* was purified in *E. coli* as an inclusion body and refolded into LDAO micelles (Section 2.6.1). FusA was imaged by cryoEM using a wide range of sample and hardware options, including optimisations in grid preparation (detergent, sample concentration and grid glow-discharge), microscope detector (K2, Falcon 4/4i), with and without an energy filter, and at additional optics parameters (low defocus, total dose), see **Fig. 4.13** (full microscope parameters for each collection are in **Table 2.2**).

Despite this broad range of optimisations and attempts, intermediate-high resolution structural solution failed. Image contrast, and thus signal, was improved by optimising grid preparation to ensure evenly thin ice – with final conditions including an amylamine glow-discharge and a pause between sample application and blotting of 10 s – and improving imaging hardware. While these yielded marginally improved 2D classes, it did not enable significantly better reconstruction (**Fig. 4.13**, collections 1-6). Upon inspection of the 3D reconstructions generated from these data, it was observed that the central plug domain that sits inside the barrel was forming symmetric cylindrical density, rather than adopting its expected asymmetric shape (**Fig. 4.14a**). In addition, the shape of the transmembrane barrel was circular, rather than the expected rounded-triangular shape, and thus did not provide a good match to the previously solved crystal structure (**Fig. 4.14b**). Together, this suggested that the images are rotationally averaging, with the pseudo-rotational-symmetry axis down the centre of the barrel dominating the signal and overwhelming the rest of the protein signal, thus preventing alignment on higher-resolution, asymmetric features (or that the protein structure is significantly disrupted). This is further supported by

consideration of the 2D classes (classified from particles contributing to the 'best' final model from indicated collections), which show strong signal for the barrel (which was circular in top-views), but only weak, featureless density for the plug and extracellular loop regions of the protein (**Fig. 4.14c**, left), compared to the asymmetric and strong signal expected (**Fig. 4.14c**, right)

Using the optimised grid preparations, a broad range of hardware options were used to collect images – including the K2, Falcon 3, 4 and 4i detectors, and a Falcon 4i with a Selectris energy filter. However, despite the improving quality of these detectors and an observable, concomitant improvement in the raw images, a high-resolution structural solution still eluded. Further adjustment of the imaging parameters, including low-defocus, stage tilt and tilt-pair collections failed to break the false rotational-averaging applied to the models. Solubilising FusA in DDM detergent rather than LDAO also did not help. Thus, despite extensive efforts, solving apo-FusA structure remained intractable.

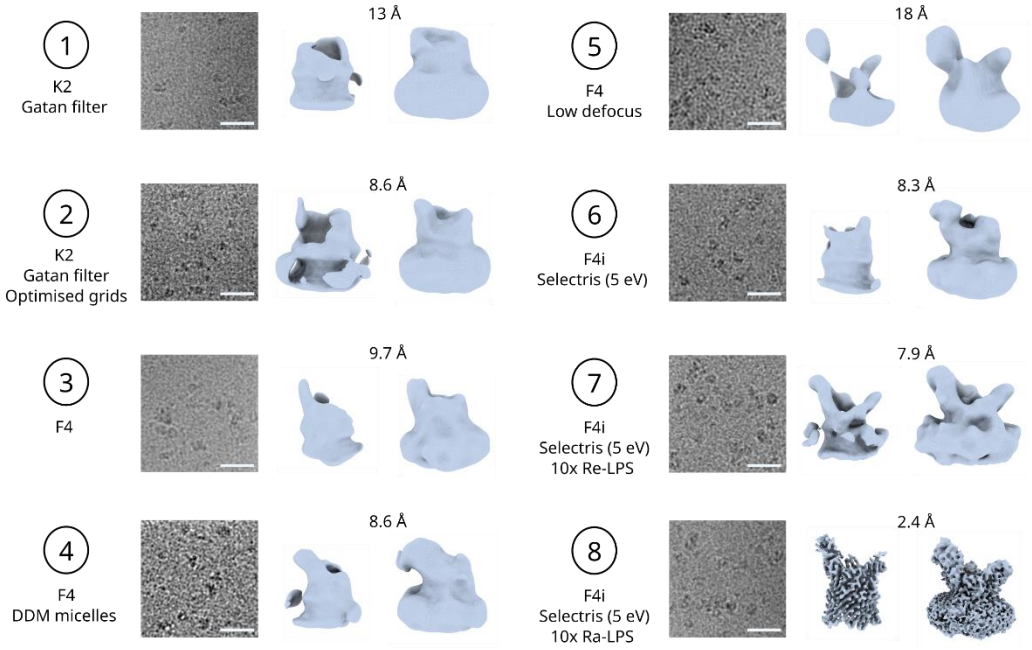


Figure 4.13: FusA requires LPS to resolve. Apo-FusA across multiple collections with various hardware and data collection parameters did not resolve (1-6, reported resolution values are nominal due to additional problems with these models, see **Fig. 4.14**). Upon addition of Re-LPS (7) or Ra-LPS (8) FusA resolved to low-resolution and high resolution respectively. High-contour (left) and low-contour (right) representations of each model are presented. Scale bars are 25 nm.

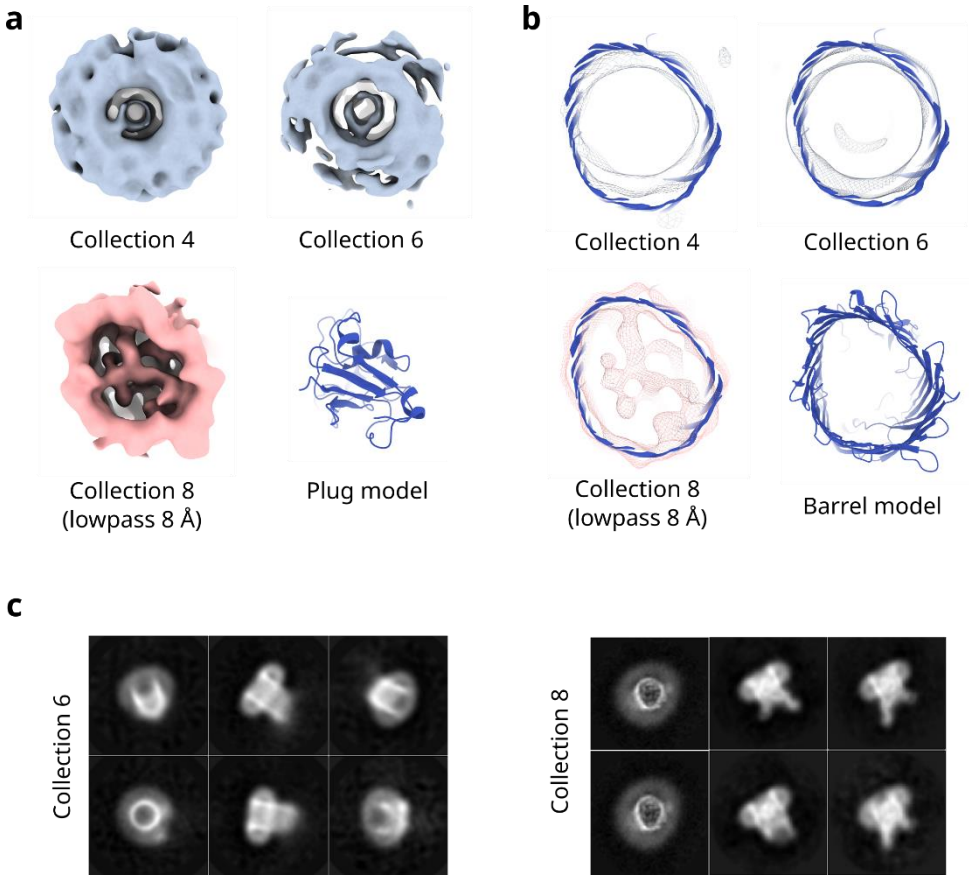


Figure 4.14: Apo-FusA failed to align well, with models showing rotational averaging. Rotational averaging was observed in **(a)** the plug domains and **(b)** the barrel of apo-FusA models with the formation of symmetrical cylinder-like density (upper) rather than the expected asymmetric shape resolved upon the addition of LPS (lower). **(c)** Comparison of best 2D classes from apo-FusA (left) and Fusa:Ra-LPS (right). Apo-FusA classes highlight the alignment to the pseudo-symmetry axis by the lack of detail and presentation of symmetrical features (compared to expected, right). Boxsize is 18.2 nm.

4.2.8 Molecular dynamics reveals specific FusA-LPS interactions

Some OMPs, including the TBDT FhuA[381] (**Fig. 1.21b**) are known to interact specifically with LPS in the OM. To probe how LPS might interact with FusA, the protein was simulated with Re-LPS and Ra-LPS using CG-MD. FusA was simulated with (A) 25 % (mol/mol) of these LPS-forms randomly placed asymmetrically in the outer leaflet of the membranes with DMPE (**Fig. 4.15a**), and (B) 10 LPS (~1.5 %, mol/mol) initialised singularly in a distant ring around FusA (**Fig. 4.15b**) (previously shown to allow significantly faster diffusion of LPS, **Fig. 4.9b**). The inner leaflet was composed of 80:15:5 DMPE:DMPG:CDL2. As the clustering and slow diffusion times of LPS prevented full equilibration and lipid-protein contact convergence in the system 25% Ra-LPS, it is not discussed further. Lipid-protein contacts in simulations with 25% Re-LPS and 1.5% Re- or Ra-LPS were found to have converged after 6 µs (**Fig. A11**).

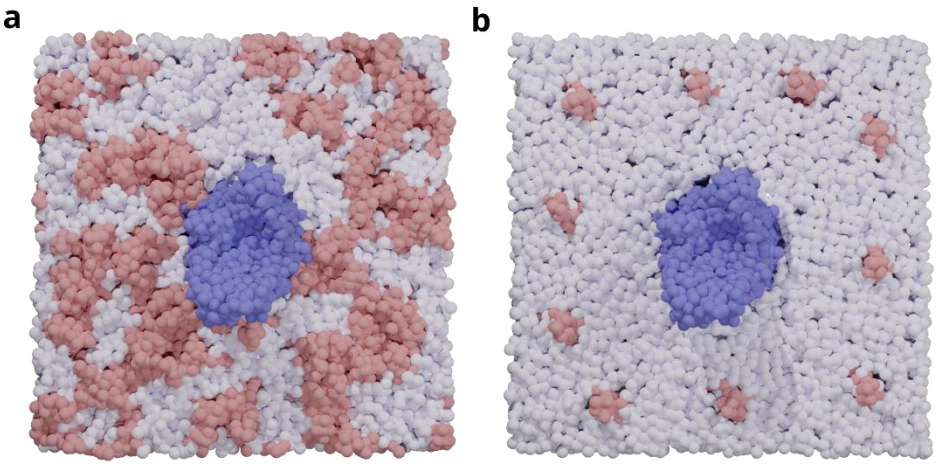


Figure 4.15: Starting frames for (a) 25 % and (b) 1.5 % Re-LPS FusA simulations. Concentrations are mol/mol. (Re-LPS: pink, FusA: blue, phospholipid: gray)

Analysing the protein-lipid contacts between the phosphates of the LPS and each residue of the protein, specific sites of contact can be identified. The 25% Re-LPS simulations show three likely Re-LPS binding sites of differing strength (**Fig. 4.16a** left, **b**). The strongest is around the β-seam, site 1, and weaker sites at β7-9 (site 2) and β13-15 (site 3), all of which have positively charged residues. Calculating lipid density around the protein over the simulation time also shows these three sites (**Fig. 4.16b**, right). However, the LPS clustering at this concentration makes it harder to identify isolated binding sites as any bound LPS is part of a cluster, which could obscure binding sites via interaction with adjacent regions of the protein. This effect

is much reduced in the sparse LPS simulations, where the Re-LPS density plot highlights the tendency of the Re-LPS to cluster around the protein with most of the bulk lipid area having negligible Re-LPS density (**Fig. 4.16a** centre, **c**). The contact analysis indicates a retained strong site around the β-seam, but weaker interaction sites with the rest of the protein. (Given the reduced concentration of Re-LPS, there is less LPS available to interact with the protein). Calculating the residence time of Re-LPS contact with each residue of the protein indicates that all expected accessible regions (external leaflet portions of β-strands and near membrane eLs) were sampled during the simulations in multiple replicas (**Fig 4.17a**) increasing confidence in the Re-LPS:FusA interaction sites. Similar to the contact analysis, the residence time per residue also indicates that the strongest interaction site is around the β-seam (**Fig. 4.17b**).

Ra-LPS, with the full sugar-core that extends several nm away from the membrane, can form larger surface contacts with FusA. Considering the protein-lipid contacts in the sparse Ra-LPS simulations shows that while sites 1 and 2 (as identified for Re-LPS:FusA) are maintained by Ra-LPS, the site 3 is missing and a new interaction location is observed (**Fig. 4.16a** right, **d**). These data, coupled with different apparent interaction strengths, suggest that FusA has LPS type specific interactions. Ra-LPS moieties bind tightly to FusA and are rarely released from the protein over the simulation sampling time. Considering the average lipid residence time as for the Re-LPS demonstrates good sampling over FusA by Ra-LPS during the simulations (**Fig. 4.17c-d**). Given the extended polysaccharide chain present in Ra-LPS, it seems likely that it is interacting specifically with FusA and thus altering the lipid binding patterns. Similar to previously reported data[509], cardiolipin in the inner leaflet is excluded from under the LPS (**Fig. A12**). Thus, FusA has strong and lipid-unique interactions with Ra-LPS and Re-LPS. This pattern of different interactions with different forms of LPS is consistent with the data presented above for OmpT, where interactions with specific regions of the LPS oligosaccharide are responsible for activity modulation.

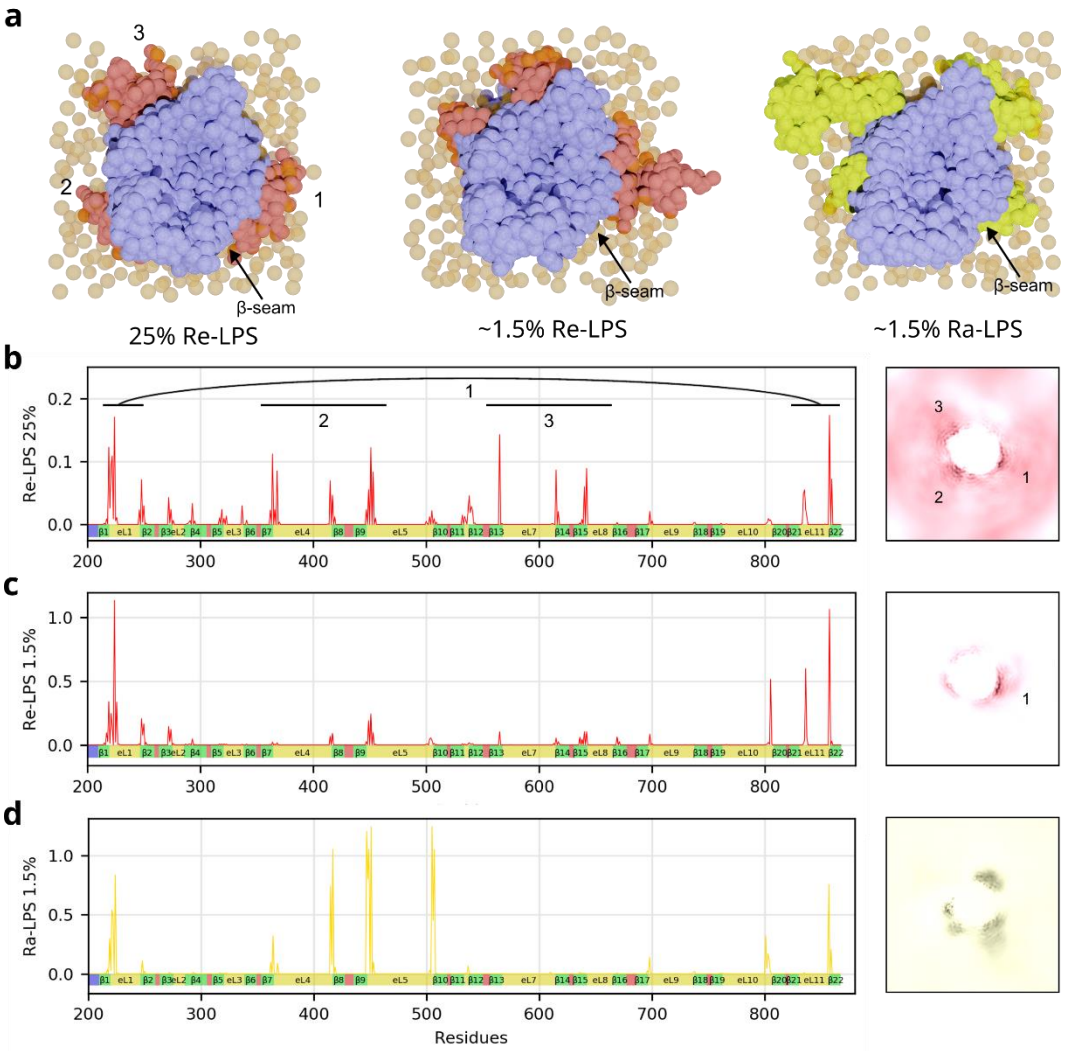


Figure 4.16: FusA specifically and uniquely interacts with Ra-LPS and Re-LPS. (a) Typical frames from simulations with 25 % Re-LPS (left) 1.5% Re-LPS (centre) and 1.5% Ra-LPS (right). Image frames determined by clustering all frames of the last 500 ns of the simulation and taking a single frame from the largest cluster. Numbering in 25% Re-LPS corresponds to numbering in **(b)** (orange: phosphates, blue: FusA, red: Re-LPS, yellow: Ra-LPS). Normalised (by lipid count and frame number) lipid-protein contacts and Re-LPS density (averaged over the simulation time) for **(b)** 25% Re-LPS, **(c)** 1.5% Re-LPS, and **(d)** 1.5% Ra-LPS. Lipid density maps are aligned as in part **(a)**. (In all parts concentrations are mol/mol).

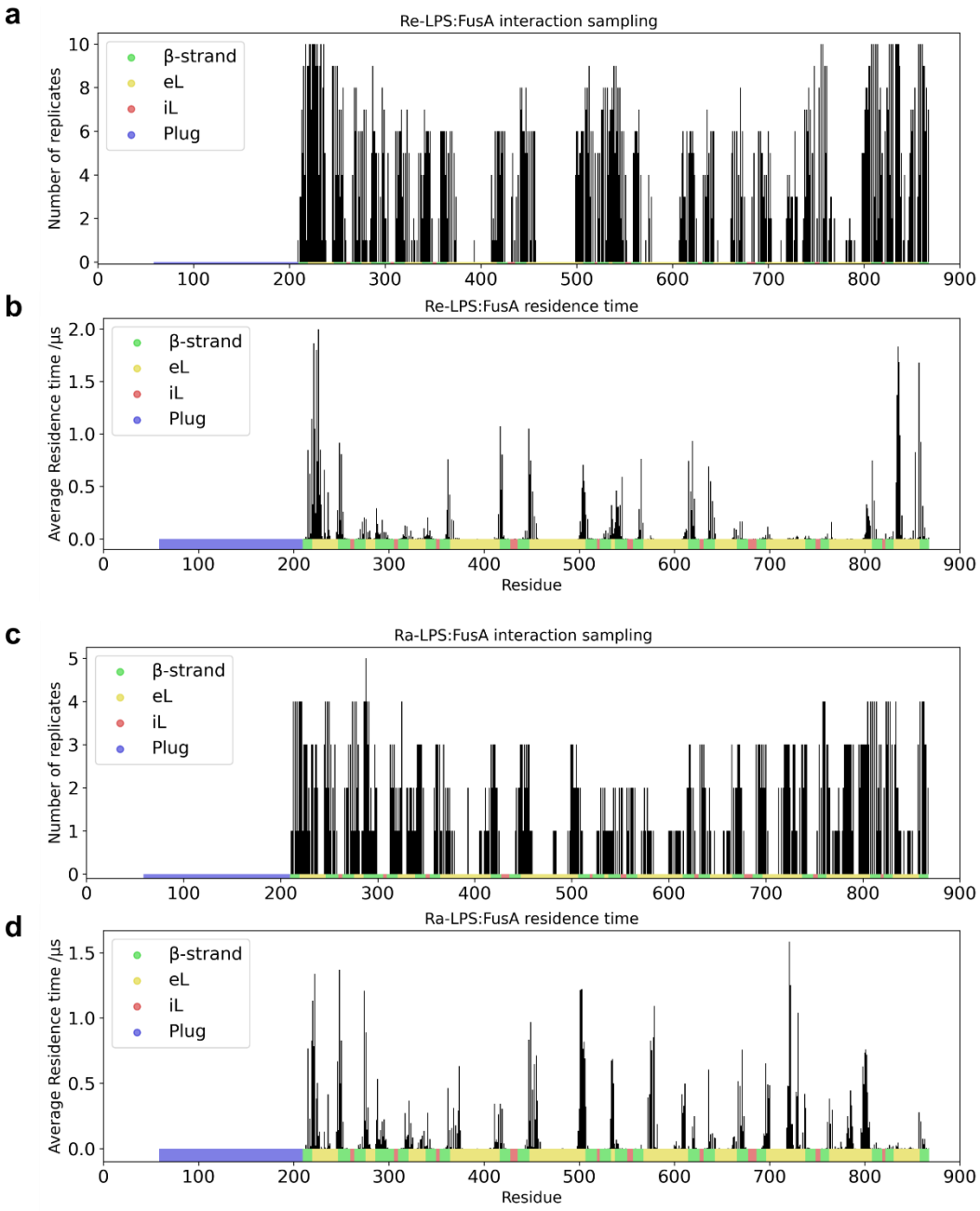


Figure 4.17: Both Ra-LPS and Re-LPS sufficiently samples FusA across replicates in the sparse LPS simulations. (a) Re-LPS:FusA sampling shown by the number of replicates where an LPS-FusA contact has been made (for any length of time). **(b)** Average residence time of Re-LPS per residue of FusA over all replicates of the ~1.5% Re-LPS simulation. **(c)** and **(d)** as (a) and (b) for Ra-LPS.

4.2.9 CryoEM resolved FusA:Ra-LPS binding

Given the specific and apparently strong LPS-FusA interactions indicated by the CG-MD, Ra-LPS and Re-LPS from K12 *E.coli* were doped into the LDAO micelles of FusA at a 10:1 LPS:protein ratio (mol/mol). The structure of LPS is well conserved between *E. coli* and *Pectobacterium (carotovora)* with the Lipid-A moiety, and the first four sugars of the polysaccharide chain being identical, with variation in the further extended and branching sugar identity (**Fig. A13**), and thus interactions are expected to be similar. CryoEM micrographs of the LPS-FusA data did not appear different from previously collected datasets (**Fig. 4.13**), and initial 2D classification yielded similar looking classes to those previously obtained. However, following multiple rounds of divisive 2D classification on specific views of the protein, high-resolution, asymmetric classes were obtained (**Fig. 4.14c**, right). Training particle picking models explicitly on each of these classes and recombining the data yielded enough particles to reconstruct good models for both the Re-LPS and Ra-LPS datasets. The Ra-LPS dataset, for which about twice the data was obtained (~8 000 micrographs vs ~4000 micrographs), finally resolved to 2.4 Å (**Fig. 4.18a-b**, **A14**), and the Re-LPS data to ~7.9 Å (**Fig. A15**). The structure of the protein, especially in the plug and transmembrane barrel region, is similar to the previously solved crystal structure (average global RMSD: 1.2 Å), although the cryoEM structure is at higher resolution (2.4 Å vs 3.2 Å). While the barrel shape of FusA:Re-LPS is consistent with FusA:Ra-LPS filtered to an equivalent resolution (**Fig. 4.18c**), there is reduced density for the plug (**Fig. 4.18d**), suggesting that it has been at least partially dislodged from the pore. Refining the optimised particles from data collections 1-6 against the high-resolution model failed to yield a better solution, indicating that the presence of the LPS is critical to enable solution. An additional, smaller data collection (~2500 micrographs) of FusA:Ra-LPS on a new sample yielded a correct but low-resolution model (9.5 Å, **Fig.**

Figure 4.18: High resolution FusA:Ra-LPS. (a) FusA:Ra-LPS resolved to 2.4 Å with at least four bound Ra-LPS moieties (yellow). **(b)** Details of well-resolved side-chains in Fusa:Ra-LPS model. **(c)** The FSC curves of the model. The masked/corrected maps do not fully drop to zero, likely due to the model resolution practically reaching Nyquist. **(d)** Comparison of overall shape and structure of FusA:Re-LPS (blue) and equivalently lowpass filtered FusA:Ra-LPS (pink) shows broadly similar barrel shapes, while **(e)** the plug domain in FusA:Re-LPS appears to be missing compared to lowpass filtered FusA:Ra-LPS (right) or the modelled plug (left). **(f)** Another small data collection of FusA:Ra-LPS yields a low-resolution but correct reconstruction shown at two contour levels (upper) and compared to an equivalent lowpass filtered model.

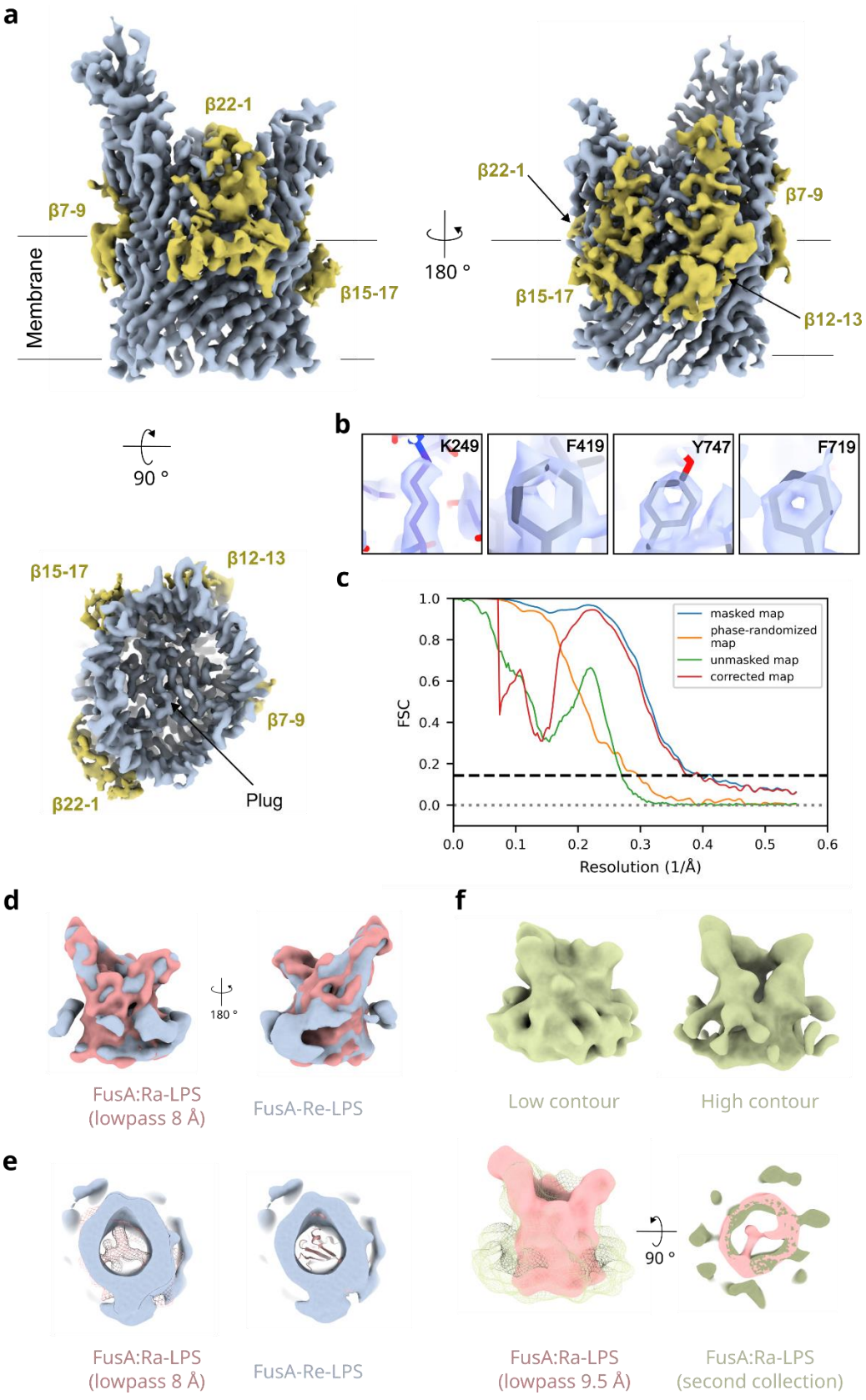


Figure 4.18: (Legend on preceding page)

4.18e). Unlike reconstructions in the absence of LPS this model has both correct loop conformations (within the resolution) and an asymmetric barrel shape (although like FusA:Re-LPS had weak plug density), confirming that the addition of LPS is required for structural resolution.

Indeed, non-protein density was observed in the Ra-LPS:FusA reconstruction, and this was assigned as LPS based on the presence of ring-shaped sugar moieties (LDAO detergent has no cyclic chemical groups) and depending on whether there was also adjacent non-protein extra-micellar density that could belong to the oligosaccharide chain. This analysis revealed four confidently assigned, extracellular LPS molecules, where the lipid-A sugars bound at β7-9, β12-13, β15-17 and at β22-β1 (**Fig. 4.18a-b, Fig. 4.19a**), and the less well resolved polysaccharide sugars were observed to varying extents, and although additional density was present, the upper LPS sugars could not always be confidently modelled. In all cases multiple acyl chains from each LPS moity are seen. The doped-in Ra-LPS could freely associate with the extracellular or intracellular side of the barrel, while acyl chains were seen on the intracellular side, they were not well-enough resolved to confidently determine the presence of ring-moieties, and hence their identities as LPS or LDAO. This suggests a binding preference for the LPS to its native, extracellular position.

The binding site at β7-9 and β22-1 (e.g. **Fig. 4.19b**) contains a canonical FhuA-LPS binding motif (**Fig. 1.22** c.f. **Fig. 4.1b**) formed by K368/R451/R453 and R219/K224/R857, respectively. Although positive residues are found adjacent to the LPS-phosphates in other bound LPS (indeed, Lipid-A is found exclusively bound to the electropositive surface regions, while it is excluded from regions of electronegative surface potential), they do not form the binding motif, indicating that additional chemical features are strengthening the binding at these regions. In some cases, grooves on the protein's surface hold the acyl chains (e.g. **Fig. 4.19c**) or oligosaccharide sugars (e.g. **Fig. 4.19d**). In all cases residues of FusA's well-defined aromatic girdle interact with the upper regions of the acyl-chains (e.g. **Fig. 4.19e**). No other residue-specific conserved patterning is seen in the LPS interactions, although more generally the binding surface is dominated by non-polar or weakly electronegative residues (expected patterns for oligosaccharide binding).

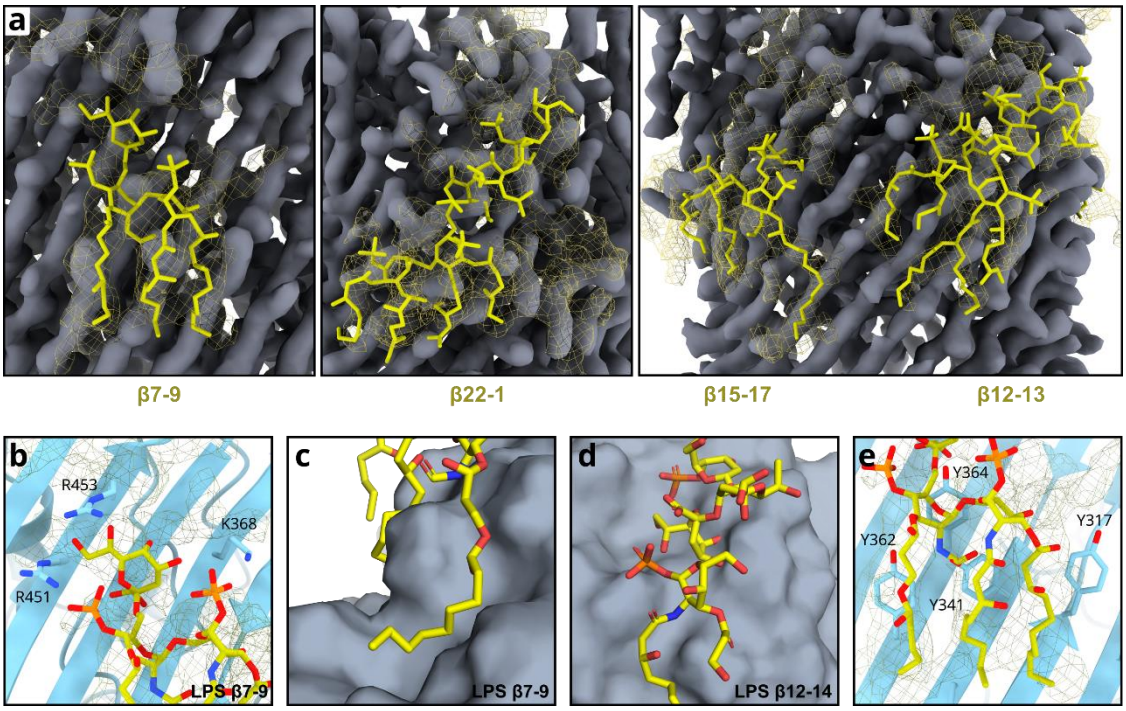


Figure 4.19: Ra-LPS binding to FusA. (a) Overview of the four fitted LPS moieties and their densities (yellow), LPS was only modelled into the density that could be confidently modelled. **(b)** LPS β7-9 Lipid-A binds at a canonical LPS binding site formed by K368/R451/R453. **(c)** Some acyl chains of LPS were observed residing in grooves on the protein surface, **(d)** as were some oligosaccharides. **(e)** All bound LPS moieties interacted with the underlying aromatic girdle.

Comparing the Ra-LPS binding sites observed by CG-MD and in the cryoEM structure generally show good agreement, with three of the four LPS molecules seen in the cryoEM density binding at each of the three binding sites predicted by the simulations (**Fig. 4.20a**), with strong similarity between the top-down views of a simulation end frame and the EM density. Intriguingly, the additional molecule of bound LPS in the structure is at the Re-LPS binding site that minimally bound Ra-LPS during the simulations, suggesting that in an excess of LPS, Ra-LPS will bind to this site also. Considering only the lipid-A sugars and phosphates, the regions of LPS best resolved also show occupancy at all three of the LPS binding sites predicted by the MD (**Fig. 4.20b**). Thus, the CG-MD approach successfully identified the major LPS binding sites, as verified by experimental structural solution, which was only possible in the presence of the bound LPS moieties.

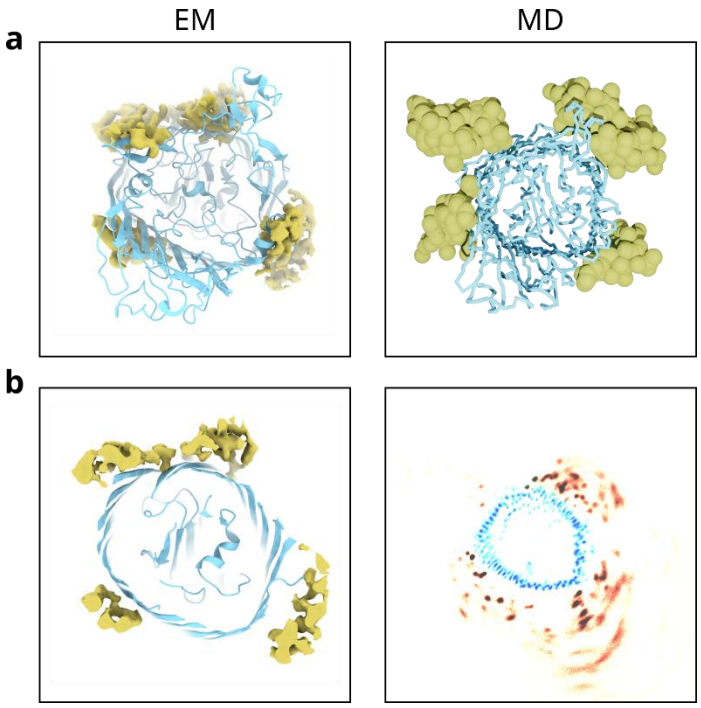


Figure 4.20: LPS binding by cryoEM and MD broadly agree. (a) Comparison of EM model (left) and a LPS-bound consensus frame from the MD (right) and **(b)** of a slice through the EM model at approximately the Lipid-A phosphate position (left) and the probability density of finding the phosphates around FusA calculated over the final 3 µs of the sparse Ra-LPS simulations (right).

4.3 Discussion

In the half-century since its discovery and protease characterisation[352,353], the omptin family, and particularly its links to virulence and pathology, have been extensively studied[623]. However, while the dependence on LPS for omptin catalytic activity is well-known[389,612,614], some details of its nature and mechanism have remained obscure. It is shown here that different sugar components of the LPS core-oligosaccharide mediate a two-stage activation of OmpT, that is partially exchangeable to environments that mimic some properties of LPS. It has previously been shown that certain modifications of LPS are able to modulate the ability to enhance OmpT activity, in particularly the need for phosphorylated heptoses and high levels of Lipid A acylation (≥5), but only with reference to shorter core-oligosaccharide truncations (Rc and shorter)[612]. These studies therefore inform on determinants of the first-stage of activation. The importance of upper core sugars has been implicated in Pla activation[389], while other studies dispute this (although under very different experimental conditions)[614], but this has not been systematically studied. Intriguingly, the lack of an O-antigen improved the activity of Pla[389], consistent with data presented here. A recent study suggests that removing the positive residues in the LPS-binding motif in the extracellular loops fails to abolish LPS mediated activation[611], suggesting that additional interactions are important, and supporting the conclusion that LPS interacts additionally at the extracellular loops.

The proposed binding motif of LPS on the surface of OmpT[612,614], supported by the cryoEM and MD data presented here, is based around the Lipid A moiety, and no previous report suggests the importance of the upper core sugars of LPS for mediating the interaction. However, it is shown here that specific interactions with eL4/5 of OmpT with the upper (Rb/Ra) sugars of LPS facilitate an additional gain in activity. A recent *in vivo* characterisation of OmpT activity in cells with various LPS truncations (Rd1 and shorter) demonstrated that activity was reduced in all cases compared to the full-core chemotype[502]. While interpreting this data is challenging due to the complex changes that occur in cells upon alteration of their LPS, it is suggestive that the upper sugars of LPS are required for optimal activation in agreement with the *in vitro* data presented here. It has been further shown here that at least part of the mechanism of activation is via overcoming the product inhibition that occurs in the absence of LPS. This is supported by a substrate-Pla structure in the absence of LPS in which the substrate binds deeper and more stably than expected in the active-site cleft[613]. Intriguingly, *Shigella flexneri* which generally shortens or eliminates LPS O-antigens during its intracellular lifecycle stages, maintains LPS oligosaccharides at a

full-core equivalent, perhaps in part to keep the virulence critical omptin SopA maximally active. While some elements of OmpT activation by the truncated LPS chemotypes remain uncertain, it seems that biology has exploited specific parts of the LPS molecule to mediate specific functionality of the omptin family.

LPS interactions with TBDTs are known to be important for function[504], and here the apparent stabilising of FusA by LPS binding is demonstrated by facilitating a high-resolution structural solution. LPS binding is known to stabilise OMP trimeric porin formation (**Fig. 1.21a**)[489], and has been shown to order the extracellular loops of BtuB[504,505] and OprH[508], thus the stabilisation of some inherent flexibility in FusA by LPS binding, giving enough signal to allow accurate particle alignment, is a plausible explanation of why LPS is required for a high resolution structural solution. The stabilisation of FusA upon native-lipid binding is intriguing and gives hints about the role of LPS in the native OM (that is, the ordering of protein structure) presumably for reasons of function or membrane-integrity. Given the findings presented here, we suggest that the addition of LPS may be a more general solution to solving the structure of relatively small OMPs by cryoEM. Indeed the recently solved FhuA structure, another small OMP solved by cryoEM also includes a stably bound and well-resolved LPS moiety[507].

As shown above for OmpT, it is well known that lipids are important for the functional role of many membrane proteins, and can support specific structural motifs, functional transitions and to help with substrate recruitment/binding, as well as forming their native environment. Over the last decade, cryoEM has become an established method to elucidate protein-lipid interactions[624]. However, these are typically identified either in nanodiscs, which can pose significant challenges for cryoEM structural determination and have a constrained possible lipid composition, or as strongly bound molecules that have remained through a traditional detergent purification. Here we have shown that it is possible to reconstitute native lipids into detergent micelles, distinguish them from micellar density and that they bind at the predicted protein binding sites (by CG-MD). Therefore, we propose the addition of native lipids back into detergent-protein and their high-resolution structural solution by cryoEM as a method to determine lipid binding sites. By using high-throughput cryoEM approaches coupled with lipid titrations of multiple lipid types, binding patterns and preferences could plausibly be determined.

A few OMPs have previously been solved structurally with unambiguous LPS bound, including FhuA (crystallography/cryoEM[381,507]) and the OmpE36 trimer

(crystallography, two distinct binding sites[489]). The structural impact of LPS binding have been determined on OprH by NMR, but the lipid itself was not observed[508]. The FusA-LPS structure presented here demonstrates a range of LPS binding modes, and highlights some of those observed before, including the previously structurally characterised FhuA-like LPS binding motif, although the data here indicate a diversity of binding modes including important contributions from both chemical and steric features. The good agreement between the identified locations of LPS and between the MD and cryoEM is an encouraging validation of the ability of the martini forcefield to simulate LPS-protein interactions accurately.

In summary, these data give insight into the importance of specific OMP-LPS interactions. Although additional characterisation is required to mechanistically understand how LPS can effectively modulate OmpT and FusA structure and/or function, the results demonstrate the critical importance of LPS-OMP interactions to maintain a structural and functional OM and highlights the matching between specific regions of LPS and proteins for optimal function.

Chapter 5: Specific protein-protein interactions modulate OMP functionality

The outer membrane (OM) of lipopolysaccharide (LPS)-diderm bacteria is crowded with OM proteins (OMPs). This high protein content is dominated by a few specific OMPs, situated in a sparse lipid context and experiencing minimal lateral diffusion, enforcing long-lived protein-protein interactions. While these features mean that specific, functional OMP-OMP interactions are biochemically plausible and evolutionary expected, they have not yet been demonstrated. Here, following characterisation of DMPG membranes as facilitating near-universal OMP folding, the very common OMP OmpA is shown to enhance the activity of OmpT by increasing the substrate affinity in a specific, 1:1 interaction that is not observed with any other common OMP. This interaction is further assessed through structural modelling and crosslinking experiments. In contrast, OmpA is found to reduce PagP's activity by ~50%, while OmpLA activity shows no significant change. An array of >500 possible OMP-OMP interactions between the abundant OMPs of the OM and all other OMPs was screened by predicting complexes using Alphafold2. The results suggest that smaller OMPs are generally more promiscuous interactors and identify several new interactions. Together, these data point to the previously unappreciated role of specific OMP-OMP interactions to modulate protein function (and clustering) in the OM, as evolution exploits the near-infinite local concentrations of highly abundant proteins in the diffusion limited OM.

Proteins discussed in this chapter were expressed and purified by Dr Bob Schiffrin (OmpLA, OmpA, BamA), Dr James Whitehouse (tOmpA) Dr Anton Calabrese (OmpF) and Dr Samuel Haysom (OmpX). TMB2.17LA, TMB2.3LA and OmpTrans3 arise from a collaboration with the lab of Prof. David Baker[625] and TMB2.17LA and TMB2.3LA were generated and purified by Dr James Whitehouse. Circular dichroism was performed with support from Nasir Khan.

5.1 Background

Lipid and protein constituents of biological membranes are typically constrained and organised by additional levels of higher-order membrane structure[30,626]. In the OM of LPS-diderm bacteria this is particularly striking, with lipids and proteins partially phase-separating into OMP islands (which can be >500nm in diameter[140]) and LPS patches (typically ~55 nm diameter[275]), each significantly enriched in their own component[275]. A combination of the very low LPR (~7:1, c.f. ~32:1 for the LPS-diderm inner membrane[160]), the stable arrays formed by trimeric porins[490], the relative immobility of the large LPS and the LPS oligosaccharides associating together via divalent cations[494] means that only very limited diffusion is possible[627]. Indeed, the rate of OMP diffusion scales with the local membrane protein density, although the extent of lateral constraint varies for specific proteins[628]. OMP-OMP interactions are thought to be mediated frequently by LPS[491], but given the scarcity of lipids and their partitioning into lipid-rich domains, many direct protein-protein interactions must occur in OMP islands, for example as characterised for OmpF-BtuB[36]. OMPs, about two thirds of which (in *E. coli*) are OmpA and the trimeric porin proteins OmpF/OmpC, typically take a few seconds to partition into the membrane via the BAM complex[139]. Folding occurs in 'folding precincts' consisting of multiple copies of BAM[252] (with folding competent BAM possibly localised to the mid-cell, at least for specific substrates[140,501]), further driving the formation of protein-protein interactions in the OM. This dense packing of proteins and the many interaction points between individual LPS moieties and multiple OMPs helps to generate the primary barrier function of the OM. A picture of the OM thus emerges as a mosaic of OMP and lipid islands within a background of the highly abundant OmpA and trimeric porins, with protein-rich domains characterised by non-specific OMP-OMP interactions (**Fig. 1.20**).

In addition to the common porins, structural proteins and transporters in the OM, there are three enzymes found in K12 *E. coli*[434]. All three are constitutively expressed and implicated in virulence[437,616,629], playing a key role in rapid management of cellular threats before typical adaptive responses can adjust the OM composition[434]. The protease OmpT and its interactions with LPS were explored in Chapter 4. Briefly, OmpT (**Fig. 5.1a**) is a relatively abundant (~6% of the OMPome[273]) 10-stranded protease responsible for cleaving host antimicrobial peptides[630] that has a facile fluorescence based activity assay[105]. The phospholipase OmpLA (or PldA, **Fig. 5.1a**, typically ~0.2% of the OMPome[273]) degrades any phospholipids that are

erroneously incorporated into the outer leaflet of the OM, to yield fatty acid and lysophospholipid[420], which are then presumably either recycled by the cell[418,419] or released into the environment. OmpLA is a 12-stranded barrel with an external facing active site, that is only fully formed upon divalent cation mediated enzyme dimerization[631]. OmpLA preferably acts on acyl-chains of 14 carbons or longer and is largely non-specific to phospholipid headgroup[411]. While the enzyme has been kinetically characterised in detergents using synthetic substrates[409], studying its activity in synthetic membranes is challenging due to their rapid degradation upon enzyme activation. The third enzyme, PagP (**Fig. 5.1a**, typically <0.1% of the OMPome[273]), is a lipid A palmitoyltransferase (transferring a palmitate group from phospholipids to LPS)[632], an adaptation that stiffens the OM and, among other roles, helps protect against cationic antimicrobial peptides[437]. In the absence of acceptor substrate (i.e. LPS) PagP also displays a slow phospholipase activity[435,436]. PagP is 8-stranded and has a short N-terminal α-helix that anchors it in the membrane at a tilt of ~25° relative to the membrane plane[443]. Interestingly, although in *E. coli* all three enzymes are implicated in pathogenesis, across a range of other pathogens one or more of these enzymes is inactivated[458,633,633], suggesting that enzyme activity may be incompatible with some pathogenic states. This highlights the likely advantages to the bacterium for precise modulation of the activity of these enzymes. However, while the lipid interactions of all three of these enzymes have been studied[369,440,634], their protein-protein interactions (beyond the OmpLA dimer) have not.

These enzymes sit in a membrane dominated by a handful of OMPs, with only six estimated to be present at >3% in the *E. coli* OM (when grown in rich media: OmpF, OmpC, OmpA, OmpX, OmpT, MipA), and ~65% of the total content formed from monomeric OmpA and the trimeric porins OmpF/OmpC regardless of environmental conditions[273]. Therefore, the majority of other OMPs in the membrane will be in contact with OmpA and/or a trimeric porin, either directly or via a single shell of lipids (owing to the low LPR of the OM), forcing stable interactions due to the constrained diffusion. Given these properties, it is plausible that biology has exploited common interactions to facilitate and modulate specific OMP functions. Indeed, functionality beyond their canonical role has been described for some common OMPs. OmpC for example (and possibly OmpF) interacts with lipoprotein MlaA to support the Mla lipid transport pathway[196,635], and the C-terminal domain of OmpA has been shown to interact with RcsF, helping regulate the Rcs stress response[478]. However, there remains no examples of functional intra-membrane ‘moonlighting’ (that is, hetero-oligomeric interactions not required for OMP canonical function) interactions between

OMPs, despite the apparent evolutionary opportunity this presents.

Here DMPG liposomes are characterised as facilitating near-universal folding for a wide variety of OMPs, and from which proteoliposomes of defined protein composition can be assembled. OmpA is then shown to approximately double the activity of OmpT when co-assembled into the same membrane by increasing the substrate affinity in a specific, 1:1 interaction that is not observed with any other common OMP. This interaction is further assessed through structural modelling and crosslinking. Through the development of a novel liposomal assay for PagP, it is shown that OmpA specifically reduces its activity by ~50%, while assaying OmpLA activity shows no significant change in the presence of any of the abundant OMPs. Next, an *in vitro* crosslinking approach is used to validate that while OMPs variably self-associate in membranes, their self-interaction is generally enhanced by the presence of LPS. An array of >500 possible OMP-OMP interactions between the abundant OMPs of the OM and the general OMPome was screened by predicting complexes using Alphafold2. This analysis suggested that smaller OMPs are generally more promiscuous interactors, as well as identifying several new, high-confidence interactions. Together, these data highlights the role of specific OMP-OMP interactions to modulate protein function in the OM, as evolution exploits the near-infinite local concentrations of highly abundant proteins in the diffusion limited OM.

5.2 Results

5.2.1 DMPG allows universal intrinsic folding of OMPs

To study OMP-OMP interactions systematically *in vitro*, a membrane system that is competent to fold a range of OMPs at relatively high protein concentration and high yield is required. PC lipids (of various acyl-chain length) have been used most commonly in the literature for intrinsic OMP folding studies[31]. However, in Chapter 3 it was shown that OmpA and BamA show faster folding kinetics, higher folding yields, and are more stable into pure DMPG liposomes compared to DMPC (**Fig. 3.9**). A rapid folding rate is particularly advantageous as it minimises aggregation in the low urea concentrations required to initialise partition into the membrane[265], while lower urea concentrations induce faster folding by destabilisation of the unfolded, soluble state. To enhance the folding rate further, folding can be performed at the membrane transition temperature, which has previously been shown to increase folding rates by ~10x (for OmpA in DMPC)[266].

Combining these considerations, intrinsic folding was attempted into DMPG liposomes at 23.5 °C (the T_m of DMPG, c.f. **Fig 3.8**) at a final urea concentration of 0.5 M. An array of different OMPs was tested to include the three enzymes of the OM (OmpT, OmpLA and PagP) and five of the six OMPs with protein counts estimated to be >3% in a typical *E. coli* OM, termed the abundant proteins throughout the chapter (OmpA, OmpF, OmpX, OmpT, MipA). The other highly abundant protein, OmpC, is a homolog of OmpF and both are trimeric porins. The transmembrane region of OmpA (tOmpA) was also considered, as was BamA given that all the above OMPs fold via BAM and thus must interact with it at least transiently. BtuB, a representative 22-stranded TonB-dependent transporter, was also included (**Fig 5.1a**). These OMPs span a range of sizes (8 to 22 strands, 18 to 87 kDa). Monitoring folding via tryptophan fluorescence shows folding curves for 8 of these 9 OMPs, and fluorescence spectra of the product indicate a folded protein in each case (**Fig. 5.2**). PagP did not fold under these conditions, not unexpected given previously reported results[446] (PagP folding is further discussed in 5.2.3). Two additional OMPs (FadL and LamB) also fold well under these conditions (**Fig. A16**).

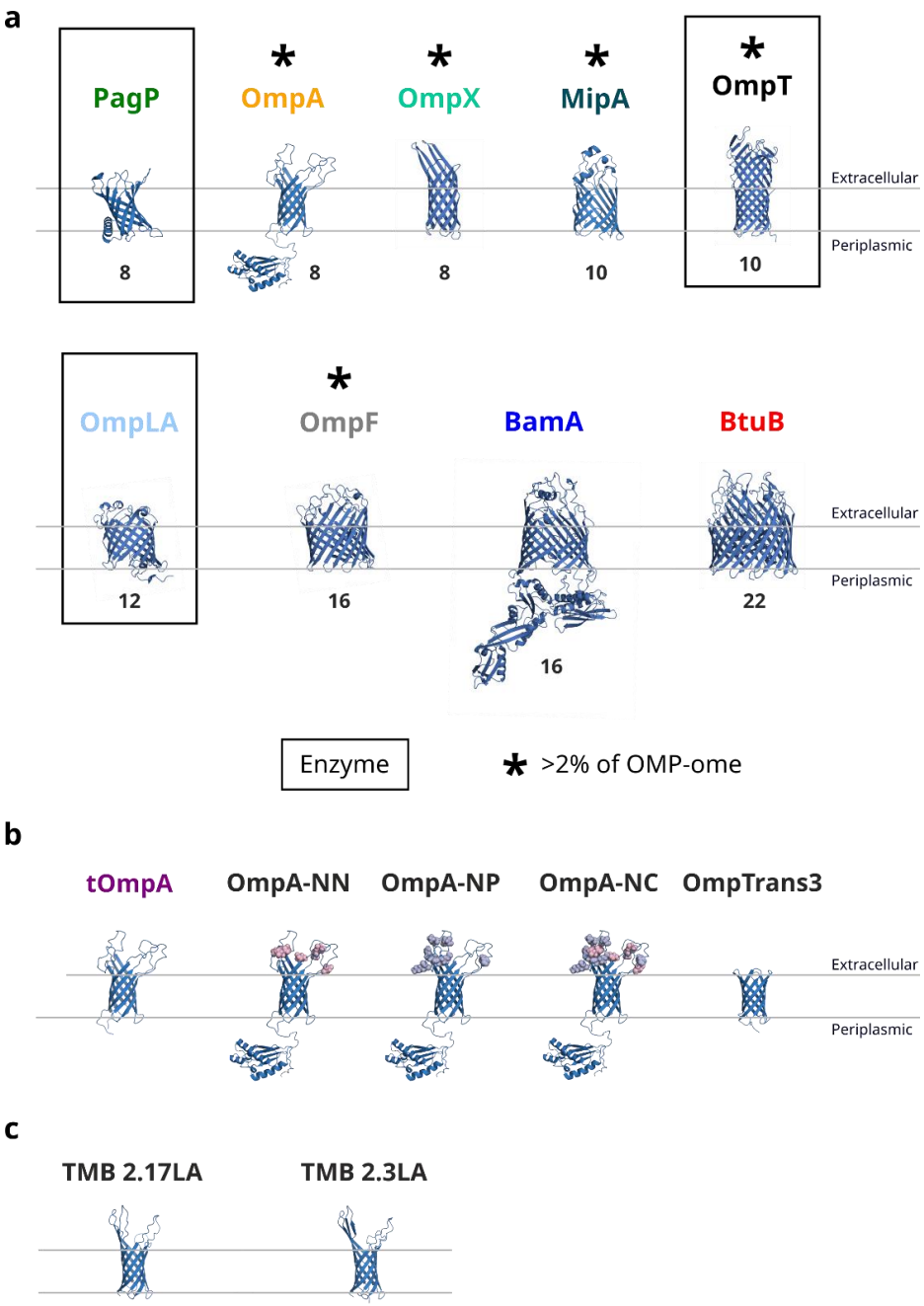


Figure 5.1: OMPs, OMP-mutants and designed β-barrel proteins used in this study. (a) The 9 native K12 *E. coli* OMPs considered, indicating the three enzymes (boxed) and the abundant OMPs (*). Underlying numbers are the strand number. Name colours are retained throughout the chapter. All models are crystal structures except MipA which is an Alphafold2 model (PDBs: PagP: 3GP6[441], OmpA: 1G90[290]/2MQE[284], tOmpA: 1G90[290], OmpX: 1QJ8[636], OmpT: 1I78[363], OmpLA: 1QD5[637], OmpF: 1MPF[638], BamA: 5D0O[243], BtuB: 1NQE[639]). **(b)** Synthetic OmpA variants, including tOmpA (transmembrane domain only), OmpTrans3 (the minimal transmembrane barrel in which the extracellular loops are swapped for the designed loops of TMB2.17[1], Alphafold2 model) and the OmpA loop charge mutants described in Section 3.2.6. **(c)** *De novo* designed transmembrane β-barrels. Two barrel scaffolds (2.17 and 2.3) have the loops of OmpA grafted on (Alphafold2 models).

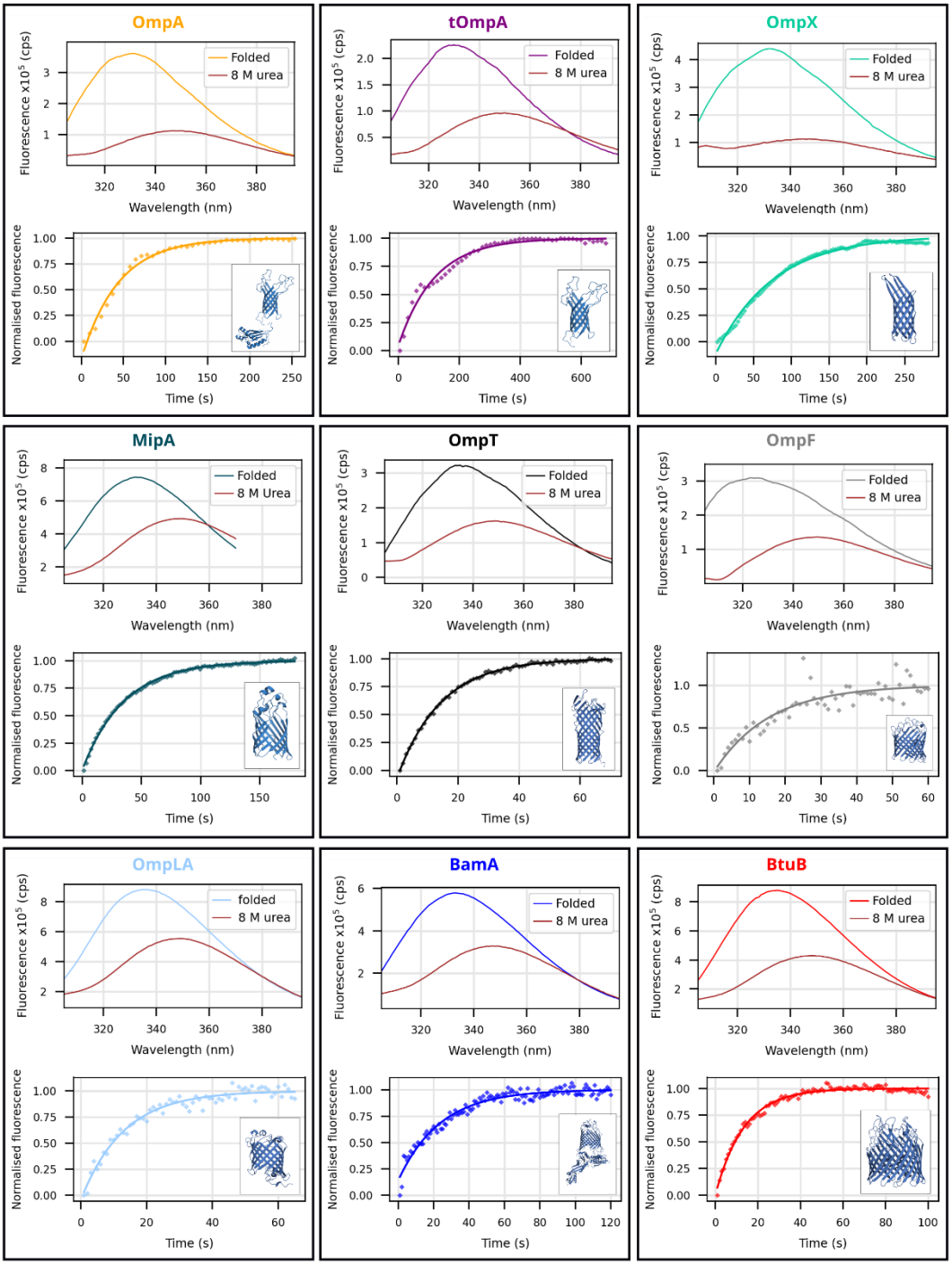


Figure 5.2: Nearly all native OMPs rapidly fold into DMPG membranes. For each of the OMPs in **Fig. 5.1a** (including tOmpA and excluding PagP which uniquely did not fold under these conditions) the folding kinetics (lower panels, followed by tryptophan fluorescence) and the folded vs unfolded (8 M urea) intrinsic fluorescence spectra are shown. All of these proteins show clear, rapid folding kinetics and the expected unfolded-folded spectral transition. (PDBs as in **Fig. 5.1**)

The intrinsic folding of MipA, which was cloned and purified for this study (see Section 2.4), has not previously been reported in the literature. Cold SDS-PAGE analysis showed it lacked a folded-unfolded band-shift, although an additional faint, presumably dimeric, band appeared in the unboiled sample (**Fig 5.3a**). To ensure that MipA is adopting a barrel structure in the membrane, circular dichroism spectroscopy was used to characterise MipA refolded into detergent and DMPG liposomes, both of which show a gain of β-strand secondary structure compared to the unfolded protein in 8 M urea (**Fig 5.3b**).

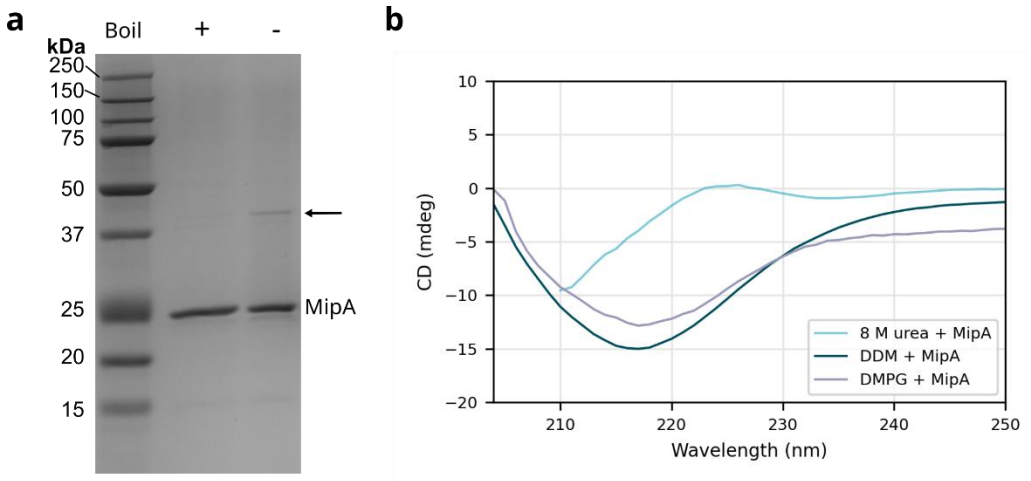


Figure 5.3: MipA characterisation. (a) MipA does not show a folded-unfolded bandshift by cold SDS-PAGE, although a minor additional band appears in the unboiled sample (arrow) at approximately the dimer molecular weight. **(b)** Circular dichroism indicates a large gain of β-sheet structure when MipA is refolded into either detergent or DMPG liposomes compared to the unfolded protein in 8 M urea.

In addition to native OMPs, mutant OMPs and *de novo* designed transmembrane barrels[625] (TMBs) are also used in the following study to explore OMP-OMP interactions. Specifically, three of the OmpA mutants described in section 3.2.6, which have altered charged residues in their extracellular loops (OmpA-NP, no positive charges; OmpA-NN, no negative charges; OmpA-NC, no charged residues) and the minimal transmembrane barrel of tOmpA where the extracellular loops have been replaced with short, designed turns (OmpTrans3)[625] (**Fig. 5.1b**). In addition, *de novo* designed barrels TMB2.17LA and TMB2.3LA, which have had the four extracellular loops of OmpA grafted onto their equivalent positions in the designed barrel (i.e. eL1 between β1-2, eL2 between β3-4, eL3 between β5-6 and eL4 between β7-8)[625] (**Fig. 5.1c**).

OmpA-NP/NN/NC were shown in Section 3.2.6 to fold (>90% yield) into DMPG liposomes at 30 °C and 0.5 M urea, although at very different rates, while it has previously been shown that TMB2.17 can fold into DMPC membranes at the transition temperature (24 °C)[625]. Each of these proteins was tested in the conditions optimised here for folding, DMPG at 23.5 °C and 0.5 M urea, and shown by tryptophan fluorescence to fold rapidly and efficiently (**Fig 5.4**).

Therefore, DMPG at its transition temperature and in the presence of 0.5 M urea, facilitates near-universal, high yield folding for a wide range of OMPs of many sizes and complexities, including unnatural designed barrels, making it facile to generate proteoliposomes consisting of a mixture of OMPs at known concentrations and ratios and hence explore OMP-OMP interactions.

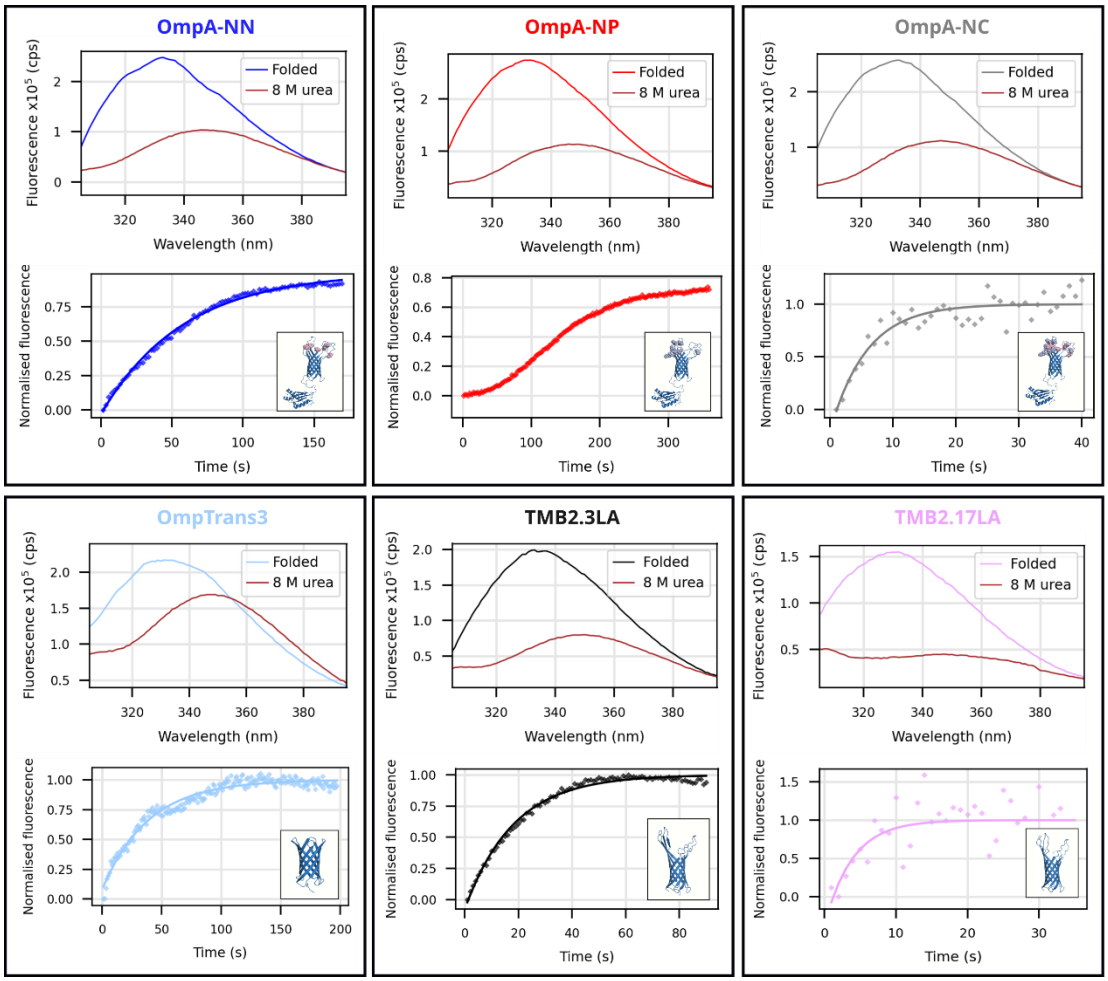


Figure 5.4: Non-native OMPs effectively fold into DMPG membranes. For each of the OMPs in **Fig. 5.1b-c** (excluding tOmpA, see **Fig. 5.2**) the folding kinetics (lower panels, followed by tryptophan fluorescence) and the folded vs unfolded (8 M urea) intrinsic fluorescence spectra are shown. All these proteins show clear, rapid folding kinetics and show the expected unfolded-folded transition in the spectra. OmpA mutant models from PDBs 1G90[290]/2MQE[284], others are Alphafold2 models.

5.2.2 The extracellular loops of OmpA enhance OmpT activity

OmpT, a pathologically relevant protease present at ~ 6% of the OM's protein content[273] has a straightforward and high throughput activity assay[105], as detailed in Chapter 4 (briefly, cleavage of a self-quenching fluorescent peptide relieves quenching and increases fluorescence). It is thus a tractable and biologically relevant protein to probe possible functional OMP-OMP interactions by assaying for changes in OmpT activity in the presence of a second OMP. DMPG-OmpT proteoliposomes were prepared with high LPRs to minimise non-specific interactions (6000:1, ~20 OmpT molecules per liposome), a second OMP was then folded into the liposomes to reach the desired ratio (typically 0-2.4x the number of OmpT molecules, unless otherwise stated) and OmpT's activity then determined.

OmpT proteoliposomes were assayed with BamA, BtuB and each of the abundant OMPs (for example **Fig 5.5a**). Calculating the initial rate demonstrated that the majority of the OMPs showed minimal concentration-dependent effects on the activity of OmpT (**Fig 5.5b**), although some small, apparently non-specific interactions are seen (for example, ~30% enhancement with BtuB and ~30% decrease with OmpX, both at 1:2.4 OmpT:OMP molar ratio). In contrast, OmpA and tOmpA both specifically enhance the activity of OmpT by ~2.2x (**Fig 5.5c**). This effect of OmpA (and tOmpA) is saturated at a ~ 1:1 protein ratio (up to at least a 1:16 OmpT:tOmpA ratio, **Fig. A17**), indicating the effect is mediated via formation of a specific complex between OmpT and the barrel domain of OmpA (as removing the soluble C-terminal domain (tOmpA) has no effect).

The physical requirements of this effect were probed by assaying the concentration dependence of OmpT activity on the OmpA extracellular loop charge mutants (**Fig 5.6a**). For OmpA-NN and OmpA-NC all the enhancement is lost, however OmpA-NP retains ~30% of the enhancement (which still saturates at ~ 1:1). The negatively charged residues are thus essential, indicating a key electrostatic interaction between the loops of OmpA and OmpT. However, negative charge is not sufficient as OmpA-NP is 3x less efficient at enhancing OmpT than wild-type OmpA, suggesting that the precise conformation of the loops, which could be disrupted by the seven mutations in OmpA-NP, is also critical. To test this, OmpT was assayed with TMB2.17LA and TMB2.3LA, the *de novo* designed barrels with the loops of OmpA. Neither of these showed any rate enhancement despite the presence of the negatively charged residues in the extracellular loops (**Fig. 5.6b**), demonstrating that the native

conformation of the loops, which will be disrupted in these unnatural TMBs, is important for enhancement. It was further shown that no rate enhancement is seen with OmpTrans3 (which has no extracellular loops, **Fig. 5.6b**), confirming the role of the loops in mediating the effect.

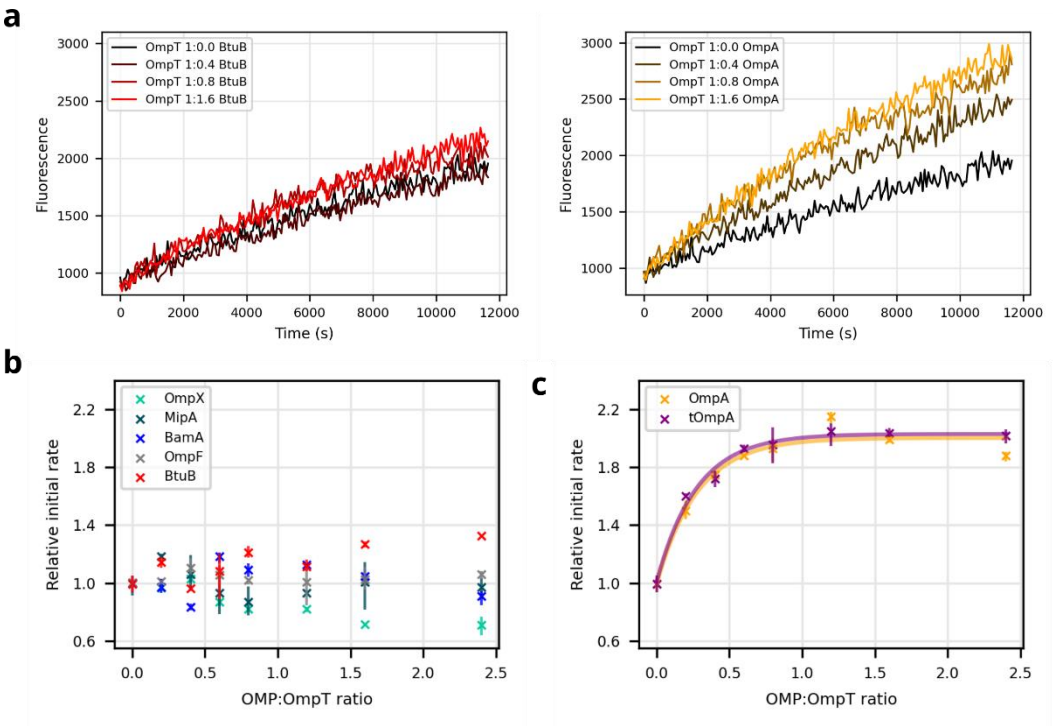


Figure 5.5: The transmembrane region of OmpA specifically enhances the activity of OmpT. (a) Sample raw OmpT enzyme kinetic traces for OmpT:BtuB and OmpT:OmpA at 1:0, 1:0.4, 1:0.8 and 1:1.6 ratios. OmpT activity dependence on the concentration of **(b)** OmpX, MipA, BamA, OmpF, and BtuB, or **(c)** OmpA and tOmpA (n ≥ 3). Rates are normalised to OmpT activity alone. (All error bars are data range, n ≥ 3)

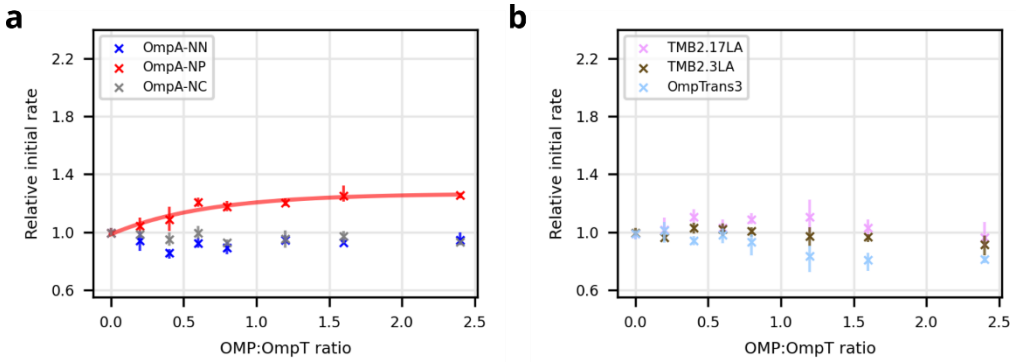


Figure 5.6: Both the precise conformation and the presence of extracellular loop negative charge are required for OmpA mediated OmpT rate enhancement. (a) Considering the extracellular loop charge mutants of OmpA shows that only OmpA-NP partially retains the ability to enhance OmpT activity. **(b)** Neither the transmembrane barrel of OmpA with loops, or the loops of OmpA alone, grafted onto the TMBs, are sufficient to enhance OmpT activity. (All error bars are data range, n ≥ 3)

It was shown in Chapter 4 that OmpT activity is increased ~25x (in DMPG) by the presence of Ra-LPS (Section 4.2.2). It was therefore considered whether OmpA and LPS compete or act additively (or synergistically) to increase OmpT activity. DMPG liposomes with 0.1 % or 0.5 % Ra-LPS (mol/mol) were synthesised and validated using silver-stain SDS-PAGE gel to show the incorporation of Ra-LPS (**Fig 5.7a**) and DLS to confirm their hydrodynamic radii (**Fig 5.7b**). Additionally, the folding of OmpA and OmpT into the DMPG-LPS liposomes was confirmed by tryptophan fluorescence (**Fig 5.7c**), as described in Section 5.2.1.

The activity of OmpT was then assayed in these liposomes at a range of OmpA concentrations (**Fig 5.7d**). In the presence of Ra-LPS, OmpT activity starts at a much higher rate as expected (Section 4.2.2), but an additional rate enhancement of 15% is seen, plateauing at ~1:1 as expected for the DMPG-0.5% Ra-LPS liposomes. Although due to the LPS-mediated enhancement OmpA's rate enhancing effect appears less, in absolute terms it is ~2x larger than in the absence of Ra-LPS, suggesting a degree of cooperativity. The enhancement and then decrease of OmpT activity in DMPG-0.1% Ra-LPS liposomes is likely due to increasing amounts of OmpA competing the Ra-LPS away from the OmpT and thus decreasing its activity (0.1% Ra-LPS is in 1:3 ratio of OmpT:Ra-LPS). Intriguingly, in the presence of Ra-LPS, OmpX is also able to enhance OmpT activity similarly to OmpA, suggesting a productive LPS mediated interaction (**Fig 5.7e**).

A substrate titration revealed that OmpT-OmpA (in DMPG) has double the affinity for the substrate compared to OmpT alone (179 μM +/- 31 μM vs 364 μM +/- 91.78 μM, 95 % CLs), while the V_{max} remains unchanged (**Fig. 5.8a**). Given that LPS had minimal effects on the enzyme's K_m (see **Fig. 4.8**), this also indicates that OmpA is acting in conjunction with LPS via a different mechanism. As shown in section 4.2.2, OmpT is basally active in DMPG and DMPS membranes, and OmpA can enhance OmpT activity in both these lipid contexts (**Fig. 5.5c, Fig. 5.8b**). However, OmpT shows no activity in DMPC or DLPC membranes, and the addition of OmpA fails to activate the enzyme in either case (**Fig. 5.8c**) (DMPC, DMPS and DLPC OMP folding validation in **Fig. A18**). This suggests that a basal level of activity, which can be mediated by negatively-charged lipids or short LPS, is required for OmpA rate-enhancement. Interestingly, in the presence of OMPs with significantly electronegative extracellular loops (BamA, MipA) OmpT is variably activated, with no detectable activity with BamA, but slight activation with MipA (**Fig. 5.8d**).

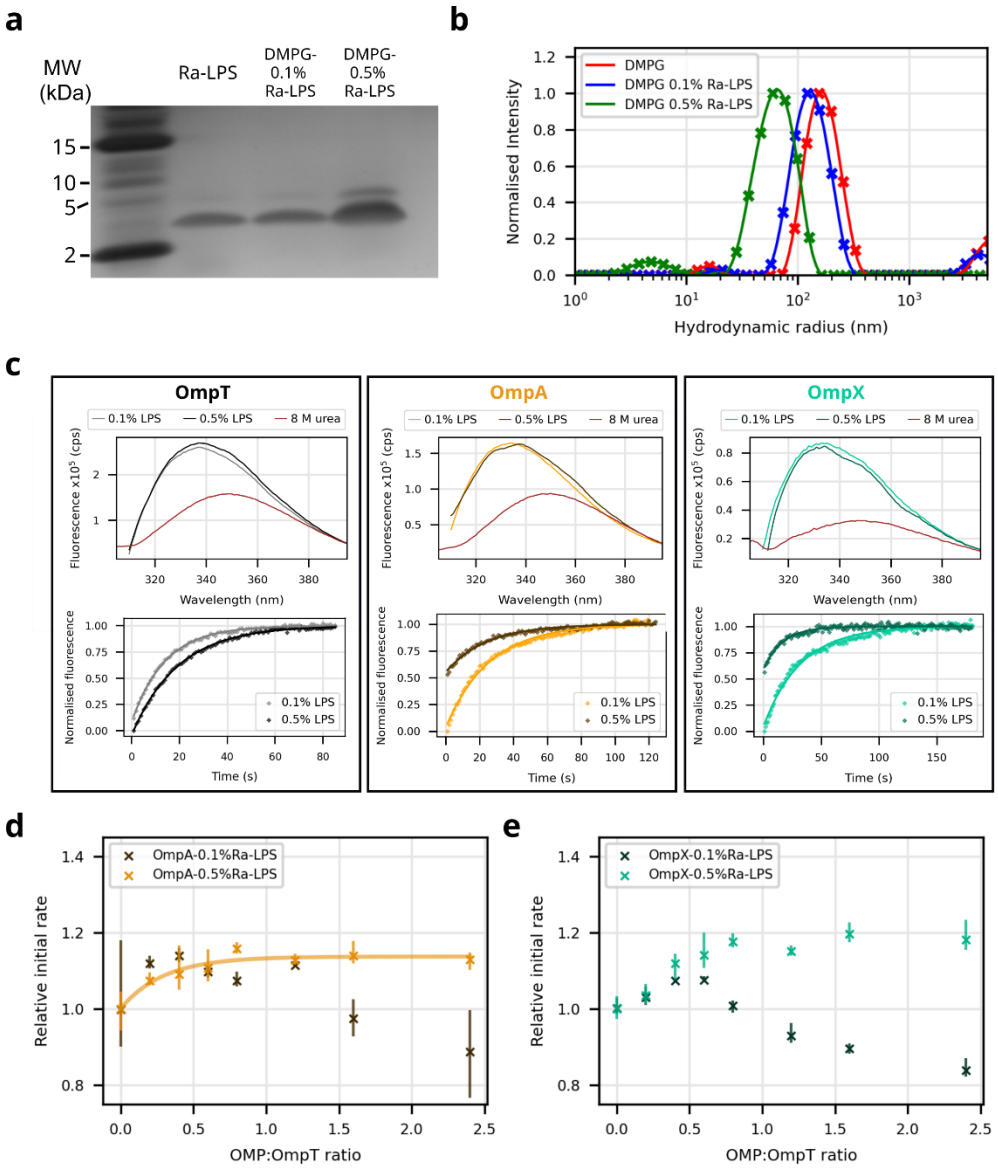


Figure 5.7: LPS and OmpA additively enhance OmpT activity. (a) Silver-stain SDS-PAGE gel of DMPG liposomes doped with Ra-LPS showing the incorporation of Ra-LPS. **(b)** DLS confirming the retention of vesicular shape of the LPS doped liposomes. **(c)** Unfolded and folded intrinsic fluorescence spectra and folding kinetics of OmpT (left), OmpA (centre) and OmpX (right) into (Ra-)LPS doped DMPG membranes. OmpT activity in DMPG-Ra-LPS liposomes with increasing concentrations of **(d)** OmpA and **(e)** OmpX. (All error bars are data range, n ≥ 3)

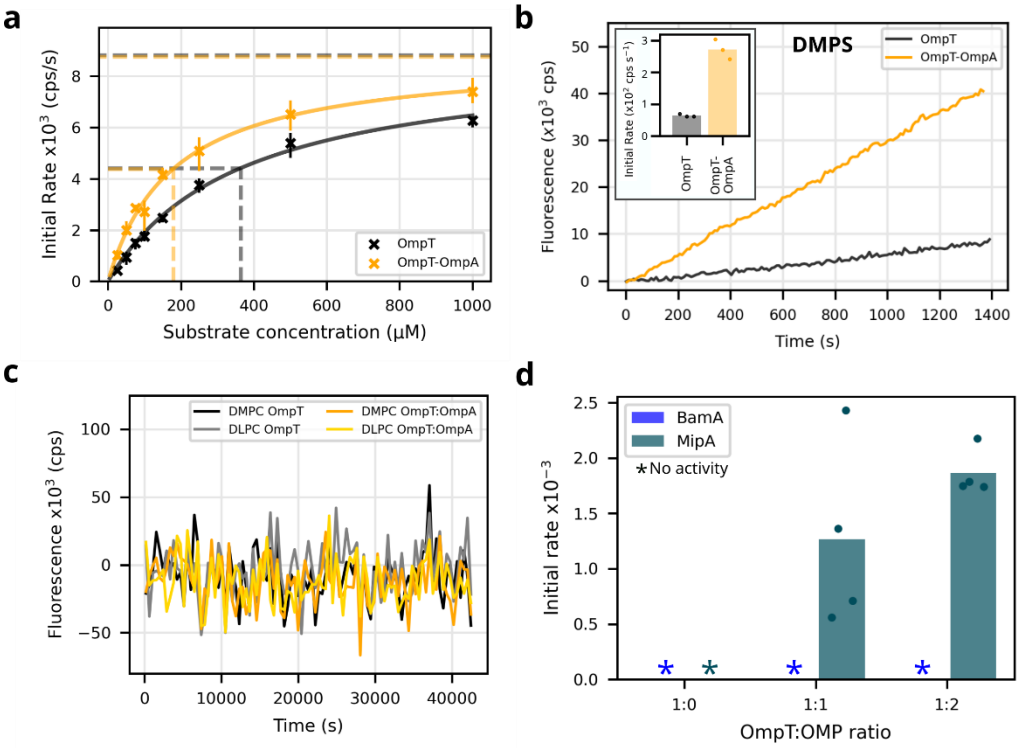


Figure 5.8: OmpA only enhances basally active OmpT. (a) Substrate titrations of OmpT and OmpT-OmpA (1:1 ratio) indicate they have the same V_{max} (upper dashed lines) but different K_m (lower dashed lines) (All error bars are data range, n=3) **(b)** OmpA enhances OmpT activity in DMPS membranes ~3x (n=3). **(c)** OmpA does not activate OmpT (1:1 ratio) in DMPC or DLPC membranes (n=4, single background subtracted example shown). **(d)** OmpT in DMPC membranes remains inactive in the presence of BamA but gains slight activity with MipA, both of which are extracellularly electronegative (n=4).

Given that MipA can activate OmpT in DMPC membranes (but did not enhance the activity in DMPG) and OmpA can enhance OmpT in DMPG membranes (but not activate in DMPC), their possible complementary effects were tested using a tripartite activity assay. OmpT-OmpA (1:1 ratio) was folded into DMPG liposomes and then a third OMP was added at variable concentrations. By exploiting the variable change in enzymatic activity of OmpT by titration of OmpA, it is possible to assess the effects of the third OMP on the interaction of OmpT-OmpA. While understanding the precise basis of observed effects in a tripartite system is complicated, it gives insight into how different OMPs might interact together in the OM. The effects of MipA and the other abundant OMPs (OmpF, OmpX) and BtuB on OmpT activity were measured (**Fig. 5.9**).

The effects with MipA are particularly striking, with MipA consistently and

significantly enhancing the rate of OmpT activity in a concentration-dependent manner. MipA alone does not alter OmpT activity (**Fig. 5.5b**), indicating a tripartite interaction between OmpT-tOmpA and the additional MipA, although the linear gradient implies weak binding. As noted above, MipA can basally activate OmpT in DMPC membranes (**Fig. 5.8d**), also in a (partially) concentration dependent manner, hinting at what it may be doing in the DMPG OmpT-tOmpA-MipA system, where tOmpA may be priming OmpT for MipA rate enhancements. Unlike MipA, OmpX reduces the activity of OmpT to a basal level, indicating a disruption of the OmpT-tOmpA complex, while both BtuB and OmpF show rate enhancements greater than that seen in the absence of OmpA, but smaller than MipA. These data yield an insight into the interaction network within the OM, where multiple OMPs can act in concert to modulate function in a manner that depends precisely on the identity of the OMP and the membrane environment.

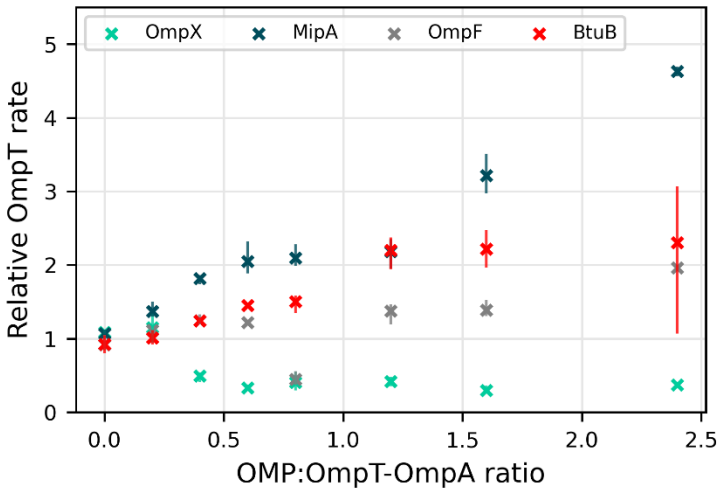


Figure 5.9: Tripartite OmpT-OmpA activity assay. Activity of OmpT in DMPG OmpT:OmpA (1:1) proteoliposomes with the addition of 0-2.4x OmpT concentration of a third OMP.

5.2.3 Understanding the OmpT-OmpA complex

It was shown above that the activity enhancement of OmpT by OmpA saturates at ~1:1 and is dependent on the transmembrane domain of OmpA and on negative charge in the extracellular loops. To visualise the potential OmpT-tOmpA dimer, the structure of the complex was predicted by Alphafold2 (AF2) (**Fig. 5.10a**), yielding a medium-confidence complex (inter-chain adjacent residue (Cα-Cα 1.2 nm cutoff) PAE ~15, **Fig. 5.10b**). Encouragingly this complex has electrostatically mediated extracellular loop interactions with two negative residues on OmpA (D149, D116) predicted to interact with positive residues on OmpT (R136, K171, R173) (**Fig. 5.10a**, inset), in agreement with the biochemical data (**Fig. 5.5a**). The AF2 predicted interface was consistent across a range of homologues from pathogenic organisms (**Fig. 5.10c**), not unexpected given the high-level of conservation between these sequences. Initial cryoEM imaging suggests that a stable detergent solubilised complex can be formed (**Fig. 5.10d**), although the data quality was not good enough for 3D reconstruction.

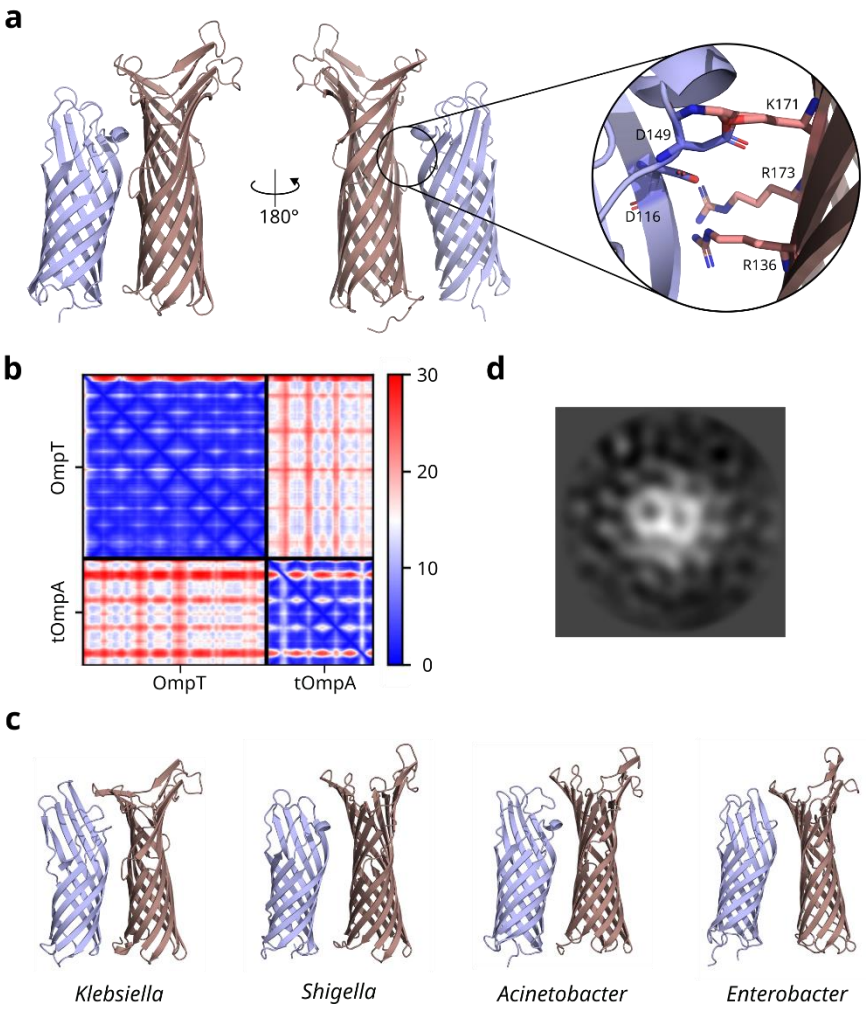


Figure 5.10: OmpT-OmpA Alphafold2 prediction. (a) Alphafold2 prediction of OmpT (blue) and tOmpA (brown) complex. Inset: the interaction surface is dominated by the extracellular loops and includes an electrostatic patch. **(b)** PAE (predicted alignment error) between OmpT and tOmpA, low values (blue) indicate high confidence. **(c)** Preliminary top-down 2D class of OmpT-tOmpA dimer in LDAO. **(d)** Alphafold2 models of OmpT-tOmpA complex from key pathogenic species.

To further validate the structural modelling and biochemical inferences, OMP-OMP crosslinking was performed. Proteoliposomes were prepared with an LPR of 1000:1 (mol/mol), buffer exchanged (to remove the 0.5 M urea used for folding and Tris buffer) with four rounds of pelleting via ultracentrifugation (see Section 2.11). Samples were then crosslinked with 0.1 % (vol/vol) glutaraldehyde (a general primary amine crosslinker) for 5 or 15 minutes. (**Fig. 5.11a**). The method was validated by assessing the buffer-exchanged proteoliposomes via DLS (**Fig. 5.11b**) and cold SDS-PAGE band-shift for tOmpA and OmpX to ensure that OMPs remain folded (**Fig. 5.11c**). The monomer fraction compared to an uncrosslinked sample was quantified by densitometry of SDS-PAGE gels (**Fig. 5.12a-d**). As OmpA and OmpT have similar masses (35 kDa and 33 KDa), tOmpA (18 kDa) was used in place of OmpA.

Initial monomer controls with tOmpA, OmpT, OmpX and OmpF indicated high levels of crosslinking (measured by loss of monomer bands) for OmpF (~100% after 15 minutes) as expected given its trimerization and array formation) and low levels for OmpX and tOmpA (~10 % after 15 minutes), with intermediate crosslinking for OmpT (~40% after 15 minutes, **Fig. 5.12e**). OmpT in proteoliposomes of OmpF-OmpT and OmpX-OmpT has similar crosslinking efficiency to the monomer, as does OmpF, although OmpX is slightly increased (~35%) (**Fig. 5.12f-h**). In contrast, the tOmpA-OmpT proteoliposomes show increased crosslinking for both OmpT and tOmpA, indicating the formation of oligomers, particularly clear after 15 minutes where both tOmpA and OmpT are about half as likely to remain monomers, compared to the individual proteins (**Fig. 5.12h**). Together, this is indicative of an oligomer formed by tOmpA-OmpT facilitating higher crosslinking of both species, which is absent in OmpX-OmpT and OmpF-OmpT proteoliposomes, demonstrating its specificity.

Together, these data demonstrate a specific interaction between OmpT and the transmembrane domain of OmpA, not seen with any other tested OMP, that enhances OmpT activity via its extracellular loops by increasing substrate affinity, dependent on both the precise conformation of OmpA's loops and their negatively charged residues, and that can act additively with LPS or other OMPs to further enhance the activity. It also highlights the multivalent functional interactions occurring within OMP clusters.

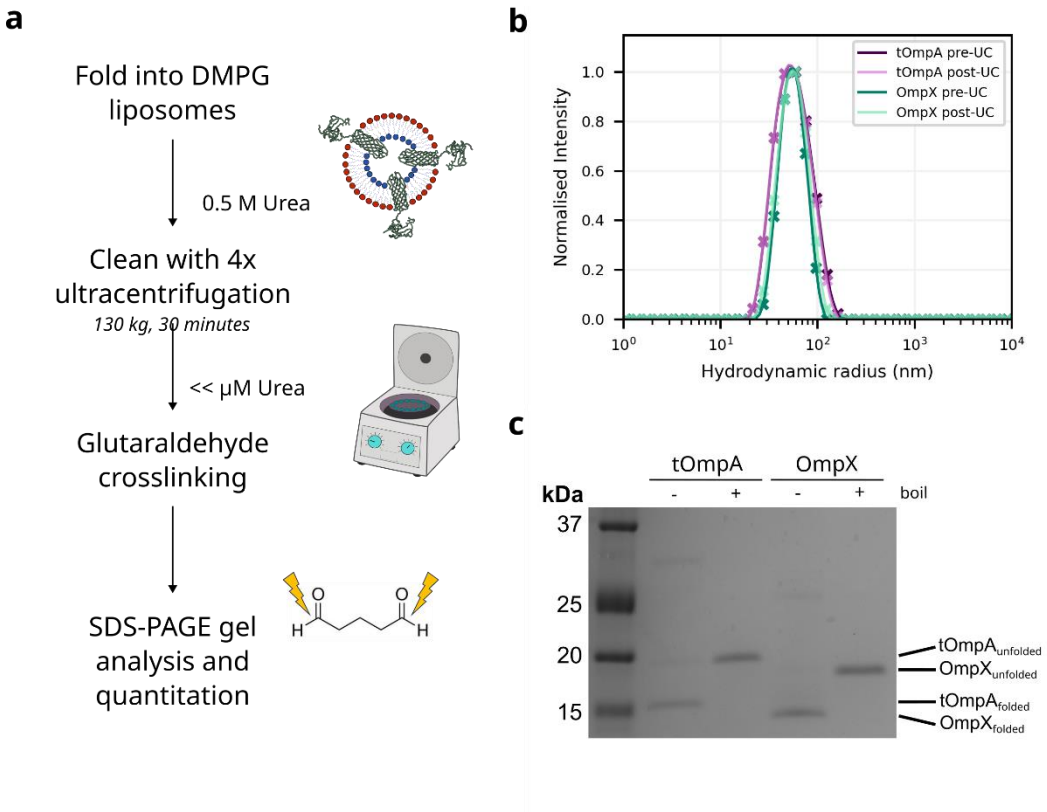


Figure 5.11: Crosslinking protocol and validation. (a) Overview of crosslinking procedure. **(b)** DLS of proteoliposomes before and after ultracentrifugation buffer exchange. **(c)** Cold SDS-PAGE of buffer exchanged tOmpA and OmpX proteoliposomes samples with and without boiling.

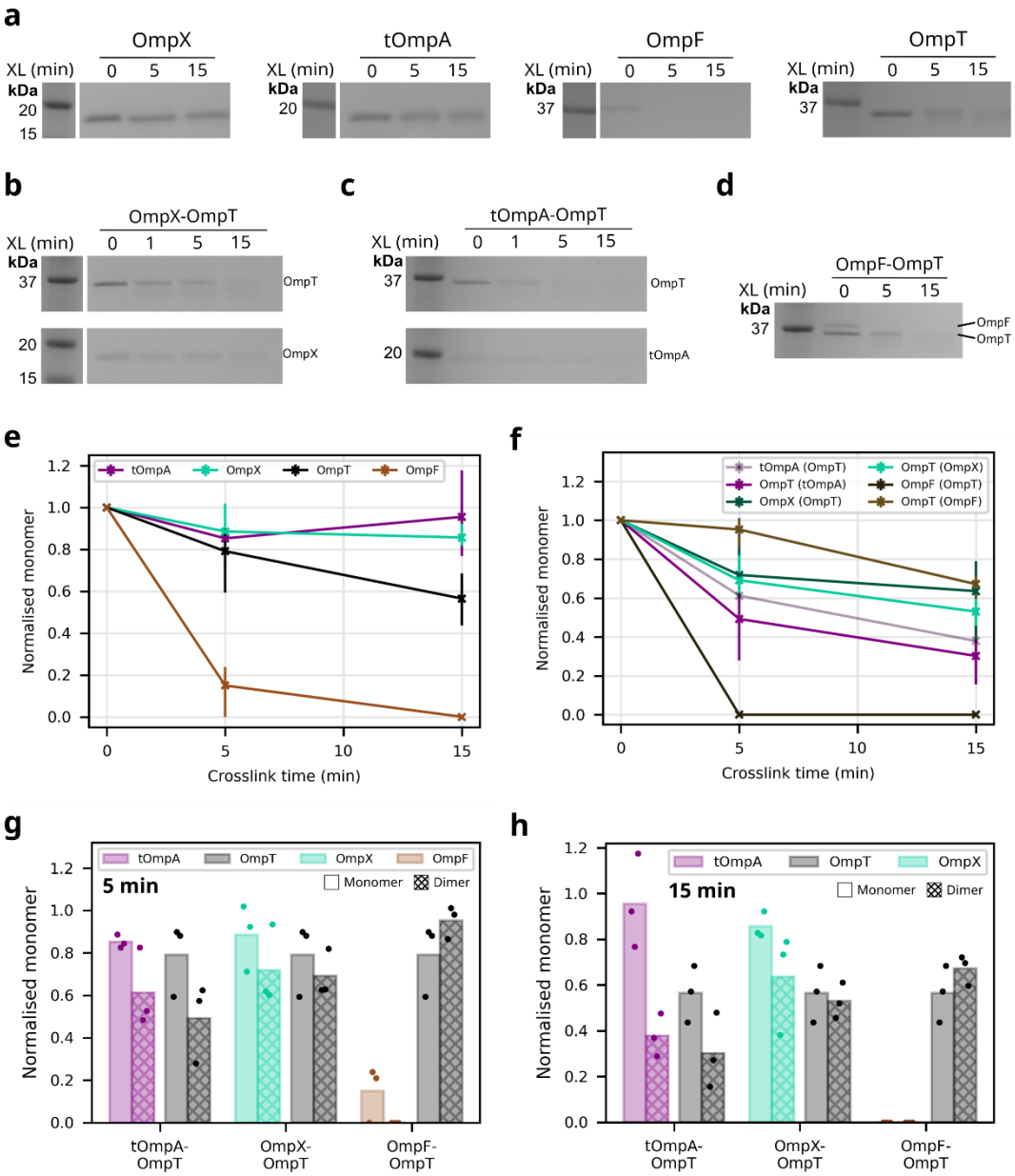


Figure 5.12: tOmpA-OmpT crosslink with greater efficiency then OmpX-OmpT and OmpF-OmpT. (a) Sample gels of crosslinking in proteoliposomes with only one protein species present (c.f. **Fig. 5.20**) and crosslinking in proteoliposomes with two protein species: **(b)** OmpX-OmpT, **(c)** tOmpA-OmpT, and **(d)** OmpF-OmpT. **(e)** SDS-PAGE gel band intensity normalised to the uncrosslinked sample for a single species of OMP in DMPG membranes and **(f)** two protein species (n=3, error bars are data range). **(g)** Comparison of monomer fraction retention in proteoliposomes of one (solid bars) or two (hashed bars) protein species after 5 minutes crosslinking or **(h)** 15 minutes crosslinking (n=3).

5.2.4 OmpA reduces PagP activity

OmpA is the single most abundant OMP in the OM of *E. coli* and given the enhancement in OmpT's activity in the presence of OmpA, it was next considered whether other OMPs have evolved modulatory activity based on interaction with OmpA. As noted above, the enzyme PagP was the only OMP tested that failed to fold into DMPG membranes. Previous reports suggest that while folding into liposomes is possible[270], PagP may become folding incompetent at low urea concentrations[446]. PagP was found to fold into DMPG liposomes at 6 M urea, confirmed by intrinsic fluorescence peak shift (**Fig. 5.13a**), circular dichroism (**Fig. 5.13b**) and gain of enzymatic activity (see **Fig. 5.14**, detailed below), although folding kinetics by intrinsic fluorescence could not be measured.

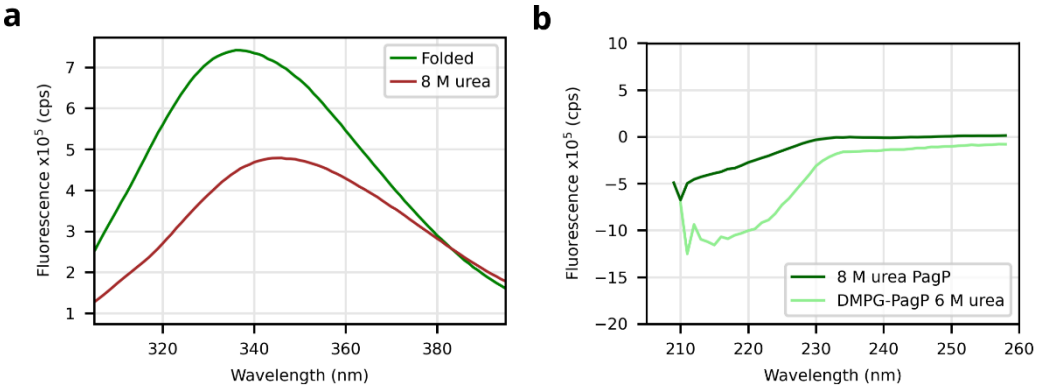


Figure 5.13: PagP folds into DMPG liposomes in the presence of 6 M urea. **(a)** Intrinsic fluorescence **(b)** and background subtracted circular dichroism spectra of unfolded (8 M urea) and DMPG-folded PagP (in the presence of 6 M urea). Circular dichroism spectra are truncated at 209 nm due to absorbance in 6 M urea.

A colorimetric activity assay for PagP activity refolded in detergent has previously been reported[446], based on cleavage of the palmitate group in the synthetic substrate palmitate-4-nitrophenol (pNP, **Fig. 5.14a**), releasing free, yellow coloured NP, allowing the reaction progress to be monitored via absorbance. While this substrate can be incorporated into liposomes during their synthesis, the reaction would then proceed uncontrollably upon the initialisation of PagP folding. In the absence of reversible PagP inhibitors, a way of rapidly delivering the pNP substrate into liposome membranes to begin the reaction is required.

As shown in Section 3.2.1, MβCD solubilised lipids can be exchanged with lipids in membranes, but historically cyclodextrins have been used at lower concentrations to deliver more hydrophobic molecules like cholesterol into bilayers[640],

suggesting its possible utility to deliver pNP to liposomes. Initially it was shown that at the lipid:MβCD ratios to be used (typically 1:2, 1:4 tested, mol/mol), the liposomes remain intact by DLS (**Fig. 5.14b**). Next it was shown that the equilibrium phase transition temperature of DMPG membranes incubated with either MβCD or MβCD-pNP increased in the presence of pNP (using fluorescent reporter laurdan, see Section 3.2.2), indicating substrate delivery and subsequent increase in the T_m (**Fig. 5.14c**). Monitoring the laurdan fluorescence (at both its fluorescence peaks of 440 nm and 490 nm) before and after addition MβCD-pNP showed a change in the T_m upon addition of MβCD-pNP, but not MβCD only, that fully occurred in the reaction deadtime (5 s) (**Fig. 5.15d**). Finally, it was shown that MβCD-pNP can functionally deliver substrate to PagP refolded in detergent (**Fig. 5.14e**) or into DMPG liposomes (**Fig. 5.14f**), yielding expected enzyme kinetic absorbance traces. The dependence on the enzyme (in liposomes) for the absorbance change was demonstrated by reducing the concentration of enzyme by 1.5x and observing a concomitant reduction in the initial rate (~1.6x reduced) and showing that in the absence of enzyme no absorbance increase was seen (**Fig. 5.14g**). Thus, MβCD facilitates a rapid delivery of PagP substrate into liposome bilayers, enabling enzyme kinetics to be measured.

Figure 5.14: Liposomal PagP activity assay development and validation. (a) Structure of pNP. **(b)** DLS shows that liposomes with 1:2 or 1:4 (mol/mol) lipid:MβCD ratios remain intact. **(c)** Equilibrium transition temperature measurements using the laurdan reporter showing the general polarisation (GP, upper) and the first differential of the GP, lower. Transition temperatures for each liposome are determined using the differential peak. **(d)** Time-resolved transition temperature measurements upon the addition of MβCD (left) or MβCD-pNP (right), measured at the gel and fluid phase fluorescence peaks of the laurdan reporter (440 and 490 nm respectively). MβCD-pNP can deliver pNP to **(e)** DDM refolded PagP and **(f)** DMPG refolded PagP efficiently to begin the reaction. **(g)** Substrate cleavage is not observed without PagP (black) and occurs in a PagP concentration dependent manner.

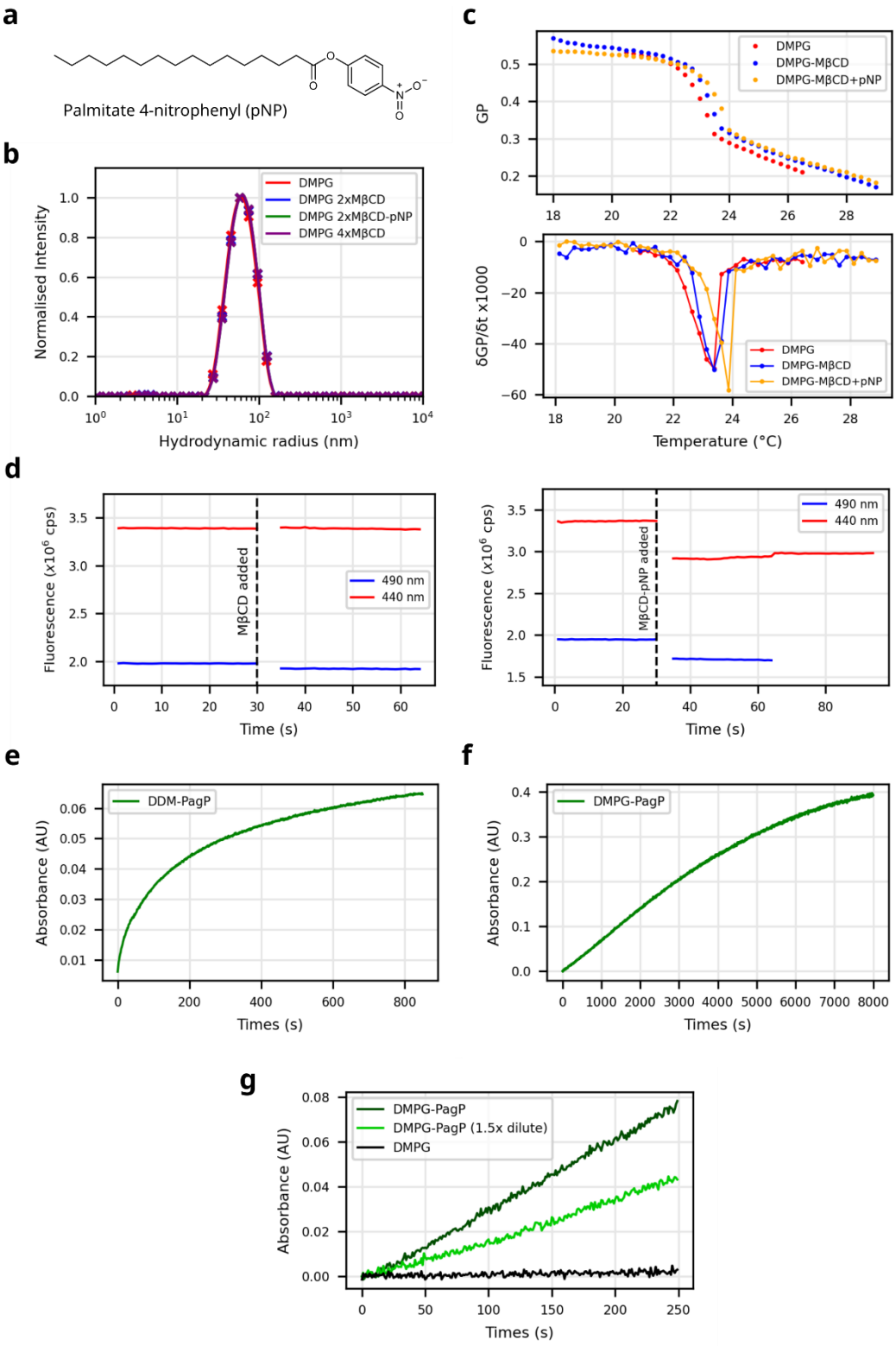


Figure 5.14: (Legend overleaf)

PagP was folded into DMPG liposomes in 6 M urea due to its unique folding requirements. Prior to folding the second OMP, the proteoliposomes were diluted to 2 M urea, which precludes folding of some of the more marginally stable OMPs like BtuB. It was shown by tryptophan fluorescence that tOmpA, OmpA and OmpX readily fold into DMPG liposomes in 2 M urea (**Fig. 5.15a**, c.f. OmpA in **Fig. 3.7b**). Thus, DMPG proteoliposomes with OmpX, OmpA or tOmpA were generated in a 1:1 molar ratio to PagP, and their PagP activity assayed (**Fig 5.15b**, single replicate shown). A striking difference is observed between the initial rates of PagP activity in the presence of OmpX, which has no effect on the activity, and the rate with OmpA and tOmpA, both of which reduce the activity (to ~35% and ~45% respectively, **Fig. 5.15c**). The effect of OmpA on PagP can be effectively demonstrated by folding OmpA into the liposomes mid-reaction and observing the concomitant decrease in PagP rate of activity (**Fig 5.15d**). Thus, OmpA is able to regulate the activity of a different enzyme in the OM, suggesting that several OMPs have adapted to use to help regulate their function.

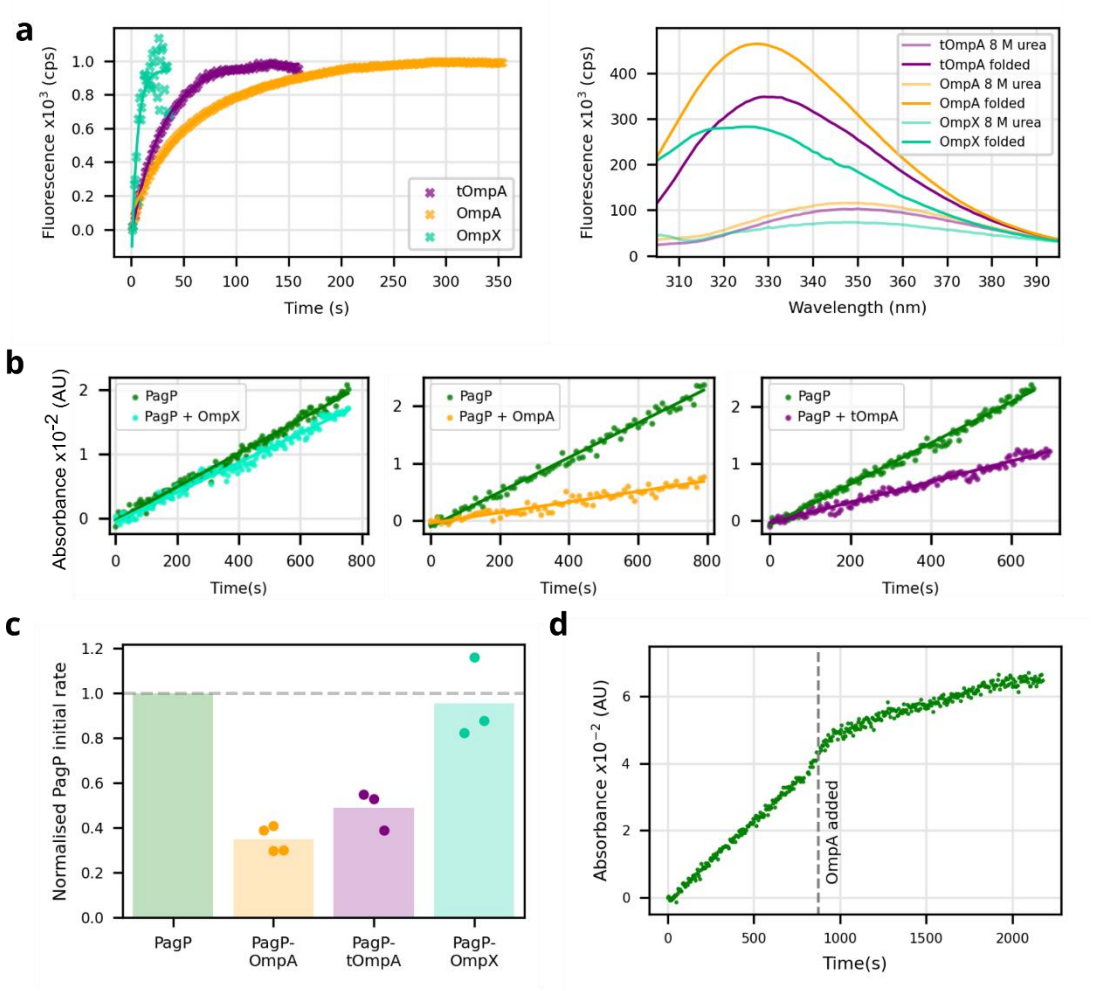


Figure 5.15: The transmembrane region of OmpA inhibits PagP activity. (a) OmpA, tOmpA and OmpX all fold efficiently into DMPG liposomes in 2 M urea. **(b)** Example raw activity traces for PagP alone compared to PagP in a 1:1 molar ratio with OmpX, OmpA or tOmpA. **(c)** Normalised initial rates of PagP activity alone or in the presence of OmpA, tOmpA or OmpX, showing rate inhibition by OmpA and tOmpA. **(d)** Folding OmpA into DMPG-PagP liposomes while measuring PagP activity induces a ~60% reduction in activity upon the addition of OmpA. (Volume increased 2% with OmpA addition).

5.2.5 OmpLA activity is not affected by the presence of other OMPs

The third enzyme of the *E. coli* OM, OmpLA (also known as PldA), is a phospholipase which requires divalent cations for activity[615]. OmpLA intrinsically refolded into DMPG liposomes (see Section 5.2.1) is stable in the absence of divalent cations for >24 hours (**Fig. 5.16a**), but upon addition of $CaCl_2$ the liposomes are rapidly degraded, (as measured by DLS where specific timepoints are quenched with a 25x molar excess of EDTA, **Fig. 5.16b**). The collapse of the liposomal structure into larger aggregates can also be monitored via light scattering at 510 nm, with a repeatable characteristic curve (**Fig. 5.16c**). Most of the scattering changes occur within 60 s, consistent with the appearance of large aggregates in the DLS by this time point. OmpLA proteoliposomes, supplemented with a 1:1 (molar) ratio of either tOmpA or OmpX show identical apparent activity traces (**Fig. 5.16d**), suggesting that no protein-protein interactions that modulate OmpLA function are occurring.

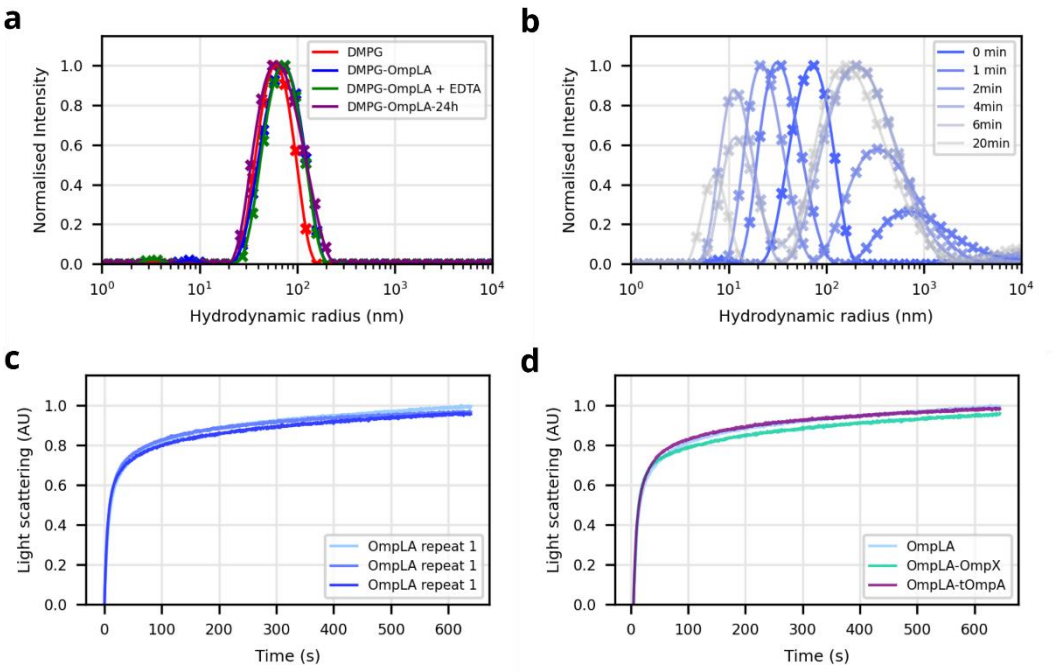


Figure 5.16: OmpLA activation initiates rapid liposomal collapse. (a) OmpLA refolded into DMPG liposomes is stable for >24 hours in the absence of divalent cations. **(b)** Adding of $CaCl_2$ to activate OmpLA leads to rapid collapse of the liposomes (reaction time-points quenched with excess EDTA). **(c)** Liposomal collapse and large aggregate formation can be followed by measuring sample OD over time (510 nm), yielding repeatable data consistent with the DLS (b). **(d)** Following OmpLA activity by OD_{510} in the presence of a 1:1 (molar) ratio of OmpX or tOmpA show no significant activity difference between samples.

While indicative, assaying activity by monitoring OD_{510} is indirect, and a more explicit activity measurement is desirable, either directly measuring the loss of substrate (i.e. liposomes, as in DLS), or the appearance of product, in this case free fatty acid (FFA) and lyso-PG. Methods to identify specific FFAs have previously been described based on their ability to be bound by bovine serum albumin[641] (BSA), which has characterized *in vitro* (and natively *in vivo*) FFA binding via six major sites[642,643] (**Fig. 5.17a**). Alternatively, ANS (8-anilinonaphthalene-1-sulfonic acid) can bind to these sites on BSA and strongly fluoresce[644]. Titrating oleic acid (OA) to pre-complexed BSA-ANS leads to a concentration dependent reduction in fluorescence as OA competes with ANS binding to BSA (**Fig. 5.17b**), indicating that BSA-ANS can be used as a measure of FFA. Incubating BSA-ANS with DMPG-OmpLA proteoliposomes and then adding $CaCl_2$ to initialise the reaction yields repeatable activity traces (**Fig. 5.17c**) and a reduction in fluorescence pre- and post-reaction (**Fig 5.17d**).

Finally, to increase throughput, this assay was optimised for a plate-reader (final parameters: 9 μM ANS, 0.5 μM BSA, 150 μM DMPG, 0.025 μM OmpLA; 1 mM $CaCl_2$ added to start the reaction; ~10 s reaction deadtime), which resulted in similar looking data (**Fig. 5.18a**, c.f. **Fig. 5.17c**). OmpLA activity was measured in proteoliposomes including a 1:1 molar ratio of each of the five measured abundant OMPs or BamA or BtuB (**Fig. 5.18b**). Calculating the initial rates, relative to the rate of OmpLA alone, shows only small difference (while OmpA, tOmpA and OmpT all statistically significantly enhance the rate (p-values: 0.013, 0.023, 0.012, respectively) the magnitude of the differences remains very small) upon the addition of any other OMP (**Fig. 5.18c**), further confirmed by determining the half-maximum time for each curve (T_{50}) (**Fig. 5.18d**). Thus, unlike OmpT and PagP, OmpLA's activity is not substantially modulated by any of the other OMPs tested here, possibly because it forms a homodimer as its active conformation[615], precluding other interactions.

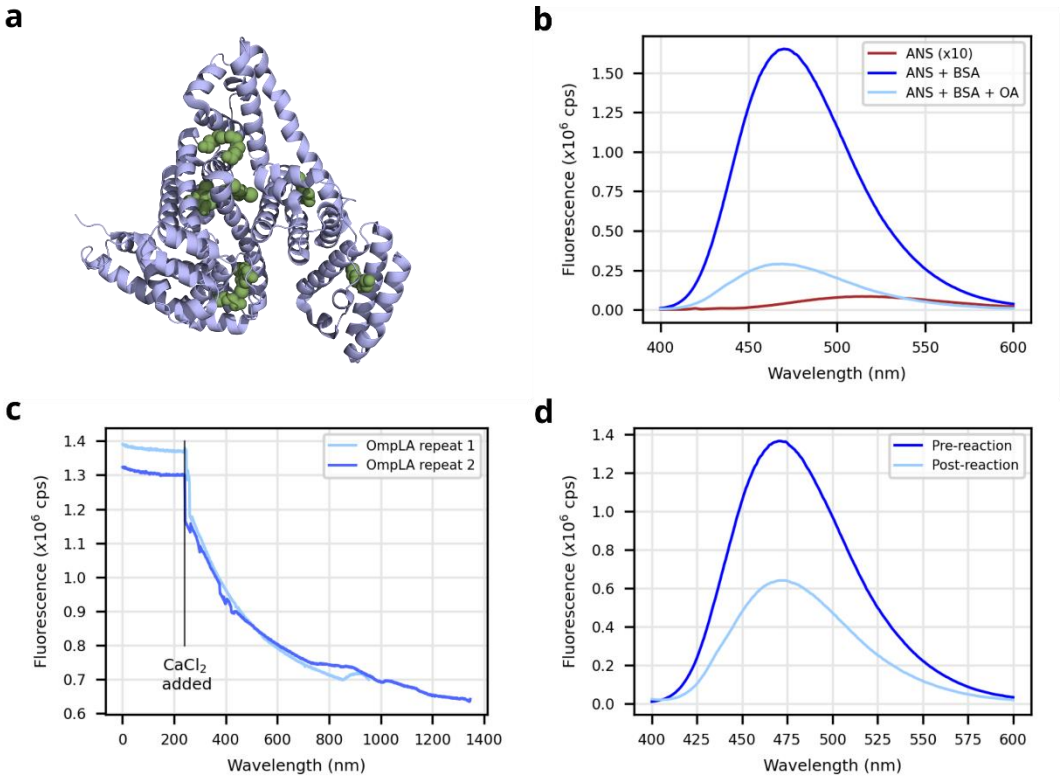


Figure 5.17: OmpLA activity can be measured by detecting free fatty acid concentrations with BSA-ANS. (a) Structure of human serine albumin (a homolog of BSA) (blue) bound to six fatty acids (green) (PDB: 1BJ5[642]). **(b)** ANS becomes fluorescent at 470 nm upon the addition of BSA (brown compared to dark blue), and this fluorescence peak can be reduced by adding free oleic acid (OA) (light blue). **(c)** Monitoring fluorescence over time upon addition of $CaCl_2$ shows measurable enzyme kinetics, and **(d)** spectra before and after the reaction show a clear reduction in BSA-ANS fluorescence (shown after 25 minutes). See also Section 2.8.2.

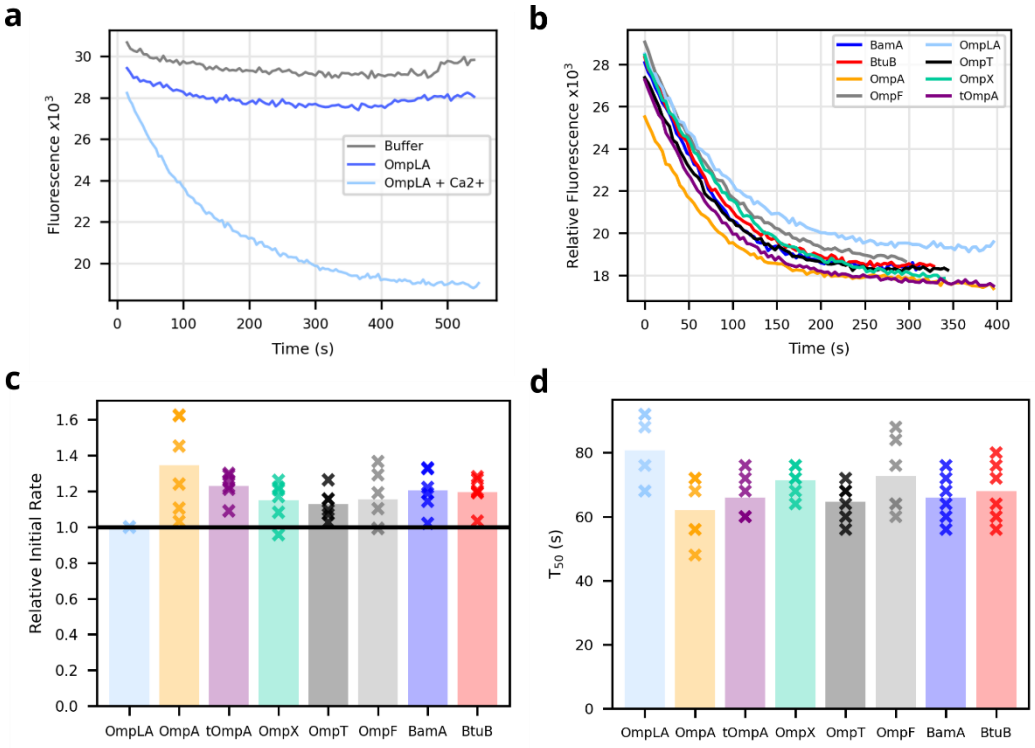


Figure 5.18: OmpLA activity is not modulated by any tested OMP. (a) Plate-reader optimised OmpLA activity assay by BSA-ANS fluorescence detection of free fatty acid (activity traces look similar to **Fig. 5.17c**). **(b)** Example OmpLA activity traces with a molar ratio of 1:1 of each OMP as indicated. **(c)** Initial rates (normalised to OmpLA alone, black line) of each OmpLA:OMP mixture indicates no significant differences in activity. **(d)** Considering the T_{50} (time to half maximum amplitude) also indicates no differences in activity.

5.2.6 LPS mildly encourages OMP homo-oligomerisation

LPS has been implicated in OMP-OMP interactions[489,491], and to better understand the role of LPS in generating and sustaining broader membrane architectural features, the clustering of multiple copies of a single OMP was considered in DMPG or DMPG with 1% (mol/mol) Ra-LPS by glutaraldehyde crosslinking. The abundant OMPs (OmpA, OmpF, OmpX, OmpT, MipA) and tOmpA were considered (the soluble domain of OmpA is known to dimerize[289]). Samples were crosslinked with 0.1 % (vol/vol) glutaraldehyde for 5 or 15 minutes, and the monomer fraction compared to an uncrosslinked sample by densitometry of SDS-PAGE gels (**Fig. 5.19**, see **Fig. 5.11** for method and validation, **Fig. A19** for full length gels).

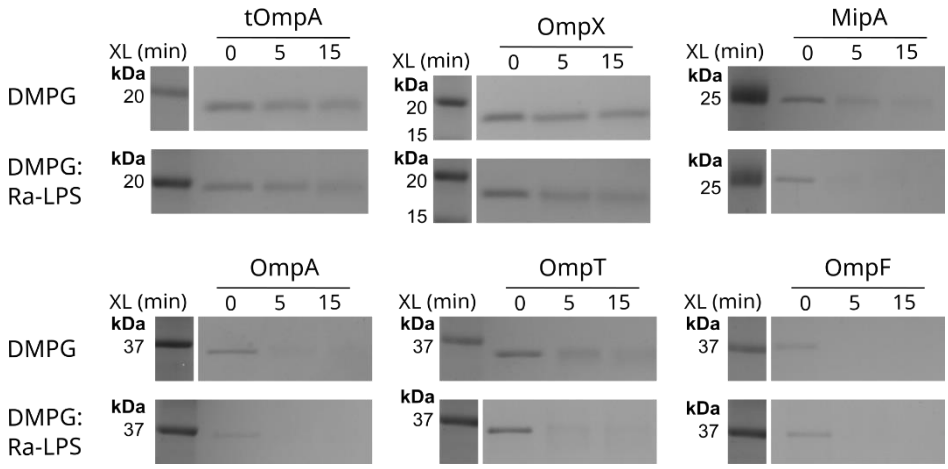


Figure 5.19: Single OMP crosslinking in the presence and absence of Ra-LPS. All samples were crosslinked with 0.1 % (vol/vol) glutaraldehyde for the indicated time and boiled prior to gel analysis. The intensity of OmpA and OmpF uncrosslinked bands is reduced as OmpA showed significant dimerization and OmpF oligomerisation without crosslinker (see **Fig. A19**).

These gel data, combined with densitometry quantification (**Fig. 5.20a**,**b**), shows expected patterns, including a general increased crosslinking efficiency with time and greater crosslinking for OMPs known to self-associate (OmpF which forms ordered arrays of trimers[627] and OmpA which forms dimers). Comparing the crosslinking of OMPs between DMPG and DMPG:Ra-LPS membranes (**Fig. 5.20c**,**d**) suggests that there is a small trend towards greater crosslinking efficiency in the presence of LPS, but only for tOmpA at 15 minutes does this comparison tend to significance (p-value: 0.05). Taking the difference between each data-point and the average crosslinking (of both DMPG and DMPG:Ra-LPS for that OMP), allows all the data at each time-point to be combined and compared (**Fig. 5.20e**,**f**). This comparison yields a highly significant difference between DMPG and DMPG:Ra-LPS at 15 minutes (p-value: 0.002) and a trending difference at 5 minutes (p-value: 0.061), with greater monomer loss in the presence of Ra-LPS. Together, this indicates that these OMPs are generally more likely to cluster in the presence of Ra-LPS compared to its absence, suggesting that small amounts of LPS may be important for mediating OMP island formation, consistent with recent reports[491].

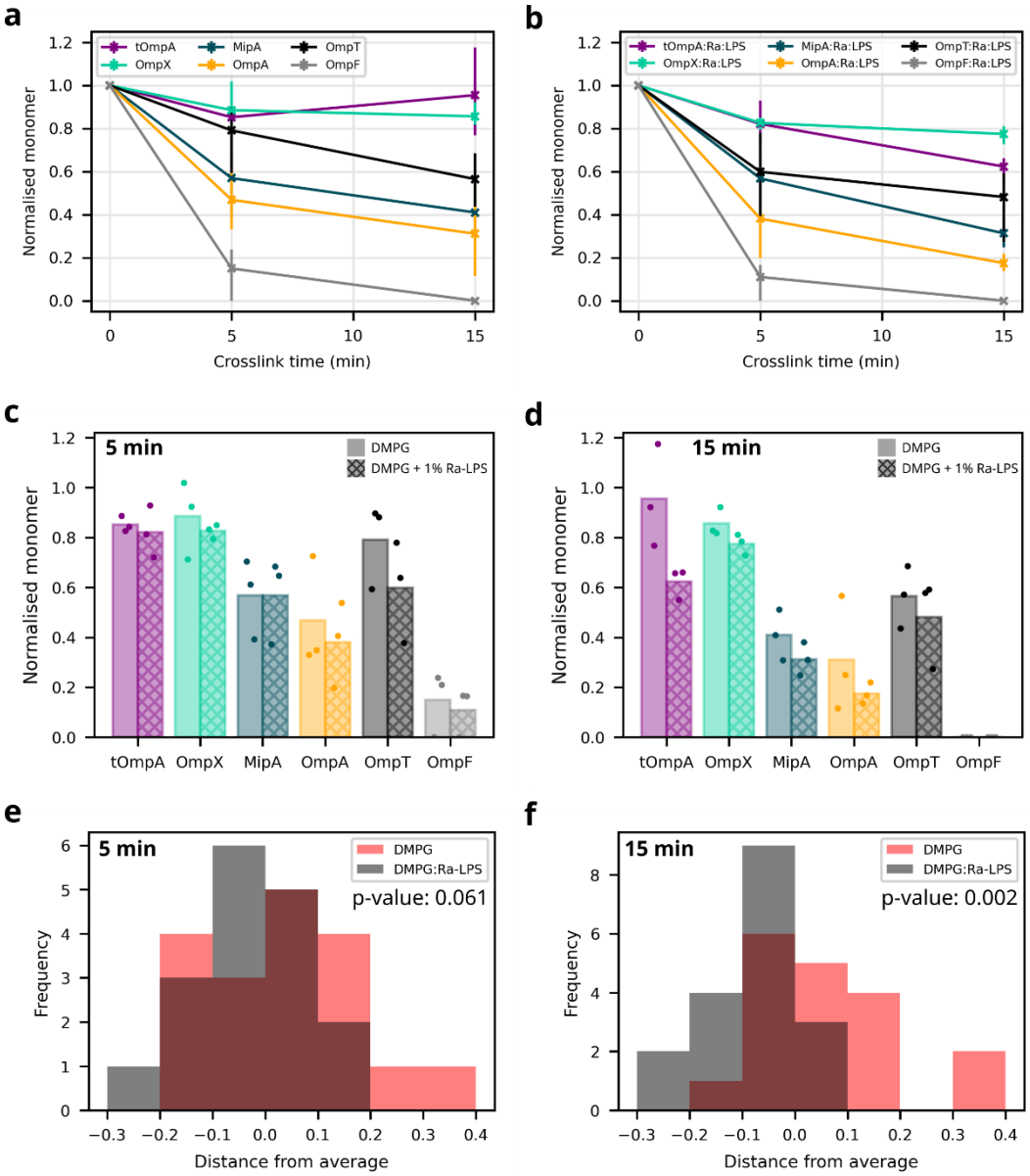


Figure 5.20: The presence of LPS mildly encourages clustering of OMPs in DMPG. (a) SDS-PAGE gel band intensity normalised to the uncrosslinked sample for OMPs in DMPG membranes and **(b)** DMPG with 1 % mol/mol Ra-LPS (n=3, error bars are data range). **(c)** Comparison of monomer fraction retention in the absence and presence of Ra-LPS after crosslinking for 5 minutes or **(d)** 15 minutes. **(e)** Comparing all the data by finding the difference between each data-point and the average crosslinking (of both DMPG and DMPG:Ra-LPS) for that OMP at the 5 minute and (**f)** 15 minute times (P-values (T-test) for the distribution differences are 0.061 and 0.002, respectively).

5.2.7 Systematic prediction of OMP-OMP complexes

Alphafold2 (AF2)[591] allows facile prediction of complexes with a good in-built metric of their quality and probability[645,646], and OMPs are known to be very well predicted[647]. Considering the modulatory role of OmpA in the activity of OmpT and PagP, and the extensive interactions that must occur natively between highly abundant OMPs and all other proteins in the OM, AF2 dimers were predicted for the seven most abundant OMPs (those estimated to comprise > 2% of the OM's protein content[273]; OmpA, OmpC, OmpF, OmpX, OmpT, MipA, Tsx) with all other *E. coli* OMPs. Non-monomeric OMPs were predicted with enough copies to make an intact barrel (i.e. three copies of the trimeric TolC). The soluble, C-terminal domain of OmpA appeared to disrupt some complex predictions, and thus tOmpA was also considered. (Predictions with trimeric OmpF or OmpC tended to be of poor quality compared to their monomeric prediction, and thus data for their monomeric predictions are presented here).

Following quality filtering (pLLDT >80), the average adjacent-residue (Cα-Cα cutoff of 1.2 nm) inter-chain PAE was calculated for each prediction (i.e., between all geometrically adjacent residues in distinct chains) (**Fig. 5.21**, ordered left-right by OMPs with increasing strand number). As well as identifying known interactions, such as the trimeric porins, this highlights multiple unexpected, medium-high probability interactions (PAE ~ <18), concentrated on the smaller OMPs tOmpA, OmpX and MipA.

Figure 5.21: Alphafold2 predicted confidence of interactions between abundant OMPs and all other OMPs. Confidence metric is the average (between five predictions) adjacent residue inter-chain PAE (i.e. only considering the PAE between residues of difference chains that are geometrically adjacent (1.5 nm cutoff). Hashed data-points indicate the average pLLDT of the model was less than 80. High values indicate low probabilities of interaction, lower values (red) are more likely to interact.

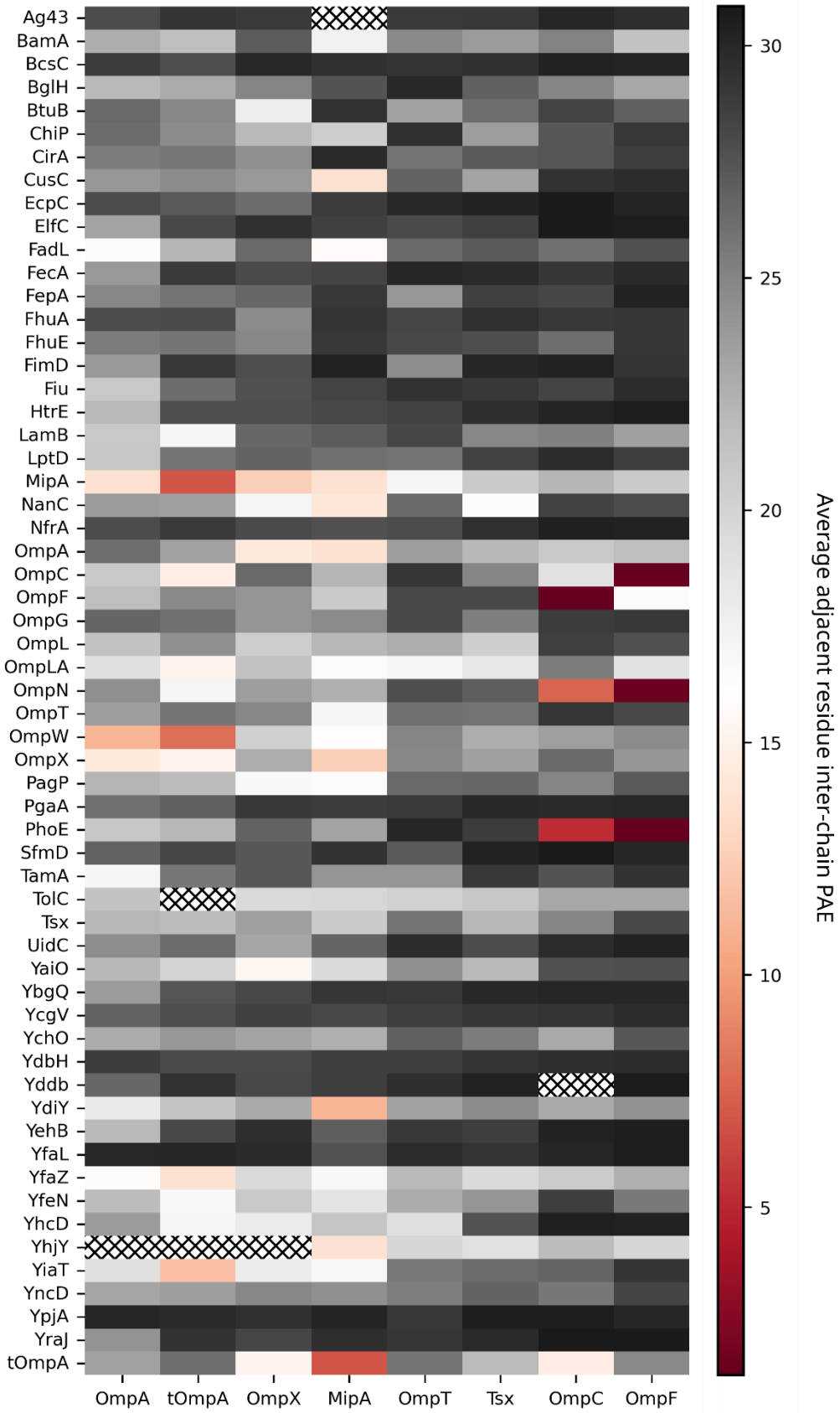


Figure 5.21: (Legend Overleaf)

To explore differences between the abundant OMPs, the distribution of PAEs over all predicted complexes were compared, **Fig. 5.22** (i.e. each column in **Fig. 5.21**). The three OMPs identified above (tOmpA, OmpX and MipA) have lower maxima and broader, negatively shifted distributions, evidenced by the skew and kurtosis of each distribution (stated on each graph panel). Analysing statistical differences between the distributions (Kolmogorov-Smirnov test) demonstrated two primary groupings (tOmpA/OmpX/MipA and OmpT/Tsx/OmpC/OmpF) and one outlier (OmpA) (**Fig 5.23a**, upper right). This grouping is further confirmed by considering the associated p-values, where non-significant differences are found within each cluster (**Fig 5.23a**, lower left). This highlights that tOmpA, OmpX and MipA are predicted to be more promiscuous than the other OMPs considered with more higher confidence predictions. This partly can be explained by the lower PAEs (i.e., better predicted complexes) that are typically found with smaller OMPs compared to larger ones (**Fig. 5.23b**), with a test dataset of predicted non-OMP pairs confirmed the PAE-size correlation is minimally attributable to an underlying feature of AF2 (**Fig. 5.23c**). Together, this analysis suggests that specific OMPs, generally smaller ones, are important for mediating OMP-OMP interactions in the OM, and coupled with the above biochemical data it particularly highlights the promiscuity of the barrel of OmpA, which this analysis predicts to have many different interaction partners.

The binding surfaces of each of the abundant OMPs was then assessed by identifying where other OMPs bind around the first OMPs β-barrel (**Fig. 5.24**). Predictions were weakly filtered (inter-chain PAE <25, pLDDT >80), aligned to the abundant OMP and the centre of mass of the transmembrane region of the interacting OMP calculated. This yields a range of different binding patterns, with MipA and OmpT appearing to have single binding sites, while other OMPs such as OmpF and OmpX, are predicted to interact at a diverse range of locations. OmpT's binding site is diametrically opposite to its LPS binding site, suggesting that evolution has evolved the exposed non-lipid binding surface for OMP interactions. About half of the OmpF and a third of the OmpC interactors considered are predicted at the trimer interface, which would yield conflicts with an intact trimer, suggesting why these predictions were often poorer quality than the OmpF/C monomers. It is possible that these interactions replace monomers of OmpF/C in the trimer (and larger hexameric lattice), or that they are excluded from occurring due to trimer formation. More broadly, this analysis highlights that the likely promiscuity of binding of (t)OmpA, OmpX and OmpF to other OMPs extends around much of the accessible intra-membrane protein surface.

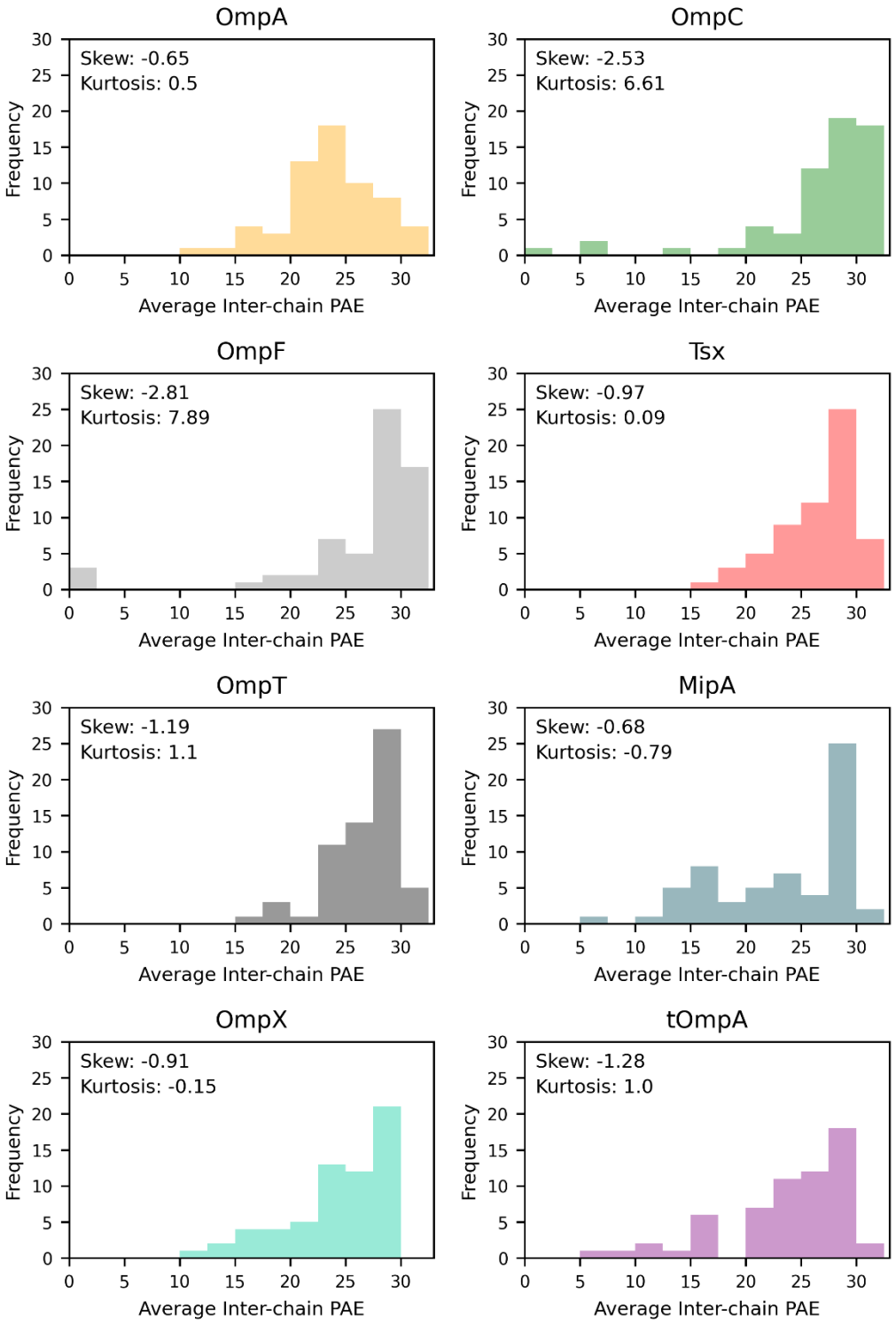


Figure 5.22: Distributions of Alphafold2 predicted confidence of interactions for abundant OMPs. X-axis of each plot is the confidence metric: the average (between five predictions) adjacent residue inter-chain PAE (i.e. only considering the PAE between residues of difference chains that are geometrically adjacent (1.5 nm cutoff). Predictions with average pLLDT <80 were excluded. Skew and kurtosis of each distribution are displayed.

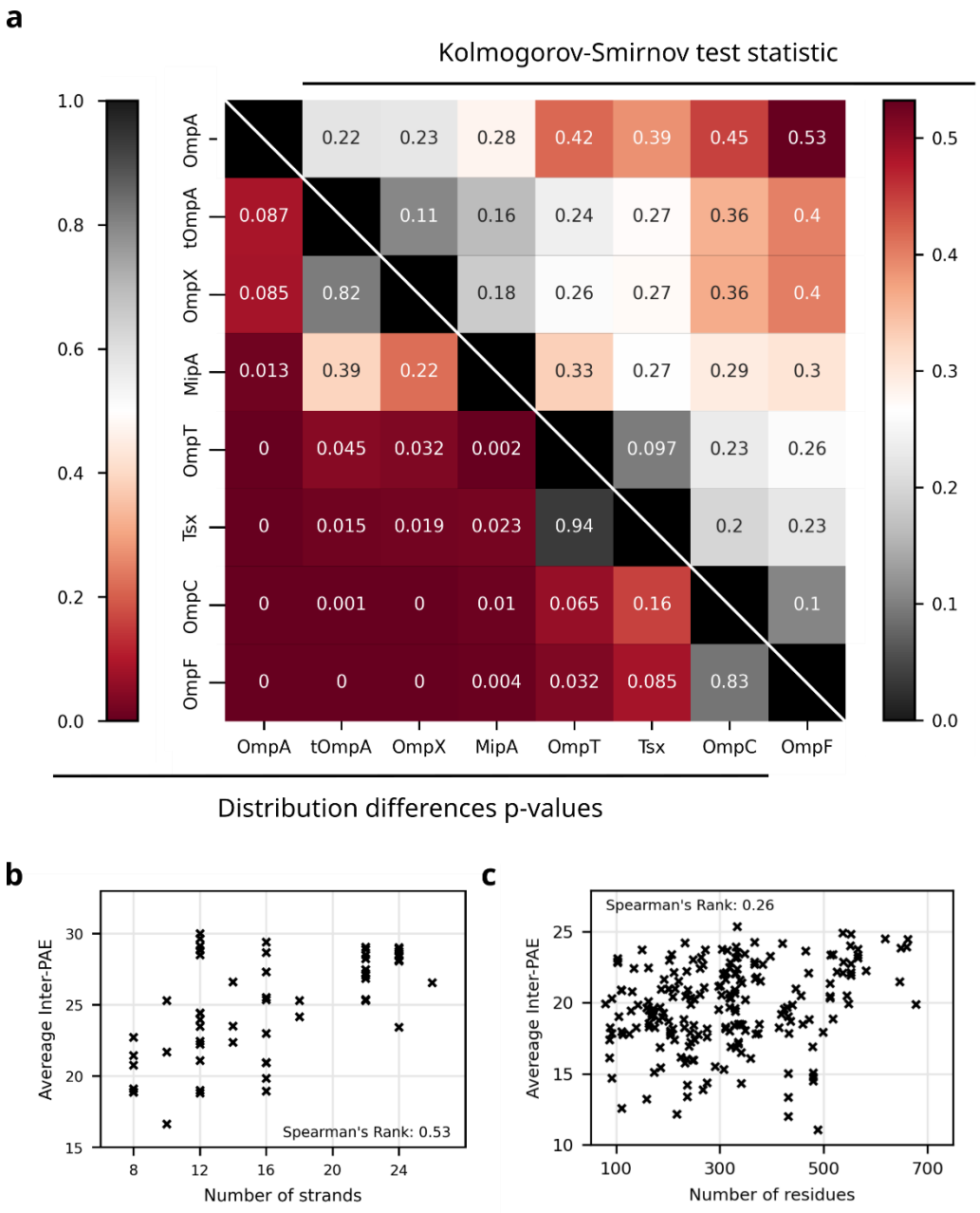


Figure 5.23: Clustering and differences between average inter-chain PAE distributions for each abundant OMP. (a) The data distributions shown in **Fig. 5.23** were compared using the Kolmogorov-Smirnov test (upper right values), smaller values indicate more similar data distributions. The associated p-values comparing each OMP pair are shown in the lower left half of the heatmap (to 3 significant figures). Two dominant clusters emerge: tOmp-OmpX-MipA and OmpT-Tsx-OmpX-OmpF. **(b)** Considering the average residue-adjacent inter-chain PAE from all OMPs to the most abundant OMPs reveals smaller OMPs are consistently predicted to be more promiscuous interactors, in contrast to **(c)** a control dataset of protein pair predictions that shows a minimal correlation between protein size and PAE.

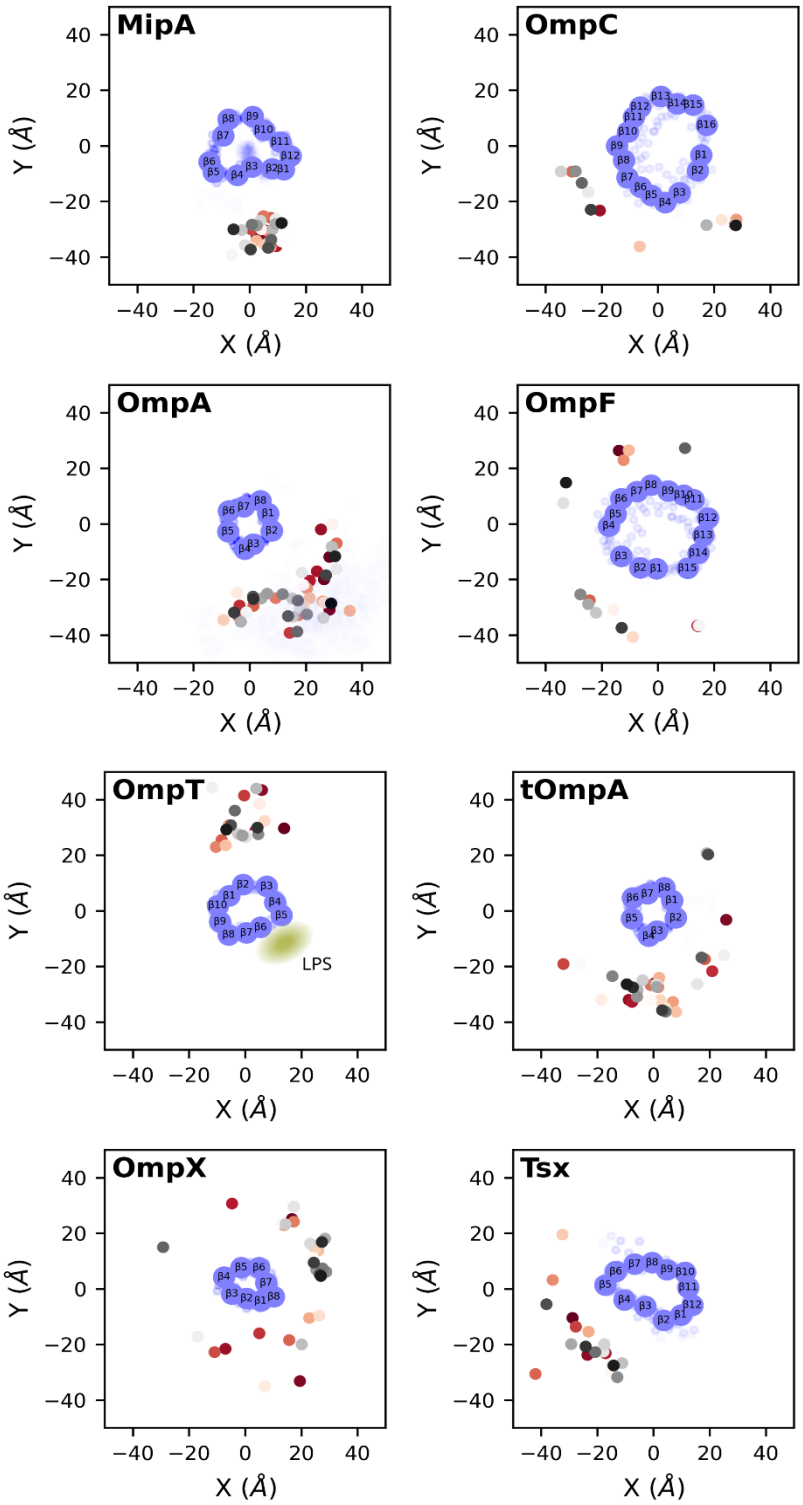


Figure 5.24: Angular distributions of predicted interaction OMPs. All predictions for the abundant OMPs (blue, inter-chain PAE <25, pLDDT >80) were aligned to the common OMP and the centre of mass of the second OMPs transmembrane domains plotted (dots, black to red based on their inter-chain PAE, as in **Fig. 5.22**). (Models with highly diverging predictions were filtered out based on inter-model RMSD). The green dot in the OmpT plot is the LPS binding site.

Next the confidence of the interaction based on the inter-strand PAE was compared with the single-OMP crosslinking data (Section 5.2.6) obtained in the presence or absence of Ra-LPS (**Fig. 5.25**). Both show good correlation, providing experimental confidence in the quality of the AF2 predictions (Spearman's rank of 0.80 and 0.90 without and with Ra-LPS respectively). However, there is more significant correlation when in the presence of Ra-LPS (p-value 0.037, compared to 0.104 in absence of Ra-LPS), indicating that despite the absence of lipid in the AF2 predictions, it is identifying some lipid-mediated coevolutionary signatures. Although only five experimental data points are compared, these high correlations suggest that rather than a mere confidence metric, a complex prediction's PAE may be informative about the strength, longevity or likelihood of an interaction to occur in the membrane on a semi-quantitative, apparently titratable scale.

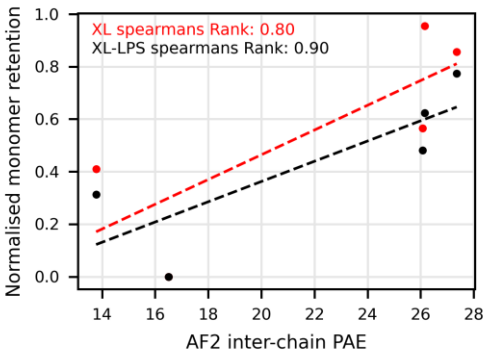


Figure 5.25: AF2 predictions agree with experimental crosslinking better in the presence of Ra-LPS. Comparison between crosslinking efficiency (monomer retention after 15 minutes, **Fig. 5.20d**) and AF2 predicted inter-chain PAE. (Red and black are with and without Ra-LPS, respectively). Note both a red and a black point occupy (~16.5, 0), but only the black point is visible.

Some of the higher confidence AF2 predictions are shown with their PAE plots in **Fig. 5.26a-i**. As noted above, they are enriched in (t)OmpA, MipA and OmpX. Indeed, four of the nine complexes shown here are between two of the abundant proteins, highlighting their increased probability to form interactions. In most cases it is challenging to interpret any possible functional consequences of the complexes, especially with proteins of unknown function (YiaT, which is a MipA homolog, and YdiY, **Fig. 5.26d,g**). Several of the predictions (e.g. OmpX-MipA, MipA-MipA, **Fig. 5.26e,f**) have large intervening space, suggestive of the inferred intermediate lipid binding discussed above. Interestingly, a high confidence interaction is found for tOmpA-OmpLA (away from the OmpLA dimerisation interface, **Fig. 5.26i**) indicating that, despite the negligible activity change (section 5.2.4), tOmpA may interact with the enzyme. Possible functionality of other complexes can be speculated based on functional correlations, for example both OmpA and MipA interact with peptidoglycan

(tOmp-MipA, **Fig. 5.26a**, while a MipA dimer may more effectively support assembly of the hypothesised peptidoglycan hydrolase complex[336] (MipA-MipA, **Fig. 5.26f**). However, some complexes (for example, tOmpA-OmpW, OmpX-MipA, OmpA-OmpX, **Fig. 5.26b, e, h**,) have no clear functional implications based on known protein functions, possibly reflecting the promiscuous binding of these proteins to form and stabilise OMP islands.

Figure 5.26: High confidence predicted interactions and their PAEs. Interactions between trimeric porin proteins not shown. For PAE plots, blue indicates high and red low confidence. In **(h)** OmpA C-terminal domain not shown in model for clarity, note both domains are in the PAE plot with the mostly red C-terminal domain unconfidently predicted, **(i)** tOmpA-OmpLA was predicted with a single copy of OmpLA, a second copy added for display to show the dimer interface.

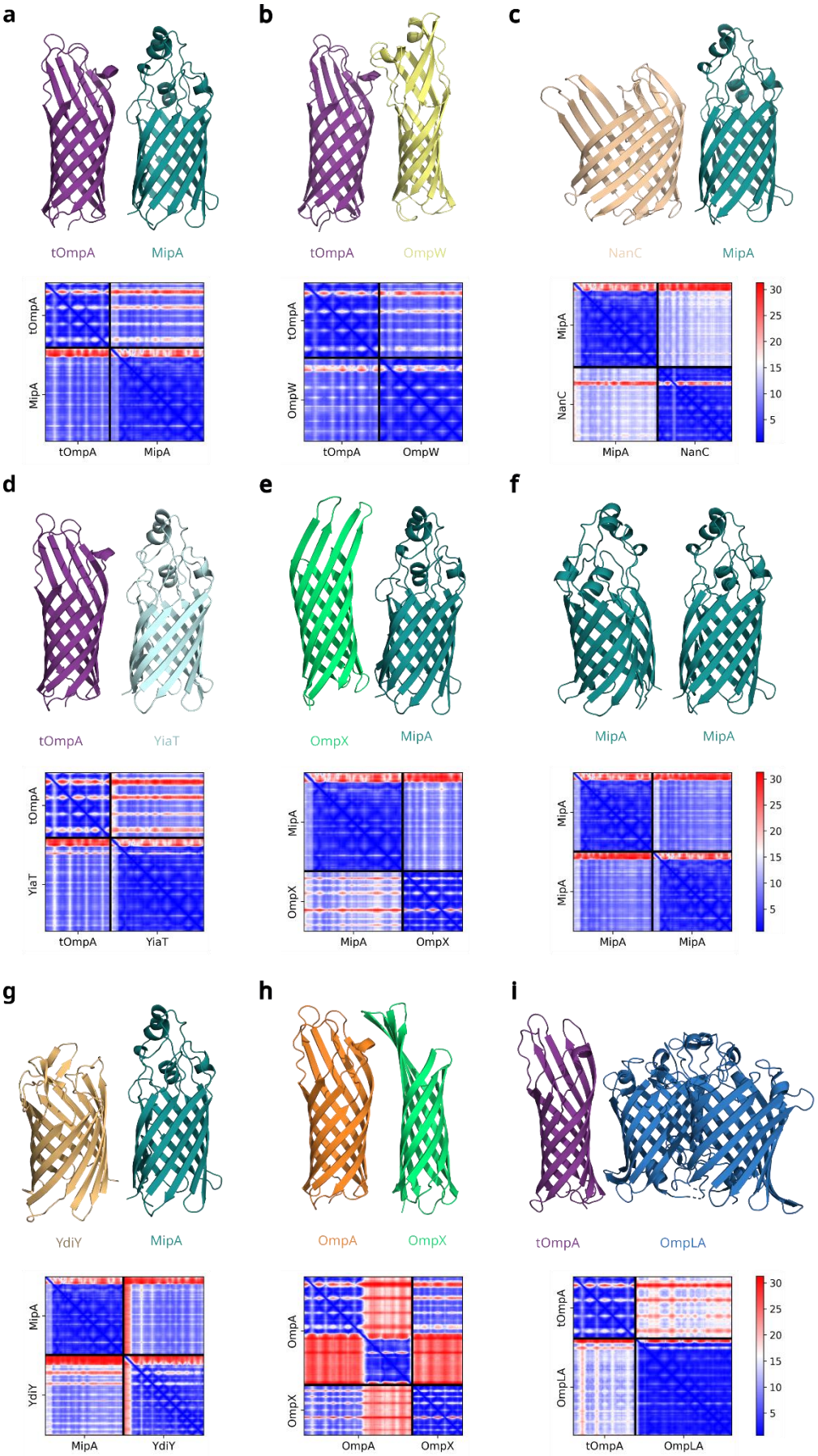


Figure 5.26: (Legend overleaf)

5.3 Discussion

The OM is crowded with proteins and there is little lipid, but while these features mean that modulation of OMP activity via specific OMP-OMP interactions is biochemically plausible and evolutionary expected[648,649], it has not previously been demonstrated. By exploiting and developing comparatively straightforward activity assays of the OM enzymes, two instances have been identified, where interaction with OmpA is able to up- or down-regulate the activity of an enzyme. OmpA and the trimeric porins OmpF/C are the most abundant OMPs, making up about two thirds of the OM's protein content[273]. However, unlike the trimeric porins, whose identity can change depending on environmental conditions[277], OmpA is consistently highly expressed allowing it to constitutively interact with other OMPs[323]. Furthermore, the trimeric porins are known to partially separate into arrays in the membrane[275], making them less available to interact, and their loops are partly sequestered in order to maintain the trimeric shape and pore structure[638], applying additional constraints for evolving other protein-protein interactions. In contrast, OmpA is evenly dispersed over the OM[650] and amongst other OMPs[493] (although likely as a dimer[651]) and defined roles of the loops do not typically block their accessibility[326]. It is interesting that the rate enhancement of OmpT activity by OmpA is explicitly mediated by the loops and not the transmembrane regions of the barrel, suggesting additional evolutionary flexibility in the loop sequences, supported by conservation patterns[652,653]. Together, these considerations suggest that OmpA is the ideal protein to have evolved additional, specific OMP-OMP functionality in the OM.

In Chapter 4 it was shown that in the absence of LPS, OmpT requires a negatively charged lipid environment to be basally active. This is further confirmed here by showing that in neutral DMPC membranes OmpT can be activated by MipA (but not OmpA) which is extracellularly electronegative (16 negative residues, 11 positive residues), indicating that in the absence of LPS, MipA may be a partial substitute, although with much lower catalytic enhancement. While the MipA activation and OmpA rate enhancement of OmpT are individually an order of magnitude less than LPS mediated activation, by acting in combination OmpA and MipA can reach a comparable OmpT activation (~2x less than LPS), offering a plausible route for OmpT activity in the LPS limiting conditions of OMP islands[275]: via interaction with two abundant OMPs. More generally the tripartite OmpT assay data glimpses the complex functional interaction landscape that could occur in the OM, with multiple OMPs acting in concert to modulate the activity of others. Unlike MipA,

the similarly extracellularly negatively charged BamA cannot activate OmpT, perhaps an adaptation to prevent premature activity of the enzyme and cleavage of the BAM complex. Specific OmpA binding by OmpT will also support OmpT function by sequestering negative charges in OmpA's loops in the interaction interface, decreasing the negativity of the surrounding environment and thus increasing the attraction of its positively-charged antimicrobial targets to the enzyme active site rather than the surrounding membrane. Thus, whether combined in an LPS-OmpT complex or not, interactions with OmpA will enhance OmpT activity, providing a greater protection to the cell.

In contrast to OmpT, PagP activity is downregulated by OmpA, but the reasons for this remain obscure. While phospholipid substrate accessibility is a defined regulator of PagP activity[412], given the low diffusivity in the OM, it is unclear how PagP can access sufficient LPS to alter the global membrane properties, especially given the precise substrate approach needed for its activity[444]. However, it is expected to localise to the LPS rich regions of the membrane, suggesting that PagP mislocalised to OMP islands could be downregulated by interacting with OmpA, possibly to prevent disruption of the tightly packed LPS-OMP networks. Alternatively, alterations in OmpA binding to peptidoglycan are known to modulate OMV release by controlling membrane planarity and stabiltiy[654], while hepta-acylated LPS is enriched in OMVs, indicating that PagP activity can enhance their formation[655]. Given both OmpA and PagP functions are used to manipulate OMV formation, the downregulation of PagP by OmpA could be part of a regulatory mechanism governing OMV formation. Although further work is needed to understand the physiological role of the PagP-OmpA interaction, it is clear that OmpA has adapted multiple regulatory functionalities.

While OmpA modulated both OmpT and PagP activity, demonstrating the benefits of evolving these interactions, OmpLA was not modulated by any of the OMPs tested. Thus, the OmpT and PagP rate modulations are not a generically evolved feature of OMPs, but appear to be specific for their functional niche, with not all OMPs developing them despite their consistent proximity to the most abundant OMPs in real OM. While OMP-OMP interactions do not always induce functional modulation, the AF2 modelling indicates that there is a higher level of co-evolution between the smaller OMPs and the general OMPome than for larger OMPs, and a tendency to have interactions predicted over more of their surface area, suggesting mediating roles in OMP clustering. The most abundant OMPs, which will have the

strongest evolutionary pressure to be general interactors, are typically smaller (two 8-stranded, two 10-stranded and two 16-stranded) allowing them to better pack inside OMP islands and minimise possible membrane defects, especially under lipid-limiting conditions. To mediate efficient packing, it is advantageous to be a generically good binder to many other OMPs and offers an explanation as to why the extracellular loops have been exploited for activity modulation rather than the barrels. Indeed, the ability to interact with multiple partners and hence form part of the glue holding OMP islands together is likely to be a specifically evolved function of these OMPs, and agrees with previous simulation data[140,491]. It also raises the intriguing possibility that the functionality of different regions of the OM could be modulated via the presence of diverse rare OMPs, generating different OMP islands of altered function.

Comparison of the AF2 modelling to experimental data showed a better correlation when LPS was present in the membranes (given that the negatively charged DMPG can replace some of the interactions of LPS, in different bilayers this effect is likely exaggerated). This is intriguing because it suggests that AF2 is better predicting interactions experimentally determined with LPS, suggesting AF2 (which relies heavily on co-evolutionary links) is identifying signatures of co-evolution between OMPs even when interactions are mediated by intervening lipid. This thus provides further evidence of the co-evolution of OMPs and LPS (c.f. Section 4.4).

In summary, these data give insights into OM organisation and the exploitation of specific features to maximise functionality, as well as identifying DMPG as a synthetic membrane competent to mediate efficient and controllable folding of OMPs and *de novo* designed proteins, critical for rigorous *in vitro* characterisation and for biotechnology where many functionally-designed proteins are based around transmembrane β-barrels[625,656]. Specific modulatory OMP-OMP interactions are identified between OmpA and enzymes of the OM, while modelling suggests that the small abundant OMPs are more promiscuous binders, facilitating OMP cluster formation. Although mechanistic details remain to be revealed, the results provide a glimpse of the complex functional interaction networks that occur in the protein-crowded OM and indicate the evolutionary exploitation of common interaction interfaces to ensure optimal cell fitness.

Chapter 6: Conclusions and perspective

Membrane biology is embedded in many of the challenges the world is currently facing, and the relatively poor understanding of how features of a complex membrane modulate and control membrane biology enhances the associated risks, especially of single-celled or bacterial pathogens. In particular, due to the membrane's vital role in bacterial fitness as the external facing layer of a broad range of pathogenic, symbiotic and exploitable species, the OM of diderm bacteria is a crucial surface for mediating external manipulation, whether by direct (e.g., antimicrobials, genetic manipulations) or indirect (e.g., diet) means. Despite its importance, the organisation of the OM have only recently been revealed, including the lattice-like nature of OMP arrays[36,491] and OMP/LPS phase separation[275]. Indeed, many questions about the ultrastructure of the OM and the interactions between its different components remain unanswered. This thesis has sought to address some of these questions, focussing on how lipid and protein organisation in the OM can modulate OMP folding and function.

The OM is an asymmetric membrane, with the presence of LPS in the outer leaflet and phospholipids in the inner leaflet, conferring a negative-outside charge dipole to the membrane. In Chapter 3 it was revealed that controllably introducing charge asymmetry across membranes *in vitro* strongly affected OMP folding and stability, with both enhanced upon the introduction of a native-like bilayer dipole. This is consistent with a co-evolution of the negatively charged LPS and OMP sequence and led to the discovery, herein described, of the conserved PEP in OMP sequences to ensure optimal folding. Additional experiments that introduced proteins into lipid symmetric membranes also suggested that, in addition to lipids, protein induced asymmetries can alter OMP folding rates. These data are particularly intriguing given the charge asymmetric properties of both the OM and β-barrel assembly machinery, BAM, suggesting that the membrane, OMP and BAM properties may have co-evolved to promote optimal protein folding and stability. Together, the data point to a possible new mechanism of BAM function, wherein it's negatively charged, extracellular loops may contribute to drive OMP folding. While demonstrated here using OMPs and membrane charge dipoles relevant to the OM, given that all biological membranes are asymmetric[657] and the known roles of charge in controlling α-helical bundle insertion rate and topology[658], it would seem likely that similar mechanisms occur in other membranes.

In addition to conferring an asymmetric charge dipole to the membrane, LPS also interacts specifically with OMPs to modulate their conformation and function. In

Chapter 4, the specific parts of the LPS oligosaccharide responsible for LPS-mediated activation of the protease OmpT were dissected. Remarkably, this revealed that distinct sugars in the LPS core oligosaccharide were required for optimal activity, with specific regions of OmpT interacting with specific LPS sugars, again highlighting the co-evolution of OMPs with their lipid environment. This also raises the intriguing, although untested, possibility that OMPs are adapted to their specific chemotypes of LPS, and that by designing precise binders it may be possible to target bacterial populations with exquisite sensitivity. LPS was further identified in Chapter 4 to be important for stabilising FusA in a state that was competent for structural determination by cryoEM, suggestive of important LPS-FusA interactions. These data thus lend weight to the increasing body of work that demonstrates the essentiality of LPS for mediating functional and structural alteration to OMPs, and particularly highlights the role of specific sugars in the LPS oligosaccharide in altering or modulating OMP stability and/or function.

While LPS is the only lipid in appreciable concentrations in the outer leaflet of the OM, the very high protein concentrations and OMP/LPS phase separation mean that many direct OMP-OMP interactions must also occur in the OM. In Chapter 5, the consequences of interactions between the three OMP enzymes (OmpT, OmpLA, PagP) and other OMPs was explored, revealing a novel and key role for OmpA, in binding these three OMPs. Given the very high frequency of OmpA in the OM (~30% of all OMPs), it is unsurprising that biology has exploited the interactions which must natively occur whether functional or not, between OmpA and other OMPs. However, these interactions are precise, with the enzyme activity of OmpT being activated by OmpA, the activity of PagP being inhibited and that of OmpLA unaffected. These three enzymes provide an experimentally tractable target for considering OMP-OMP interactions because of their relatively straightforward activity assays. However, given the protein crowding in the OM, it is plausible that many other functional, but as yet unappreciated, OMP-OMP interactions may occur (hinted at by the breadth of confident OMP-OMP dimer predictions by Alphafold2). Indeed, it is possible that the manipulation of global OMP content and/or OMP island composition provides a mechanism for specific OMP functional modulation by controlling the OMP-OMP interactions that can occur.

This study has generated many avenues for further research. Details of the role of charge and bilayer charge asymmetry in OMP folding remain unclear, and many facets of this merit additional study (including expanding the scope to more

OMPs). In particular, understanding the role of the PEP (and the positive outside rule more broadly) in OMP folding *in vivo* and how it interacts with BAM to ensure efficient folding would be exciting to pursue. The possibility that the negatively charged regions of the extracellular loops of BamA are helping to drive OMP folding would represent a new mechanism of action of BAM and is also worthy of future exploration (perhaps via *in vivo* and *in vitro* characterisation of BamA mutants). It was shown that for OmpT activity specific sugars of the oligosaccharide are important, but the sequence motifs in OMPs mediating LPS and sugar interactions required for function are uncertain. A broad bioinformatics study comparing OMP sequence/structure with LPS identity may be able to identify co-evolving patterns and thus pinpoint sequence motifs important for mediating these interactions. Coupling this with MD and *in vitro* OMP functional studies with different types of LPS and OMP mutants would be a powerful approach to derive general rules for LPS functional modulation of OMPs. An exploration of OMP-OMP interactions more broadly than presented here, and their functional consequences, would be valuable in understanding how the OM is able be both a formidable barrier, as well as a milieu in which enzymes and proteins can function. The Alphafold2 predictions presented in Chapter 5 provide a starting point for this work, but new *in vitro* (and ideally *in vivo*) functional assays would need to be developed to test these predictions. Indeed, it is possible that better understanding of functional OMP-OMP or OMP-LPS interactions in the OM will provide a new target for antimicrobials.

Collectively, the work in this thesis reveals new insights into how native features of the OM modulate OMP folding and function, thus giving hints at how the OM's organisation and composition allow it to operate synergistically to ensure optimal fitness. Indeed, given that the OM is synthesised, maintained, and remodelled in response to the environment without access to a periplasmic energy source, the importance of the stability and organisation of its ultrastructure is emphasised. This overall system architecture is critical for understanding how the OM functions, whether considering modulation of OMP folding, OMP activity or global membrane properties (e.g., membrane permeability and fluidity), all of which are affected by OMP-OMP and OMP-LPS/phospholipid interactions, membrane biophysical qualities and the compositional makeup of the OM. Recent work (e.g., [139,274,275,491]), including that presented here, takes steps towards putting our knowledge of the different parts of the OM together to gain a holistic understanding. The same Francis Bacon whose vision began this thesis remained humble enough to recognise that '[w]e rise to great

heights by a winding staircase of small steps.'[1] It seems likely that in the near future, based on improving tools and techniques in biophysics, computational and structural biology, we will soon rise high enough to observe how the properties of the native OM controls its form and function for the first time.

References

1. Bacon, Francis. *Novum Organum, sive Indicia Vera de Interpretatione Naturae.* London (1620). Tr. John Spedding, *Instauratio Magna.* (1858) Boston

2. Hooke, Robert. *Micrographia: or, some Physiological Descriptions of Minute Bodies made by Magnifying Glasses, with Observations and Inquiries thereupon*. (Royal Society, 1665).

3. Baker, H. & Leeuwenhoek, Mr. An Account of Mr. Leeuwenhoek's Microscopes; By Mr. Henry Baker, F. R. S. *Philos. Trans. 1683-1775* **41**, 503–519 (1739).

4. Schleiden MJ. Beiträge zur Phytogenesis. *Arch Anat Physiol Wiss Med* 137–176 (1838).

5. Schwann, T. Mikroscopische Untersuchungen über die Übereinstimmung in der Struktur und dem Wachstum der Thiere und Pflanzen. *W Engelmann Leipz. Vol 176* **176**, (1839).

6. Traube, Moritz. *Experimente zur Theorie der Zellenbildung*. (Centralblatt fu¨r die medicinischen Wissenschaften, 1864).

7. Pfeffer, Wilhelm. *Os- motische Untersuchungen: Studien zur Zellmechanik*. (Wilhelm Engelmann, 1877).

8. Meyer, H. Zur Theorie der Alkoholnarkose. *Arch. Für Exp. Pathol. Pharmakol.* **42**, 109–118 (1899).

9. Gorter, E. & Grendel, F. ON BIMOLECULAR LAYERS OF LIPOIDS ON THE CHROMOCYTES OF THE BLOOD. *J. Exp. Med.* **41**, 439–443 (1925).

10. Fricke, H. THE ELECTRIC CAPACITY OF SUSPENSIONS WITH SPECIAL REFERENCE TO BLOOD. *J. Gen. Physiol.* **9**, 137–152 (1925).

11. Korn, E. D. Structure of biological membranes. *Science* **153**, 1491–1498 (1966).

12. Lajos, L., Jobst, K. & Bacsó, K. THE LIPID CONTENT OF THE AMNIOTIC MEMBRANE AND THE PRODUCTION OF VERNIX CASEOSA. *BJOG Int. J. Obstet. Gynaecol.* **57**, 753–756 (1950).

13. Frey-Wyssling, A. Morphological aspects of the lipo-proteins in chloroplasts. *Discuss. Faraday Soc.* **6**, 130–134 (1949).

14. Palade, G. E. The fine structure of mitochondria. *Anat. Rec.* **114**, 427–451 (1952).

15. Sjostrand, F. S. Electron microscopy of mitochondria and cytoplasmic double membranes. *Nature* **171**, 30–32 (1953).

16. Robertson, J. D. The ultrastructure of cell membranes and their derivatives. *Biochem. Soc. Symp.* **16**, 3–43 (1959).

17. Danielli, J. F. & Davson, H. A contribution to the theory of permeability of thin films. *J. Cell. Comp. Physiol.* **5**, 495–508 (1935).

18. Benson, A. A. On the orientation of lipids in chloroplast and cell membranes. *J. Am. Oil Chem. Soc.* **43**, 265–270 (1966).

19. Vanderkooi, G. & Green, D. E. Biological membrane structure, I. The protein crystal model for membranes. *Proc. Natl. Acad. Sci. U. S. A.* **66**, 615–621 (1970).

20. Kendrew, J. C. *et al.* Structure of Myoglobin: A Three-Dimensional Fourier Synthesis at 2 Å. Resolution. *Nature* **185**, 422–427 (1960).

21. Tanford, C. The Hydrophobic Effect and the Organization of Living Matter. *Science* **200**, 1012–1018 (1978).

22. Kendrew, J. C. Myoglobin and the structure of proteins. *Science* **139**, 1259–1266 (1963).

23. Singer, S. J. & Nicolson, G. L. The fluid mosaic model of the structure of cell membranes. *Science* **175**, 720–731 (1972).

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

25. Nicolson, G. L. & Ferreira de Mattos, G. The Fluid-Mosaic model of cell membranes: A brief introduction, historical features, some general principles, and its adaptation to current information. *Biochim. Biophys. Acta Biomembr.* **1865**, 184135 (2023).

26. Levental, I. & Lyman, E. Regulation of membrane protein structure and function by their lipid nano-environment. *Nat. Rev. Mol. Cell Biol.* **24**, 107–122 (2023).

27. Fahy, E. *et al.* Update of the LIPID MAPS comprehensive classification system for lipids1. *J. Lipid Res.* **50**, S9–S14 (2009).

28. Gerle, C. Essay on Biomembrane Structure. *J. Membr. Biol.* **252**, 115–130 (2019).

29. Nicolson, G. L. & Ferreira de Mattos, G. A Brief Introduction to Some Aspects of the Fluid-Mosaic Model of Cell Membrane Structure and Its Importance in Membrane Lipid Replacement. *Membranes* **11**, 947 (2021).

30. Sezgin, E., Levental, I., Mayor, S. & Eggeling, C. The mystery of membrane organization: composition, regulation and roles of lipid rafts. *Nat. Rev. Mol. Cell Biol.* **18**, 361–374 (2017).

31. Horne, J. E., Brockwell, D. J. & Radford, S. E. Role of the lipid bilayer in outer membrane protein folding in Gram-negative bacteria. *J. Biol. Chem.* **295**, 10340–10367 (2020).

32. Oddershede, L., Dreyer, J. K., Grego, S., Brown, S. & Berg-Sørensen, K. The motion of a single molecule, the lambda-receptor, in the bacterial outer membrane. *Biophys. J.* **83**, 3152–3161 (2002).

33. Oh, D., Yu, Y., Lee, H., Wanner, B. L. & Ritchie, K. Dynamics of the serine chemoreceptor in the Escherichia coli inner membrane: a high-speed single-molecule tracking study. *Biophys. J.* **106**, 145–153 (2014).

34. Saffman, P. G. & Delbrück, M. Brownian motion in biological membranes. *Proc. Natl. Acad. Sci. U. S. A.* **72**, 3111–3113 (1975).

35. Jacobson, K., Liu, P. & Lagerholm, B. C. The Lateral Organization and Mobility of Plasma Membrane Components. *Cell* **177**, 806–819 (2019).

36. Chavent, M. *et al.* How nanoscale protein interactions determine the mesoscale dynamic organisation of bacterial outer membrane proteins. *Nat. Commun.* **9**, 2846 (2018).

37. Lee, A. G. Lipid-protein interactions in biological membranes: a structural perspective. *Biochim. Biophys. Acta* **1612**, 1–40 (2003).

38. Niggli, V. Lipid Interactions of Cytoskeletal Proteins. in *Advances in Molecular and Cell Biology* vol. 37 221–250 (Elsevier, 2006).

39. Burt, A. *et al.* Complete structure of the chemosensory array core signalling unit in an E. coli minicell strain. *Nat. Commun.* **11**, 743 (2020).

40. Honigmann, A. *et al.* Phosphatidylinositol 4,5-bisphosphate clusters act as molecular beacons for vesicle recruitment. *Nat. Struct. Mol. Biol.* **20**, 679–686 (2013).

41. Engelman, D. M. Membranes are more mosaic than fluid. *Nature* **438**, 578–580 (2005).

42. Penn, W. D. *et al.* Probing biophysical sequence constraints within the transmembrane domains of rhodopsin by deep mutational scanning. *Sci. Adv.* **6**, eaay7505 (2020).

43. Chaturvedi, D. & Mahalakshmi, R. Transmembrane β-barrels: Evolution, folding and energetics. *Biochim. Biophys. Acta Biomembr.* **1859**, 2467–2482 (2017).

44. Hegde, R. S. & Keenan, R. J. The mechanisms of integral membrane protein biogenesis. *Nat. Rev. Mol. Cell Biol.* **23**, 107–124 (2022).

45. Doyle, M. T. & Bernstein, H. D. Function of the Omp85 Superfamily of Outer Membrane Protein Assembly Factors and Polypeptide Transporters. *Annu. Rev. Microbiol.* **76**, 259–279 (2022).

46. Cao, T. B. & Saier, M. H. The general protein secretory pathway: phylogenetic analyses leading to evolutionary conclusions. *Biochim. Biophys. Acta BBA - Biomembr.* **1609**, 115–125 (2003).

47. Antunes, L. C. *et al.* Phylogenomic analysis supports the ancestral presence of LPS-outer membranes in the Firmicutes. *eLife* **5**, e14589 (2016).

48. Tsirigotaki, A., De Geyter, J., Šoštaric´, N., Economou, A. & Karamanou, S. Protein export through the bacterial Sec pathway. *Nat. Rev. Microbiol.* **15**, 21–36 (2017).

49. Van den Berg, B. *et al.* X-ray structure of a protein-conducting channel. *Nature* **427**, 36–44 (2004).

50. Hanada, M., Nishiyama, K. I., Mizushima, S. & Tokuda, H. Reconstitution of an efficient protein translocation machinery comprising SecA and the three membrane proteins, SecY, SecE, and SecG (p12). *J. Biol. Chem.* **269**, 23625–23631 (1994).

51. Driessen, A. J. M. & Nouwen, N. Protein Translocation Across the Bacterial Cytoplasmic Membrane. *Annu. Rev. Biochem.* **77**, 643–667 (2008).

52. Braun, V. & Hantke, K. Lipoproteins: Structure, Function, Biosynthesis. *Subcell. Biochem.* **92**, 39–77 (2019).

53. El Arnaout, T. & Soulimane, T. Targeting Lipoprotein Biogenesis: Considerations towards Antimicrobials. *Trends Biochem. Sci.* **44**, 701–715 (2019).

54. Rivera-Cuevas, Y. & Carruthers, V. B. The multifaceted interactions between pathogens and host ESCRT machinery. *PLoS Pathog.* **19**, e1011344 (2023).

55. Lemmon, M. A. Membrane recognition by phospholipid-binding domains. *Nat. Rev. Mol. Cell Biol.* **9**, 99–111 (2008).

56. Iadanza, M. G. *et al.* Lateral opening in the intact β-barrel assembly machinery captured by cryo-EM. *Nat. Commun.* **7**, 12865 (2016).

57. White, J. B. R. *et al.* Outer membrane utilisomes mediate glycan uptake in gut Bacteroidetes. *Nature* 1–7 (2023) doi:10.1038/s41586-023-06146-w.

58. Saotome, K. *et al.* Structural analysis of cancer-relevant TCR-CD3 and peptide-MHC complexes by cryoEM. *Nat. Commun.* **14**, 2401 (2023).

59. Takeda, K. *et al.* Crystal structure of the M intermediate of bacteriorhodopsin: allosteric structural changes mediated by sliding movement of a transmembrane helix. *J. Mol. Biol.* **341**, 1023–1037 (2004).

60. van den Berg, B., Black, P. N., Clemons, W. M. & Rapoport, T. A. Crystal structure of the long-chain fatty acid transporter FadL. *Science* **304**, 1506–1509 (2004).

61. Abellón-Ruiz, J. *et al.* Structural basis for maintenance of bacterial outer membrane lipid asymmetry. *Nat. Microbiol.* **2**, 1616–1623 (2017).

62. Hyvönen, M. & Saraste, M. Structure of the PH domain and Btk motif from Bruton's tyrosine kinase: molecular explanations for X-linked agammaglobulinaemia. *EMBO J.* **16**, 3396–3404 (1997).

63. Melero, A. & Jiménez-Rojo, N. Cracking the membrane lipid code. *Curr. Opin. Cell Biol.* **83**, 102203 (2023).

64. Tang, X. *et al.* Structural insights into outer membrane asymmetry maintenance in Gram-negative bacteria by MlaFEDB. *Nat. Struct. Mol. Biol.* **28**, 81–91 (2021).

65. Devaux, P. F., Herrmann, A., Ohlwein, N. & Kozlov, M. M. How lipid flippases can modulate membrane structure. *Biochim. Biophys. Acta BBA - Biomembr.* **1778**, 1591–1600 (2008).

66. Botte, M. *et al.* Cryo-EM structures of a LptDE transporter in complex with Pro-macrobodies offer insight into lipopolysaccharide translocation. *Nat. Commun.* **13**, 1826 (2022).

67. Fadeel, B. & Xue, D. The ins and outs of phospholipid asymmetry in the plasma membrane: roles in health and disease. *Crit. Rev. Biochem. Mol. Biol.* **44**, 264–277 (2009).

68. Chaurio, R. A. *et al.* Phospholipids: Key Players in Apoptosis and Immune Regulation. *Molecules* **14**, 4892–4914 (2009).

69. Bertin, A. *et al.* Human ESCRT-III polymers assemble on positively curved membranes and induce helical membrane tube formation. *Nat. Commun.* **11**, 2663 (2020).

70. Buyan, A. *et al.* Piezo1 Forms Specific, Functionally Important Interactions with Phosphoinositides and Cholesterol. *Biophys. J.* **119**, 1683–1697 (2020).

71. Bogdanov, M. *et al.* Phospholipid distribution in the cytoplasmic membrane of Gram-negative bacteria is highly asymmetric, dynamic, and cell shape-dependent. *Sci. Adv.* **6**, eaaz6333 (2020).

72. Bury, M. *et al.* Ophiobolin A induces paraptosis-like cell death in human glioblastoma cells by decreasing BKCa channel activity. *Cell Death Dis.* **4**, e561–e561 (2013).

73. London, E. Membrane Structure-Function Insights from Asymmetric Lipid Vesicles. *Acc. Chem. Res.* **52**, 2382–2391 (2019).

74. Marquardt, D., Geier, B. & Pabst, G. Asymmetric lipid membranes: towards more realistic model systems. *Membranes* **5**, 180–196 (2015).

75. Krompers, M. & Heerklotz, H. A Guide to Your Desired Lipid-Asymmetric Vesicles. *Membranes* **13**, 267 (2023).

76. Szente, L., Singhal, A., Domokos, A. & Song, B. Cyclodextrins: Assessing the Impact of Cavity Size, Occupancy, and Substitutions on Cytotoxicity and Cholesterol Homeostasis. *Mol. J. Synth. Chem. Nat. Prod. Chem.* **23**, 1228 (2018).

77. Frijlink, H. W. *et al.* The effect of parenterally administered cyclodextrins on cholesterol levels in the rat. *Pharm. Res.* **8**, 9–16 (1991).

78. Gimpl, G. & Gehrig-Burger, K. Probes for studying cholesterol binding and cell biology. *Steroids* **76**, 216–231 (2011).

79. Doktorova, M. *et al.* Preparation of asymmetric phospholipid vesicles for use as cell membrane models. *Nat. Protoc.* **13**, 2086–2101 (2018).

80. Markones, M. *et al.* Engineering Asymmetric Lipid Vesicles: Accurate and Convenient Control of the Outer Leaflet Lipid Composition. *Langmuir ACS J. Surf. Colloids* **34**, 1999–2005 (2018).

81. Markones, M. *et al.* Stairway to Asymmetry: Five Steps to Lipid-Asymmetric Proteoliposomes. *Biophys. J.* **118**, 294–302 (2020).

82. Chiantia, S. & London, E. Acyl chain length and saturation modulate interleaflet coupling in asymmetric bilayers: effects on dynamics and structural order. *Biophys. J.* **103**, 2311–2319 (2012).

83. Lin, Q. & London, E. Ordered raft domains induced by outer leaflet sphingomyelin in cholesterol-rich asymmetric vesicles. *Biophys. J.* **108**, 2212–2222 (2015).

84. Cheng, H.-T. & London, E. Preparation and properties of asymmetric large unilamellar vesicles: interleaflet coupling in asymmetric vesicles is dependent on temperature but not curvature. *Biophys. J.* **100**, 2671–2678 (2011).

85. Lin, Q. & London, E. The influence of natural lipid asymmetry upon the conformation of a membrane-inserted protein (perfringolysin O). *J. Biol. Chem.* **289**, 5467–5478 (2014).

86. Scott, H. L., Heberle, F. A., Katsaras, J. & Barrera, F. N. Phosphatidylserine Asymmetry Promotes the Membrane Insertion of a Transmembrane Helix. *Biophys. J.* **116**, 1495–1506 (2019).

87. von Heijne, G. Control of topology and mode of assembly of a polytopic membrane protein by positively charged residues. *Nature* **341**, 456–458 (1989).

88. Piller, P., Semeraro, E. F., Rechberger, G. N., Keller, S. & Pabst, G. Allosteric modulation of integral protein activity by differential stress in asymmetric membranes. *PNAS Nexus* **2**, pgad126 (2023).

89. Jiang, Y. *et al.* Membrane-mediated protein interactions drive membrane protein organization. *Nat. Commun.* **13**, 7373 (2022).

90. Hao, B., Zhou, W. & Theg, S. M. Hydrophobic mismatch is a key factor in protein transport across lipid bilayer membranes via the Tat pathway. *J. Biol. Chem.* **298**, 101991 (2022).

91. Parker, J. L. & Newstead, S. Gateway to the Golgi: molecular mechanisms of nucleotide sugar transporters. *Curr. Opin. Struct. Biol.* **57**, 127–134 (2019).

92. Budin, I. *et al.* Viscous control of cellular respiration by membrane lipid composition. *Science* **362**, 1186–1189 (2018).

93. Sinensky, M. Homeoviscous Adaptation—A Homeostatic Process that Regulates the Viscosity of Membrane Lipids in Escherichia coli. *Proc. Natl. Acad. Sci.* **71**, 522–525 (1974).

94. Cox, C. D., Bavi, N. & Martinac, B. Bacterial Mechanosensors. *Annu. Rev. Physiol.* **80**, 71–93 (2018).

95. Coste, B. *et al.* Piezo1 and Piezo2 are essential components of distinct mechanically activated cation channels. *Science* **330**, 55–60 (2010).

96. Reddy, B., Bavi, N., Lu, A., Park, Y. & Perozo, E. Molecular basis of force-from-lipids gating in the mechanosensitive channel MscS. *eLife* **8**, e50486 (2019).

97. Ge, J. *et al.* Molecular mechanism of prestin electromotive signal amplification. *Cell* **184**, 4669-4679.e13 (2021).

98. Orsini, F. *et al.* Atomic force microscopy imaging of lipid rafts of human breast cancer cells. *Biochim. Biophys. Acta* **1818**, 2943–2949 (2012).

99. Castello-Serrano, I., Lorent, J. H., Ippolito, R., Levental, K. R. & Levental, I. Myelin-Associated MAL and PLP Are Unusual among Multipass Transmembrane Proteins in Preferring Ordered Membrane Domains. *J. Phys. Chem. B* **124**, 5930–5939 (2020).

100. Parton, R. G., Kozlov, M. M. & Ariotti, N. Caveolae and lipid sorting: Shaping the cellular response to stress. *J. Cell Biol.* **219**, e201905071 (2020).

101. Stone, M. B., Shelby, S. A., Núñez, M. F., Wisser, K. & Veatch, S. L. Protein sorting by lipid phase-like domains supports emergent signaling function in B lymphocyte plasma membranes. *eLife* **6**, e19891 (2017).

102. Hong, H. & Bowie, J. U. Dramatic destabilization of transmembrane helix interactions by features of natural membrane environments. *J. Am. Chem. Soc.* **133**, 11389–11398 (2011).

103. Romanenko, V. G. *et al.* Cholesterol sensitivity and lipid raft targeting of Kir2.1 channels. *Biophys. J.* **87**, 3850–3861 (2004).

104. Hansen, S. B., Tao, X. & MacKinnon, R. Structural basis of PIP2 activation of the classical inward rectifier K+ channel Kir2.2. *Nature* **477**, 495–498 (2011).

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

106. Corey, R. A., Stansfeld, P. J. & Sansom, M. S. P. The energetics of protein-lipid interactions as viewed by molecular simulations. *Biochem. Soc. Trans.* **48**, 25–37 (2020).

107. Liu, Q. *et al.* A proximity-tagging system to identify membrane protein–protein interactions. *Nat. Methods* **15**, 715–722 (2018).

108. Moore, D. T., Berger, B. W. & DeGrado, W. F. Protein-protein interactions in the membrane: sequence, structural, and biological motifs. *Struct. Lond. Engl. 1993* **16**, 991–1001 (2008).

109. Phillips, R., Ursell, T., Wiggins, P. & Sens, P. Emerging roles for lipids in shaping membrane-protein function. *Nature* **459**, 379–385 (2009).

110. Cymer, F., Veerappan, A. & Schneider, D. Transmembrane helix-helix interactions are modulated by the sequence context and by lipid bilayer properties. *Biochim. Biophys. Acta* **1818**, 963–973 (2012).

111. van den Bogaart, G. *et al.* Membrane protein sequestering by ionic protein-lipid interactions. *Nature* **479**, 552–555 (2011).

112. Hiroaki, Y. *et al.* Implications of the Aquaporin-4 Structure on Array Formation and Cell Adhesion. *J. Mol. Biol.* **355**, 628–639 (2006).

113. Roche, J. V. & Törnroth-Horsefield, S. Aquaporin Protein-Protein Interactions. *Int. J. Mol. Sci.* **18**, 2255 (2017).

114. Vercellino, I. & Sazanov, L. A. Structure and assembly of the mammalian mitochondrial supercomplex CIII2CIV. *Nature* **598**, 364–367 (2021).

115. Aboulwafa, M., Zhang, Z. & Saier, M. H. Protein:Protein interactions in the cytoplasmic membrane apparently influencing sugar transport and phosphorylation activities of the e. coli phosphotransferase system. *PloS One* **14**, e0219332 (2019).

116. Hirst, J. Open questions: respiratory chain supercomplexes—why are they there and what do they do? *BMC Biol.* **16**, 1–4 (2018).

117. Diwanji, D. *et al.* Structures of the HER2–HER3–NRG1β complex reveal a dynamic dimer interface. *Nature* **600**, 339–343 (2021).

118. Membrane bending by protein phase separation - PubMed. https://pubmed.ncbi.nlm.nih.gov/33688043/.

119. Gram, H. C. J. & Friedlaender, C. *Ueber die isolirte Färbung der Schizomyceten: in Schnitt-und Trockenpräparaten*. (Theodor Fischer's medicinischer Buchhandlung, 1884).

120. Bladen, H. A. & Mergenhagen, S. E. ULTRASTRUCTURE OF VEILLONELLA AND MORPHOLOGICAL CORRELATION OF AN OUTER MEMBRANE WITH PARTICLES ASSOCIATED WITH ENDOTOXIC ACTIVITY. *J. Bacteriol.* **88**, 1482–1492 (1964).

121. Sutcliffe, I. C. A phylum level perspective on bacterial cell envelope architecture. *Trends Microbiol.* **18**, 464–470 (2010).

122. Skennerton, C. T. *et al.* Phylogenomic analysis of Candidatus 'Izimaplasma' species: free-living representatives from a Tenericutes clade found in methane seeps. *ISME J.* **10**, 2679–2692 (2016).

123. Coleman, G. A. *et al.* A rooted phylogeny resolves early bacterial evolution. *Science* **372**, eabe0511 (2021).

124. Cavalier-Smith, T. Deep phylogeny, ancestral groups and the four ages of life. *Philos. Trans. R. Soc. B Biol. Sci.* **365**, 111–132 (2010).

125. Chiaradia, L. *et al.* Dissecting the mycobacterial cell envelope and defining the composition of the native mycomembrane. *Sci. Rep.* **7**, 12807 (2017).

126. Brown, D. R., Whitcomb, R. F. & Bradbury, J. M. Revised minimal standards for description of new species of the class Mollicutes (division Tenericutes). *Int. J. Syst. Evol. Microbiol.* **57**, 2703–2719 (2007).

127. Fagan, R. P. & Fairweather, N. F. Biogenesis and functions of bacterial S-layers. *Nat. Rev. Microbiol.* **12**, 211–222 (2014).

128. Whitfield, C., Wear, S. S. & Sande, C. Assembly of Bacterial Capsular Polysaccharides and Exopolysaccharides. *Annu. Rev. Microbiol.* **74**, 521–543 (2020).

129. Megrian, D., Taib, N., Witwinowski, J., Beloin, C. & Gribaldo, S. One or two membranes? Diderm Firmicutes challenge the Gram-positive/Gram-negative divide. *Mol. Microbiol.* **113**, 659–671 (2020).

130. Léonard, R. R. *et al.* Was the Last Bacterial Common Ancestor a Monoderm after All? *Genes* **13**, 376 (2022).

131. Beveridge, T. J. & Davies, J. A. Cellular responses of Bacillus subtilis and Escherichia coli to the Gram stain. *J. Bacteriol.* **156**, 846–858 (1983).

132. Hernández, S. B. & Cava, F. New approaches and techniques for bacterial cell wall analysis. *Curr. Opin. Microbiol.* **60**, 88–95 (2021).

133. Viljoen, A., Foster, S. J., Fantner, G. E., Hobbs, J. K. & Dufrêne, Y. F. Scratching the Surface: Bacterial Cell Envelopes at the Nanoscale. *mBio* **11**, e03020-19 (2020).

134. Pasquina-Lemonche, L. *et al.* The architecture of the Gram-positive bacterial cell wall. *Nature* **582**, 294–297 (2020).

135. Dover, R. S., Bitler, A., Shimoni, E., Trieu-Cuot, P. & Shai, Y. Multiparametric AFM reveals turgor-responsive net-like peptidoglycan architecture in live streptococci. *Nat. Commun.* **6**, 7193 (2015).

136. Ultee, E. *et al.* Teichoic acids anchor distinct cell wall lamellae in an apically growing bacterium. *Commun. Biol.* **3**, 1–9 (2020).

137. Whatmore, A. M. & Reed, R. H. Determination of turgor pressure in Bacillus subtilis: a possible role for K+ in turgor regulation. *J. Gen. Microbiol.* **136**, 2521–2526 (1990).

138. Mathew, B. J. *et al.* Role of Streptococcus pneumoniae extracellular glycosidases in immune evasion. *Front. Cell. Infect. Microbiol.* **13**, 1109449 (2023).

139. Lithgow, T., Stubenrauch, C. J. & Stumpf, M. P. H. Surveying membrane landscapes: a new look at the bacterial cell surface. *Nat. Rev. Microbiol.* 1–17 (2023) doi:10.1038/s41579-023-00862-w.

140. Rassam, P. *et al.* Supramolecular assemblies underpin turnover of outer membrane proteins in bacteria. *Nature* **523**, 333–336 (2015).

141. Schiffrin, B., Brockwell, D. J. & Radford, S. E. Outer membrane protein folding from an energy landscape perspective. *BMC Biol.* **15**, 123 (2017).

142. Noinaj, N., Guillier, M., Barnard, T. J. & Buchanan, S. K. TonB-dependent transporters: regulation, structure, and function. *Annu. Rev. Microbiol.* **64**, 43–60 (2010).

143. Silhavy, T. J., Kahne, D. & Walker, S. The bacterial cell envelope. *Cold Spring Harb. Perspect. Biol.* **2**, a000414 (2010).

144. Cronan, J. E. Bacterial Membrane Lipids: Where Do We Stand? *Annu. Rev. Microbiol.* **57**, 203–224 (2003).

145. Bertani, B. & Ruiz, N. Function and Biogenesis of Lipopolysaccharides. *EcoSal Plus* **8**, (2018).

146. Egan, A. J. F., Errington, J. & Vollmer, W. Regulation of peptidoglycan synthesis and remodelling. *Nat. Rev. Microbiol.* **18**, 446–460 (2020).

147. Asmar, A. T. *et al.* Communication across the bacterial cell envelope depends on the size of the periplasm. *PLoS Biol.* **15**, e2004303 (2017).

148. Cayley, D. S., Guttman, H. J. & Record, M. T. Biophysical characterization of changes in amounts and activity of Escherichia coli cell and compartment water and turgor pressure in response to osmotic stress. *Biophys. J.* **78**, 1748–1764 (2000).

149. Rojas, E. R. *et al.* The outer membrane is an essential load-bearing element in Gram-negative bacteria. *Nature* **559**, 617–621 (2018).

150. Goemans, C., Denoncin, K. & Collet, J.-F. Folding mechanisms of periplasmic proteins. *Biochim. Biophys. Acta BBA - Mol. Cell Res.* **1843**, 1517–1528 (2014).

151. Sharma, P., Garg, N., Sharma, A., Capalash, N. & Singh, R. Nucleases of bacterial pathogens as virulence factors, therapeutic targets and diagnostic markers. *Int. J. Med. Microbiol.* **309**, 151354 (2019).

152. Troman, L. & Collinson, I. Pushing the Envelope: The Mysterious Journey Through the Bacterial Secretory Machinery, and Beyond. *Front. Microbiol.* **12**, (2021).

153. Mandela, E. *et al.* Adaptation of the periplasm to maintain spatial constraints essential for cell envelope processes and cell viability. *eLife* **11**, e73516 (2022).

154. Collet, J.-F., Cho, S.-H., Iorga, B. I. & Goemans, C. V. How the assembly and protection of the bacterial cell envelope depend on cysteine residues. *J. Biol. Chem.* **295**, 11984–11994 (2020).

155. López-Lara, I. M. & Geiger, O. Bacterial lipid diversity. *Biochim. Biophys. Acta Mol. Cell Biol. Lipids* **1862**, 1287–1299 (2017).

156. Nikaido, H. Outer membrane of Salmonella typhimurium. Transmembrane diffusion of some hydrophobic substances. *Biochim. Biophys. Acta* **433**, 118–132 (1976).

157. Funahara, Y. & Nikaido, H. Asymmetric localization of lipopolysaccharides on the outer membrane of Salmonella typhimurium. *J. Bacteriol.* **141**, 1463–1465 (1980).

158. Kamio, Y. & Nikaido, H. Outer membrane of Salmonella typhimurium: accessibility of phospholipid head groups to phospholipase c and cyanogen bromide activated dextran in the external medium. *Biochemistry* **15**, 2561–2570 (1976).

159. Nikaido, H. Molecular basis of bacterial outer membrane permeability revisited. *Microbiol. Mol. Biol. Rev. MMBR* **67**, 593–656 (2003).

160. Lessen, H. J., Fleming, P. J., Fleming, K. G. & Sodt, A. J. Building Blocks of the Outer Membrane: Calculating a General Elastic Energy Model for β-Barrel Membrane Proteins. *J. Chem. Theory Comput.* **14**, 4487–4497 (2018).

161. Ursell, T. S., Trepagnier, E. H., Huang, K. C. & Theriot, J. A. Analysis of surface protein expression reveals the growth pattern of the gram-negative outer membrane. *PLoS Comput. Biol.* **8**, e1002680 (2012).

162. Wong, L. H., Gatta, A. T. & Levine, T. P. Lipid transfer proteins: the lipid commute via shuttles, bridges and tubes. *Nat. Rev. Mol. Cell Biol.* **20**, 85–101 (2019).

163. Rodriguez-Loureiro, I., Latza, V. M., Fragneto, G. & Schneck, E. Conformation of Single and Interacting Lipopolysaccharide Surfaces Bearing O-Side Chains. *Biophys. J.* **114**, 1624–1635 (2018).

164. Basauri, A. *et al.* Biochemical interactions between LPS and LPS-binding molecules. *Crit. Rev. Biotechnol.* **40**, 292–305 (2020).

165. Girard, R. *et al.* Lipopolysaccharides from *Legionella* and *Rhizobium* stimulate mouse bone marrow granulocytes via Toll-like receptor 2. *J. Cell Sci.* **116**, 293–302 (2003).

166. Albitar-Nehme, S. *et al.* Comparison of lipopolysaccharide structures of Bordetella pertussis clinical isolates from pre- and post-vaccine era. *Carbohydr. Res.* **378**, 56–62 (2013).

167. Sándor, V. *et al.* NACE-ESI-MS/MS method for separation and characterization of phosphorylation and acylation isomers of lipid A. *Electrophoresis* **41**, 1178–1188 (2020).

168. van Belkum, A. *et al.* Guidelines for the validation and application of typing methods for use in bacterial epidemiology. *Clin. Microbiol. Infect.* **13**, 1–46 (2007).

169. Stenutz, R., Weintraub, A. & Widmalm, G. The structures of Escherichia coli O-polysaccharide antigens. *FEMS Microbiol. Rev.* **30**, 382–403 (2006).

170. Wang, X. & Quinn, P. J. Lipopolysaccharide: Biosynthetic pathway and structure modification. *Prog. Lipid Res.* **49**, 97–107 (2010).

171. Henderson, J. C. *et al.* The Power of Asymmetry: Architecture and Assembly of the Gram-Negative Outer Membrane Lipid Bilayer. *Annu. Rev. Microbiol.* **70**, 255–278 (2016).

172. Nascimento, A., Pontes, F. J. S., Lins, R. D. & Soares, T. A. Hydration, ionic valence and cross-linking propensities of cations determine the stability of lipopolysaccharide (LPS) membranes. *Chem. Commun.* **50**, 231–233 (2013).

173. Delcour, A. H. Outer membrane permeability and antibiotic resistance. *Biochim. Biophys. Acta* **1794**, 808–816 (2009).

174. O'Shea, R. & Moser, H. E. Physicochemical properties of antibacterial compounds: implications for drug discovery. *J. Med. Chem.* **51**, 2871–2878 (2008).

175. Pagnout, C. *et al.* Pleiotropic effects of rfa-gene mutations on Escherichia coli envelope properties. *Sci. Rep.* **9**, 9696 (2019).

176. Geurtsen, J., Steeghs, L., Hove, J. T., van der Ley, P. & Tommassen, J. Dissemination of lipid A deacylases (pagL) among gram-negative bacteria: identification of active-site histidine and serine residues. *J. Biol. Chem.* **280**, 8248–8259 (2005).

177. Heimesaat, M. M., Backert, S., Alter, T. & Bereswill, S. Molecular Targets in Campylobacter Infections. *Biomolecules* **13**, 409 (2023).

178. Mi, W. *et al.* Structural basis of MsbA-mediated lipopolysaccharide transport. *Nature* **549**, 233–237 (2017).

179. Ward, A., Reyes, C. L., Yu, J., Roth, C. B. & Chang, G. Flexibility in the ABC transporter MsbA: Alternating access with a twist. *Proc. Natl. Acad. Sci. U. S. A.* **104**, 19005–19010 (2007).

180. Tang, X. *et al.* Cryo-EM structures of lipopolysaccharide transporter LptB2FGC in lipopolysaccharide or AMP-PNP-bound states reveal its transport mechanism. *Nat. Commun.* **10**, 4175 (2019).

181. Dong, H., Zhang, Z., Tang, X., Paterson, N. G. & Dong, C. Structural and functional insights into the lipopolysaccharide ABC transporter LptB2FG. *Nat. Commun.* **8**, 222 (2017).

182. Li, Y., Orlando, B. J. & Liao, M. Structural basis of lipopolysaccharide extraction by the LptB2FGC complex. *Nature* **567**, 486–490 (2019).

183. Sperandeo, P., Martorana, A. M. & Polissi, A. Lipopolysaccharide Biosynthesis and Transport to the Outer Membrane of Gram-Negative Bacteria. *Subcell. Biochem.* **92**, 9–37 (2019).

184. Botos, I. *et al.* Structural and Functional Characterization of the LPS Transporter LptDE from Gram-negative Pathogens. *Struct. Lond. Engl. 1993* **24**, 965–976 (2016).

185. Lundquist, K. P. & Gumbart, J. C. Presence of substrate aids lateral gate separation in LptD. *Biochim. Biophys. Acta Biomembr.* **1862**, 183025 (2020).

186. Dong, H. *et al.* Structural basis for outer membrane lipopolysaccharide insertion. *Nature* **511**, 52–56 (2014).

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

188. Hu, K.-Y. & Saier Jr., M. H. Bioinformatic analyses of gram-negative bacterial OstA outer membrane assembly homologues. *Curr. Genomics* **7**, 447–461 (2006).

189. Weiss, J., Hutzler, M. & Kao, L. Environmental modulation of lipopolysaccharide chain length alters the sensitivity of Escherichia coli to the neutrophil bactericidal/permeability-increasing protein. *Infect. Immun.* **51**, 594–599 (1986).

190. Lau, P. C. Y., Lindhout, T., Beveridge, T. J., Dutcher, J. R. & Lam, J. S. Differential Lipopolysaccharide Core Capping Leads to Quantitative and Correlated Modifications of Mechanical and Structural Properties in Pseudomonas aeruginosa Biofilms. *J. Bacteriol.* **191**, 6618–6631 (2009).

191. Palmer, L. D. *et al.* Modulating Isoprenoid Biosynthesis Increases Lipooligosaccharides and Restores Acinetobacter baumannii Resistance to Host and Antibiotic Stress. *Cell Rep.* **32**, 108129 (2020).

192. Low, W.-Y., Thong, S. & Chng, S.-S. ATP disrupts lipid-binding equilibrium to drive retrograde transport critical for bacterial outer membrane asymmetry. *Proc. Natl. Acad. Sci. U. S. A.* **118**, e2110055118 (2021).

193. Hughes, G. W. *et al.* Evidence for phospholipid export from the bacterial inner membrane by the Mla ABC transport system. *Nat. Microbiol.* **4**, 1692–1705 (2019).

194. Bayer, M. E. Areas of adhesion between wall and membrane of Escherichia coli. *J. Gen. Microbiol.* **53**, 395–404 (1968).

195. Yeow, J. & Chng, S.-S. Of zones, bridges and chaperones - phospholipid transport in bacterial outer membrane assembly and homeostasis. *Microbiol. Read. Engl.* **168**, (2022).

196. Chong, Z.-S., Woo, W.-F. & Chng, S.-S. Osmoporin OmpC forms a complex with MlaA to maintain outer membrane lipid asymmetry in Escherichia coli. *Mol. Microbiol.* **98**, 1133–1146 (2015).

197. Arruda, S., Bomfim, G., Knights, R., Huima-Byron, T. & Riley, L. W. Cloning of an M. tuberculosis DNA fragment associated with entry and survival inside cells. *Science* **261**, 1454–1457 (1993).

198. Nakayama, T. & Zhang-Akiyama, Q.-M. pqiABC and yebST, Putative mce Operons of Escherichia coli, Encode Transport Pathways and Contribute to Membrane Integrity. *J. Bacteriol.* **199**, e00606-16 (2017).

199. Isom, G. L. *et al.* MCE domain proteins: conserved inner membrane lipid-binding proteins required for outer membrane homeostasis. *Sci. Rep.* **7**, 8608 (2017).

200. Koh, Y. S. & Roe, J. H. Dual regulation of the paraquat-inducible gene pqi-5 by SoxS and RpoS in Escherichia coli. *Mol. Microbiol.* **22**, 53–61 (1996).

201. Isom, G. L. *et al.* LetB Structure Reveals a Tunnel for Lipid Transport across the Bacterial Envelope. *Cell* **181**, 653-664.e19 (2020).

202. Liu, C., Ma, J., Wang, J., Wang, H. & Zhang, L. Cryo-EM Structure of a Bacterial Lipid Transporter YebT. *J. Mol. Biol.* **432**, 1008–1019 (2020).

203. Ekiert, D. C. *et al.* Architectures of Lipid Transport Systems for the Bacterial Outer Membrane. *Cell* **169**, 273-285.e17 (2017).

204. Gerding, M. A., Ogata, Y., Pecora, N. D., Niki, H. & de Boer, P. A. J. The trans-envelope Tol-Pal complex is part of the cell division machinery and required for proper outer-membrane invagination during cell constriction in E. coli. *Mol. Microbiol.* **63**, 1008–1025 (2007).

205. Shrivastava, R., Jiang, X. & Chng, S.-S. Outer membrane lipid homeostasis via retrograde phospholipid transport in Escherichia coli. *Mol. Microbiol.* **106**, 395–408 (2017).

206. Bonsor, D. A. *et al.* Allosteric beta-propeller signalling in TolB and its manipulation by translocating colicins. *EMBO J.* **28**, 2846–2857 (2009).

207. Deprez, C. *et al.* Solution structure of the E.coli TolA C-terminal domain reveals conformational changes upon binding to the phage g3p N-terminal domain. *J. Mol. Biol.* **346**, 1047–1057 (2005).

208. Josts, I. *et al.* The Structure of a Conserved Domain of TamB Reveals a Hydrophobic β Taco Fold. *Struct. Lond. Engl. 1993* **25**, 1898-1906.e5 (2017).

209. Ruiz, N., Davis, R. M. & Kumar, S. YhdP, TamB, and YdbH Are Redundant but Essential for Growth and Lipid Homeostasis of the Gram-Negative Outer Membrane. *mBio* **12**, e0271421 (2021).

210. Zhai, L. *et al.* Novel sterol binding domains in bacteria. 2022.05.15.491920 Preprint at https://doi.org/10.1101/2022.05.15.491920 (2023).

211. Grabowicz, M. Lipoproteins and Their Trafficking to the Outer Membrane. *EcoSal Plus* **8**, (2019).

212. Mao, G. *et al.* Crystal structure of E. coli lipoprotein diacylglyceryl transferase. *Nat. Commun.* **7**, 10198 (2016).

213. Vogeley, L. *et al.* Structural basis of lipoprotein signal peptidase II action and inhibition by the antibiotic globomycin. *Science* **351**, 876–880 (2016).

214. Wiseman, B. & Högbom, M. Conformational changes in Apolipoprotein N-acyltransferase (Lnt). *Sci. Rep.* **10**, 639 (2020).

215. Narita, S.-I. & Tokuda, H. Bacterial lipoproteins; biogenesis, sorting and quality control. *Biochim. Biophys. Acta Mol. Cell Biol. Lipids* **1862**, 1414–1423 (2017).

216. Sharma, S. *et al.* Mechanism of LolCDE as a molecular extruder of bacterial triacylated lipoproteins. *Nat. Commun.* **12**, 4687 (2021).

217. Taniguchi, N., Matsuyama, S.-I. & Tokuda, H. Mechanisms underlying energy-independent transfer of lipoproteins from LolA to LolB, which have similar unclosed {beta}-barrel structures. *J. Biol. Chem.* **280**, 34481–34488 (2005).

218. Webb, C. T. *et al.* Dynamic association of BAM complex modules includes surface exposure of the lipoprotein BamC. *J. Mol. Biol.* **422**, 545–555 (2012).

219. Dong, C. *et al.* Wza the translocon for E. coli capsular polysaccharides defines a new class of membrane protein. *Nature* **444**, 226–229 (2006).

220. Schulz, G. E. The structure of bacterial outer membrane proteins. *Biochim. Biophys. Acta* **1565**, 308–317 (2002).

221. Schulz, G. E. beta-Barrel membrane proteins. *Curr. Opin. Struct. Biol.* **10**, 443–447 (2000).

222. Lauber, F., Deme, J. C., Lea, S. M. & Berks, B. C. Type 9 secretion system structures reveal a new protein transport mechanism. *Nature* **564**, 77–82 (2018).

223. Yan, Z., Yin, M., Xu, D., Zhu, Y. & Li, X. Structural insights into the secretin translocation channel in the type II secretion system. *Nat. Struct. Mol. Biol.* **24**, 177–183 (2017).

224. Slusky, J. S. G. & Dunbrack, R. L. Charge asymmetry in the proteins of the outer membrane. *Bioinforma. Oxf. Engl.* **29**, 2122–2128 (2013).

225. Schiffrin, B. *et al.* Effects of Periplasmic Chaperones and Membrane Thickness on BamA-Catalyzed Outer-Membrane Protein Folding. *J. Mol. Biol.* **429**, 3776–3792 (2017).

226. Koebnik, R., Locher, K. P. & Van Gelder, P. Structure and function of bacterial outer membrane proteins: barrels in a nutshell. *Mol. Microbiol.* **37**, 239–253 (2000).

227. Wu, R. *et al.* Plasticity within the barrel domain of BamA mediates a hybrid-barrel mechanism by BAM. *Nat. Commun.* **12**, 7131 (2021).

228. Doyle, M. T. *et al.* Cryo-EM structures reveal multiple stages of bacterial outer membrane protein folding. *Cell* **185**, 1143-1156.e13 (2022).

229. Xiao, L. *et al.* Structures of the β-barrel assembly machine recognizing outer membrane protein substrates. *FASEB J. Off. Publ. Fed. Am. Soc. Exp. Biol.* **35**, e21207 (2021).

230. Auclair, S. M., Bhanu, M. K. & Kendall, D. A. Signal peptidase I: cleaving the way to mature proteins. *Protein Sci. Publ. Protein Soc.* **21**, 13–25 (2012).

231. Mas, G., Thoma, J. & Hiller, S. The Periplasmic Chaperones Skp and SurA. *Subcell. Biochem.* **92**, 169–186 (2019).

232. Devlin, T. *et al.* FkpA enhances membrane protein folding using an extensive interaction surface. *Protein Sci. Publ. Protein Soc.* **32**, e4592 (2023).

233. Chang, Z. The function of the DegP (HtrA) protein: Protease versus chaperone. *IUBMB Life* **68**, 904–907 (2016).

234. Zhang, S. *et al.* Degp degrades a wide range of substrate proteins in Escherichia coli under stress conditions. *Biochem. J.* **476**, 3549–3564 (2019).

235. Schiffrin, B. *et al.* Dynamic interplay between the periplasmic chaperone SurA and the BAM complex in outer membrane protein folding. *Commun. Biol.* **5**, 560 (2022).

236. Marx, D. C. *et al.* SurA is a cryptically grooved chaperone that expands unfolded outer membrane proteins. *Proc. Natl. Acad. Sci. U. S. A.* **117**, 28026–28035 (2020).

237. Bitto, E. & McKay, D. B. The Periplasmic Molecular Chaperone Protein SurA Binds a Peptide Motif That Is Characteristic of Integral Outer Membrane Proteins *. *J. Biol. Chem.* **278**, 49316–49322 (2003).

238. Schiffrin, B. *et al.* Skp is a multivalent chaperone of outer-membrane proteins. *Nat. Struct. Mol. Biol.* **23**, 786–793 (2016).

239. Chamachi, N. *et al.* Chaperones Skp and SurA dynamically expand unfolded OmpX and synergistically disassemble oligomeric aggregates. *Proc. Natl. Acad. Sci. U. S. A.* **119**, e2118919119 (2022).

240. Yang, C., Peng, S., Chen, C. & Zhao, X. S. Molecular mechanism of networking among DegP, Skp and SurA in periplasm for biogenesis of outer membrane proteins. *Biochem. J.* **477**, 2949–2965 (2020).

241. Combs, A. N. & Silhavy, T. J. The sacrificial adaptor protein Skp functions to remove stalled substrates from the β-barrel assembly machine. *Proc. Natl. Acad. Sci.* **119**, e2114997119 (2022).

242. He, W. *et al.* Chaperone Spy Protects Outer Membrane Proteins from Folding Stress via Dynamic Complex Formation. *mBio* **12**, e0213021 (2021).

243. Han, L. *et al.* Structure of the BAM complex and its implications for biogenesis of outer-membrane proteins. *Nat. Struct. Mol. Biol.* **23**, 192–196 (2016).

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

245. Browning, D. F. *et al.* Cross-species chimeras reveal BamA POTRA and β-barrel domains must be fine-tuned for efficient OMP insertion. *Mol. Microbiol.* **97**, 646–659 (2015).

246. Thoma, J., Sun, Y., Ritzmann, N. & Müller, D. J. POTRA Domains, Extracellular Lid, and Membrane Composition Modulate the Conformational Stability of the β Barrel Assembly Factor BamA. *Struct. Lond. Engl. 1993* **26**, 987-996.e3 (2018).

247. Malinverni, J. C. *et al.* YfiO stabilizes the YaeT complex and is essential for outer membrane protein assembly in Escherichia coli. *Mol. Microbiol.* **61**, 151–164 (2006).

248. Tata, M. & Konovalova, A. Improper Coordination of BamA and BamD Results in Bam Complex Jamming by a Lipoprotein Substrate. *mBio* **10**, e00660-19 (2019).

249. Hagan, C. L., Wzorek, J. S. & Kahne, D. Inhibition of the β-barrel assembly machine by a peptide that binds BamD. *Proc. Natl. Acad. Sci.* **112**, 2011–2016 (2015).

250. Kumar, S. & Konovalova, A. BamE directly interacts with BamA and BamD coordinating their functions. *Mol. Microbiol.* (2023) doi:10.1111/mmi.15127.

251. Thewasano, N., Germany, E. M., Maruno, Y., Nakajima, Y. & Shiota, T. Categorization of Escherichia coli outer membrane proteins by dependence on accessory proteins of the β-barrel assembly machinery complex. *J. Biol. Chem.* **299**, 104821 (2023).

252. Gunasinghe, S. D. *et al.* The WD40 Protein BamB Mediates Coupling of BAM Complexes into Assembly Precincts in the Bacterial Outer Membrane. *Cell Rep.* **23**, 2782–2794 (2018).

253. Iadanza, M. G. *et al.* Distortion of the bilayer and dynamics of the BAM complex in lipid nanodiscs. *Commun. Biol.* **3**, 1–14 (2020).

254. Wu, R., Stephenson, R., Gichaba, A. & Noinaj, N. The big BAM theory: An open and closed case. *Biochim. Biophys. Acta Biomembr.* **1862**, 183062 (2020).

255. White, P. *et al.* The role of membrane destabilisation and protein dynamics in BAM catalysed OMP folding. *Nat. Commun.* **12**, 4174 (2021).

256. Noinaj, N., Kuszak, A. J., Balusek, C., Gumbart, J. C. & Buchanan, S. K. Lateral opening and exit pore formation are required for BamA function. *Struct. Lond. Engl. 1993* **22**, 1055–1062 (2014).

257. Konovalova, A., Kahne, D. E. & Silhavy, T. J. Outer Membrane Biogenesis. *Annu. Rev. Microbiol.* **71**, 539–556 (2017).

258. Tomasek, D. *et al.* Structure of a nascent membrane protein as it folds on the BAM complex. *Nature* **583**, 473–478 (2020).

259. Shen, C. *et al.* Structural basis of BAM-mediated outer membrane β-barrel protein assembly. *Nature* **617**, 185–193 (2023).

260. Troman, L., Alvira, S., Daum, B., Gold, V. A. M. & Collinson, I. Interaction of the periplasmic chaperone SurA with the inner membrane protein secretion (SEC) machinery. *Biochem. J.* **480**, 283 (2023).

261. Alvira, S. *et al.* Inter-membrane association of the Sec and BAM translocons for bacterial outer-membrane biogenesis. *eLife* **9**, e60669 (2020).

262. Kleinschmidt, J. H. Folding of β-barrel membrane proteins in lipid bilayers - Unassisted and assisted folding and insertion. *Biochim. Biophys. Acta* **1848**, 1927–1943 (2015).

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

264. Kleinschmidt, J. H., Wiener, M. C. & Tamm, L. K. Outer membrane protein A of E. coli folds into detergent micelles, but not in the presence of monomeric detergent. *Protein Sci.* **8**, 2065–2071 (1999).

265. Burgess, N. K., Dao, T. P., Stanley, A. M. & Fleming, K. G. Beta-barrel proteins that reside in the Escherichia coli outer membrane in vivo demonstrate varied folding behavior in vitro. *J. Biol. Chem.* **283**, 26748–26758 (2008).

266. Danoff, E. J. & Fleming, K. G. Membrane defects accelerate outer membrane β-barrel protein folding. *Biochemistry* **54**, 97–99 (2015).

267. Patel, G. J. & Kleinschmidt, J. H. The lipid bilayer-inserted membrane protein BamA of Escherichia coli facilitates insertion and folding of outer membrane protein A from its complex with Skp. *Biochemistry* **52**, 3974–3986 (2013).

268. Gessmann, D. *et al.* Outer membrane β-barrel protein folding is physically controlled by periplasmic lipid head groups and BamA. *Proc. Natl. Acad. Sci. U. S. A.* **111**, 5878–5883 (2014).

269. Tiwari, P. B. & Mahalakshmi, R. Interplay of protein primary sequence, lipid membrane, and chaperone in β-barrel assembly. *Protein Sci. Publ. Protein Soc.* **30**, 624–637 (2021).

270. McMorran, L. M., Bartlett, A. I., Huysmans, G. H. M., Radford, S. E. & Brockwell, D. J. Dissecting the Effects of Periplasmic Chaperones on the In Vitro Folding of the Outer Membrane Protein PagP. *J. Mol. Biol.* **425**, 3178–3191 (2013).

271. Jansen, C., Heutink, M., Tommassen, J. & de Cock, H. The assembly pathway of outer membrane protein PhoE of Escherichia coli. *Eur. J. Biochem.* **267**, 3792–3800 (2000).

272. Reid, J., Fung, H., Gehring, K., Klebba, P. E. & Nikaido, H. Targeting of porin to the outer membrane of Escherichia coli. Rate of trimer assembly and identification of a dimer intermediate. *J. Biol. Chem.* **263**, 7753–7759 (1988).

273. Li, G.-W., Burkhardt, D., Gross, C. & Weissman, J. S. Quantifying absolute protein synthesis rates reveals principles underlying allocation of cellular resources. *Cell* **157**, 624–635 (2014).

274. Sun, J., Rutherford, S. T., Silhavy, T. J. & Huang, K. C. Physical properties of the bacterial outer membrane. *Nat. Rev. Microbiol.* **20**, 236–248 (2022).

275. Benn, G. *et al.* Phase separation in the outer membrane of Escherichia coli. *Proc. Natl. Acad. Sci.* **118**, e2112237118 (2021).

276. Manioglu, S. *et al.* Antibiotic polymyxin arranges lipopolysaccharide into crystalline structures to solidify the bacterial membrane. *Nat. Commun.* **13**, 6195 (2022).

277. Pratt, L. A., Hsing, W., Gibson, K. E. & Silhavy, T. J. From acids to osmZ: multiple factors influence synthesis of the OmpF and OmpC porins in Escherichia coli. *Mol. Microbiol.* **20**, 911–917 (1996).

278. Marzoa, J. *et al.* Analysis of outer membrane porin complexes of Neisseria meningitidis in wild-type and specific knock-out mutant strains. *Proteomics* **9**, 648–656 (2009).

279. Rocker, A. *et al.* Global Trends in Proteome Remodeling of the Outer Membrane Modulate Antimicrobial Permeability in Klebsiella pneumoniae. *mBio* **11**, e00603-20 (2020).

280. Guan, Q., Bhowmick, B., Upadhyay, A. & Han, Q. Structure and Functions of Bacterial Outer Membrane Protein A, A Potential Therapeutic Target for Bacterial Infection. *Curr. Top. Med. Chem.* **21**, 1129–1138 (2021).

281. Vogel, H. & Jähnig, F. Models for the structure of outer-membrane proteins of Escherichia coli derived from raman spectroscopy and prediction methods. *J. Mol. Biol.* **190**, 191–199 (1986).

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

283. Pautsch, A. & Schulz, G. E. High-resolution structure of the OmpA membrane domain. *J. Mol. Biol.* **298**, 273–282 (2000).

284. Ishida, H., Garcia-Herrero, A. & Vogel, H. J. The periplasmic domain of Escherichia coli outer membrane protein A can undergo a localized temperature dependent structural transition. *Biochim. Biophys. Acta* **1838**, 3014–3024 (2014).

285. Grizot, S. & Buchanan, S. K. Structure of the OmpA-like domain of RmpM from Neisseria meningitidis. *Mol. Microbiol.* **51**, 1027–1037 (2004).

286. Samsudin, F., Ortiz-Suarez, M. L., Piggot, T. J., Bond, P. J. & Khalid, S. OmpA: A Flexible Clamp for Bacterial Cell Wall Attachment. *Struct. Lond. Engl. 1993* **24**, 2227–2235 (2016).

287. Park, J. S. *et al.* Mechanism of anchoring of OmpA protein to the cell wall peptidoglycan of the gram-negative bacterial outer membrane. *FASEB J. Off. Publ. Fed. Am. Soc. Exp. Biol.* **26**, 219–228 (2012).

288. Choi, U. & Lee, C.-R. Distinct Roles of Outer Membrane Porins in Antibiotic Resistance and Membrane Integrity in Escherichia coli. *Front. Microbiol.* **10**, 953 (2019).

289. Marcoux, J. *et al.* Mass Spectrometry Defines the C-Terminal Dimerization Domain and Enables Modeling of the Structure of Full-Length OmpA. *Structure* **22**, 781–790 (2014).

290. Arora, A., Abildgaard, F., Bushweller, J. H. & Tamm, L. K. Structure of outer membrane protein A transmembrane domain by NMR spectroscopy. *Nat. Struct. Biol.* **8**, 334–338 (2001).

291. Sugawara, E. & Nikaido, H. Pore-forming activity of OmpA protein of Escherichia coli. *J. Biol. Chem.* **267**, 2507–2511 (1992).

292. Sugawara, E., Steiert, M., Rouhani, S. & Nikaido, H. Secondary structure of the outer membrane proteins OmpA of Escherichia coli and OprF of Pseudomonas aeruginosa. *J. Bacteriol.* **178**, 6067–6069 (1996).

293. Bond, P. J., Faraldo-Gómez, J. D. & Sansom, M. S. P. OmpA: a pore or not a pore? Simulation and modeling studies. *Biophys. J.* **83**, 763–775 (2002).

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

295. Reusch, R. N. Insights into the structure and assembly of Escherichia coli outer membrane protein A. *FEBS J.* **279**, 894–909 (2012).

296. Stathopoulos, C. An alternative topological model for Escherichia coli OmpA. *Protein Sci. Publ. Protein Soc.* **5**, 170–173 (1996).

297. Arora, A., Rinehart, D., Szabo, G. & Tamm, L. K. Refolded outer membrane protein A of Escherichia coli forms ion channels with two conductance states in planar lipid bilayers. *J. Biol. Chem.* **275**, 1594–1600 (2000).

298. Zakharian, E. & Reusch, R. N. Kinetics of folding of Escherichia coli OmpA from narrow to large pore conformation in a planar bilayer. *Biochemistry* **44**, 6701–6707 (2005).

299. Zakharian, E. & Reusch, R. N. Outer membrane protein A of Escherichia coli forms temperature-sensitive channels in planar lipid bilayers. *FEBS Lett.* **555**, 229–235 (2003).

300. Hounsome, J. D. A. *et al.* Outer membrane protein a of bovine and ovine isolates of Mannheimia haemolytica is surface exposed and contains host Species-Specific Epitopes. *Infect. Immun.* **79**, 4332–4341 (2011).

301. Koebnik, R. Structural and functional roles of the surface-exposed loops of the beta-barrel membrane protein OmpA from Escherichia coli. *J. Bacteriol.* **181**, 3688–3694 (1999).

302. Schwarz, H., Riede, I., Sonntag, I. & Henning, U. Degrees of relatedness of T-even type E. coli phages using different or the same receptors and topology of serologically cross-reacting sites. *EMBO J.* **2**, 375–380 (1983).

303. Smajs, D., Pilsl, H. & Braun, V. Colicin U, a novel colicin produced by Shigella boydii. *J. Bacteriol.* **179**, 4919–4928 (1997).

304. Klimke, W. A. & Frost, L. S. Genetic analysis of the role of the transfer gene, traN, of the F and R100-1 plasmids in mating pair stabilization during conjugation. *J. Bacteriol.* **180**, 4036–4043 (1998).

305. Klimke, W. A. *et al.* The mating pair stabilization protein, TraN, of the F plasmid is an outer-membrane protein with two regions that are important for its function in conjugation. *Microbiol. Read. Engl.* **151**, 3527–3540 (2005).

306. Weiser, J. N. & Gotschlich, E. C. Outer membrane protein A (OmpA) contributes to serum resistance and pathogenicity of Escherichia coli K-1. *Infect. Immun.* **59**, 2252–2258 (1991).

307. Kim, K. S. Strategy of Escherichia coli for crossing the blood-brain barrier. *J. Infect. Dis.* **186 Suppl 2**, S220-224 (2002).

308. Torres, A. G. & Kaper, J. B. Multiple elements controlling adherence of enterohemorrhagic Escherichia coli O157:H7 to HeLa cells. *Infect. Immun.* **71**, 4985–4995 (2003).

309. Torres, A. G., Jeter, C., Langley, W. & Matthysse, A. G. Differential binding of Escherichia coli O157:H7 to alfalfa, human epithelial cells, and plastic is mediated by a variety of surface structures. *Appl. Environ. Microbiol.* **71**, 8008–8015 (2005).

310. Sukumaran, S. K., Selvaraj, S. K. & Prasadarao, N. V. Inhibition of apoptosis by Escherichia coli K1 is accompanied by increased expression of BclXL and blockade of mitochondrial cytochrome c release in macrophages. *Infect. Immun.* **72**, 6012–6022 (2004).

311. Sukumaran, S. K., Shimada, H. & Prasadarao, N. V. Entry and intracellular replication of Escherichia coli K1 in macrophages require expression of outer membrane protein A. *Infect. Immun.* **71**, 5951–5961 (2003).

312. Selvaraj, S. K. & Prasadarao, N. V. Escherichia coli K1 inhibits proinflammatory cytokine induction in monocytes by preventing NF-kappaB activation. *J. Leukoc. Biol.* **78**, 544–554 (2005).

313. Orme, R., Douglas, C. W. I., Rimmer, S. & Webb, M. Proteomic analysis of Escherichia coli biofilms reveals the overexpression of the outer membrane protein OmpA. *Proteomics* **6**, 4269–4277 (2006).

314. Barrios, A. F. G., Zuo, R., Ren, D. & Wood, T. K. Hha, YbaJ, and OmpA regulate Escherichia coli K12 biofilm formation and conjugation plasmids abolish motility. *Biotechnol. Bioeng.* **93**, 188–200 (2006).

315. Shah, C., Hari-Dass, R. & Raynes, J. G. Serum amyloid A is an innate immune opsonin for Gram-negative bacteria. *Blood* **108**, 1751–1757 (2006).

316. Belaaouaj, A., Kim, K. S. & Shapiro, S. D. Degradation of outer membrane protein A in Escherichia coli killing by neutrophil elastase. *Science* **289**, 1185–1188 (2000).

317. Soulas, C. *et al.* Outer membrane protein A (OmpA) binds to and activates human macrophages. *J. Immunol. Baltim. Md 1950* **165**, 2335–2340 (2000).

318. Jeannin, P. *et al.* Outer membrane protein A (OmpA): a new pathogen-associated molecular pattern that interacts with antigen presenting cells-impact on vaccine strategies. *Vaccine* **20 Suppl 4**, A23-27 (2002).

319. Nilsson, G., Belasco, J. G., Cohen, S. N. & von Gabain, A. Growth-rate dependent regulation of mRNA stability in Escherichia coli. *Nature* **312**, 75–77 (1984).

320. Lugtenberg, B., Peters, R., Bernheimer, H. & Berendsen, W. Influence of cultural conditions and mutations on the composition of the outer membrane proteins of Escherichia coli. *Mol. Gen. Genet. MGG* **147**, 251–262 (1976).

321. Vytvytska, O. *et al.* Host factor I, Hfq, binds to Escherichia coli ompA mRNA in a growth rate-dependent fashion and regulates its stability. *Proc. Natl. Acad. Sci. U. S. A.* **95**, 14118–14123 (1998).

322. Udekwu, K. I. *et al.* Hfq-dependent regulation of OmpA synthesis is mediated by an antisense RNA. *Genes Dev.* **19**, 2355–2366 (2005).

323. Smith, S. G. J., Mahon, V., Lambert, M. A. & Fagan, R. P. A molecular Swiss army knife: OmpA structure, function and expression. *FEMS Microbiol. Lett.* **273**, 1–11 (2007).

324. SCHWEIZER, M., HINDENNACH, I., GARTEN, W. & HENNING, U. Major Proteins of the Escherichia coli Outer Cell Envelope Membrane. Interaction of Protein II* with Lipopolysaccharide. *Eur. J. Biochem.* **82**, 211–217 (1978).

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

326. Confer, A. W. & Ayalew, S. The OmpA family of proteins: roles in bacterial pathogenesis and immunity. *Vet. Microbiol.* **163**, 207–222 (2013).

327. Nie, D. *et al.* Outer membrane protein A (OmpA) as a potential therapeutic target for Acinetobacter baumannii infection. *J. Biomed. Sci.* **27**, 26 (2020).

328. Gaddy, J. A., Tomaras, A. P. & Actis, L. A. The Acinetobacter baumannii 19606 OmpA protein plays a role in biofilm formation on abiotic surfaces and in the interaction of this pathogen with eukaryotic cells. *Infect. Immun.* **77**, 3150–3160 (2009).

329. Choi, C. H. *et al.* Acinetobacter baumannii outer membrane protein A targets the nucleus and induces cytotoxicity. *Cell. Microbiol.* **10**, 309–319 (2008).

330. Jin, J. S. *et al.* Acinetobacter baumannii secretes cytotoxic outer membrane protein A via outer membrane vesicles. *PloS One* **6**, e17027 (2011).

331. Yun, S. H. *et al.* Antibiotic treatment modulates protein components of cytotoxic outer membrane vesicles of multidrug-resistant clinical strain, Acinetobacter baumannii DU202. *Clin. Proteomics* **15**, 28 (2018).

332. De Mot, R. & Vanderleyden, J. The C-terminal sequence conservation between OmpA-related outer membrane proteins and MotB suggests a common function in both gram-positive and gram-negative bacteria, possibly in the interaction of these domains with peptidoglycan. *Mol. Microbiol.* **12**, 333–334 (1994).

333. Molloy, M. P. *et al.* Proteomic analysis of the Escherichia coli outer membrane. *Eur. J. Biochem.* **267**, 2871–2881 (2000).

334. Hays, M. P., Kumar, A., Martinez-Becerra, F. J. & Hardwidge, P. R. Immunization with the MipA, Skp, or ETEC_2479 Antigens Confers Protection against Enterotoxigenic E. coli Strains Expressing Different Colonization Factors in a Mouse Pulmonary Challenge Model. *Front. Cell. Infect. Microbiol.* **6**, 181 (2016).

335. Kumar, A. *et al.* Protective Enterotoxigenic Escherichia coli Antigens in a Murine Intranasal Challenge Model. *PLoS Negl. Trop. Dis.* **9**, e0003924 (2015).

336. Vollmer, W., von Rechenberg, M. & Höltje, J. V. Demonstration of molecular interactions between the murein polymerase PBP1B, the lytic transglycosylase MltA, and the scaffolding protein MipA of Escherichia coli. *J. Biol. Chem.* **274**, 6726–6734 (1999).

337. Meisel, U., Höltje, J.-V. & Vollmer, W. Overproduction of inactive variants of the murein synthase PBP1B causes lysis in Escherichia coli. *J. Bacteriol.* **185**, 5342–5348 (2003).

338. Thompson, K. M., Rhodius, V. A. & Gottesman, S. SigmaE regulates and is regulated by a small RNA in Escherichia coli. *J. Bacteriol.* **189**, 4243–4256 (2007).

339. Li, H., Zhang, D., Lin, X. & Peng, X. Outer membrane proteomics of kanamycin-resistant Escherichia coli identified MipA as a novel antibiotic resistance-related protein. *FEMS Microbiol. Lett.* **362**, fnv074 (2015).

340. Rivas, L., Fegan, N. & Dykes, G. A. Expression and putative roles in attachment of outer membrane proteins of Escherichia coli O157 from planktonic and sessile culture. *Foodborne Pathog. Dis.* **5**, 155–164 (2008).

341. Elad, T., Benovich, E., Magrisso, S. & Belkin, S. Toxicant identification by a luminescent bacterial bioreporter panel: application of pattern classification algorithms. *Environ. Sci. Technol.* **42**, 8486–8491 (2008).

342. Abdallah, F. B. *et al.* Identification of Outer Membrane Proteins Altered in Response to UVC-Radiation in Vibrio parahaemolyticus and Vibrio alginolyticus. *Indian J. Microbiol.* **52**, 660–665 (2012).

343. Ben Abdallah, F. *et al.* Identification of outer membrane proteins of Vibrio parahaemolyticus and Vibrio alginolyticus altered in response to γ-irradiation or long-term starvation. *Res. Microbiol.* **161**, 869–875 (2010).

344. Janet-Maitre, M. *et al.* Pseudomonas aeruginosa MipA-MipB envelope proteins act as new sensors of polymyxin. 2023.08.14.553335 Preprint at https://doi.org/10.1101/2023.08.14.553335 (2023).

345. Nilaweera, T. D., Nyenhuis, D. A. & Cafiso, D. S. Structural intermediates observed only in intact Escherichia coli indicate a mechanism for TonB-dependent transport. *eLife* **10**, e68548 (2021).

346. Klebba, P. E. ROSET Model of TonB Action in Gram-Negative Bacterial Iron Acquisition. *J. Bacteriol.* **198**, 1013–1021 (2016).

347. Grinter, R. *et al.* Protease-associated import systems are widespread in Gram-negative bacteria. *PLoS Genet.* **15**, e1008435 (2019).

348. Subashchandrabose, S., Smith, S. N., Spurbeck, R. R., Kole, M. M. & Mobley, H. L. T. Genome-wide detection of fitness genes in uropathogenic Escherichia coli during systemic infection. *PLoS Pathog.* **9**, e1003788 (2013).

349. Wurpel, D. J., Moriel, D. G., Totsika, M., Easton, D. M. & Schembri, M. A. Comparative analysis of the uropathogenic Escherichia coli surface proteome by tandem mass-spectrometry of artificially induced outer membrane vesicles. *J. Proteomics* **115**, 93–106 (2015).

350. Wojnowska, M. & Walker, D. FusB Energizes Import across the Outer Membrane through Direct Interaction with Its Ferredoxin Substrate. *mBio* **11**, e02081-20 (2020).

351. Grinter, R. *et al.* Structure of the bacterial plant-ferredoxin receptor FusA. *Nat. Commun.* **7**, 13308 (2016).

352. Schnaitman, C. A. Outer membrane proteins of Escherichia coli: I. Effect of preparative conditions on the migration of protein in polyacrylamide gels. *Arch. Biochem. Biophys.* **157**, 541–552 (1973).

353. Fiss, E. H., Hollifield, W. C. & Neilands, J. B. Absence of ferric enterobactin receptor modification activity in mutants of Escherichia coli K-12 lacking protein a. *Biochem. Biophys. Res. Commun.* **91**, 29–34 (1979).

354. Sugimura, K. & Higashi, N. A novel outer-membrane-associated protease in Escherichia coli. *J. Bacteriol.* **170**, 3650–3654 (1988).

355. Sugimura, K. & Nishihara, T. Purification, characterization, and primary structure of Escherichia coli protease VII with specificity for paired basic residues: identity of protease VII and OmpT. *J. Bacteriol.* **170**, 5625–5632 (1988).

356. Lundrigan, M. D. & Webb, R. M. Prevalence of ompT among Escherichia coli isolates of human origin. *FEMS Microbiol. Lett.* **97**, 51–56 (1992).

357. Foxman, B., Zhang, L., Palin, K., Tallman, P. & Marrs, C. F. Bacterial Virulence Characteristics of Escherichia coli Isolates from First-Time Urinary Tract Infection. *J. Infect. Dis.* **171**, 1514–1521 (1995).

358. Webb, R. M. & Lundrigan, M. D. Omptin escherichia col/correlates with severity of disease in urinary tract infections. *Med. Microbiol. Lett.* **5**, 8–14 (1996).

359. Egile, C., D'Hauteville, H., Parsot, C. & Sansonetti, P. J. SopA, the outer membrane protease responsible for polar localization of IcsA in Shigella flexneri. *Mol. Microbiol.* **23**, 1063–1073 (1997).

360. Lähteenmäki, K., Kukkonen, M. & Korhonen, T. K. The Pla surface protease/adhesin of Yersinia pestis mediates bacterial invasion into human endothelial cells. *FEBS Lett.* **504**, 69–72 (2001).

361. Hui, C.-Y. *et al.* Escherichia coli outer membrane protease OmpT confers resistance to urinary cationic peptides. *Microbiol. Immunol.* **54**, 452–459 (2010).

362. Stumpe, S., Schmid, R., Stephens, D. L., Georgiou, G. & Bakker, E. P. Identification of OmpT as the Protease That Hydrolyzes the Antimicrobial Peptide Protamine before It Enters Growing Cells ofEscherichia coli. *J. Bacteriol.* **180**, 4002–4006 (1998).

363. Vandeputte-Rutten, L. *et al.* Crystal structure of the outer membrane protease OmpT from Escherichia coli suggests a novel catalytic site. *EMBO J.* **20**, 5033–5039 (2001).

364. Dekker, N., Cox, R. C., Kramer, R. A. & Egmond, M. R. Substrate Specificity of the Integral Membrane Protease OmpT Determined by Spatially Addressed Peptide Libraries. *Biochemistry* **40**, 1694–1701 (2001).

365. Okuno, K., Yabuta, M., Ohsuye, K., Ooi, T. & Kinoshita, S. An analysis of target preferences of Escherichia coli outer-membrane endoprotease OmpT for use in therapeutic peptide production: efficient cleavage of substrates with basic amino acids at the P4 and P6 positions. *Biotechnol. Appl. Biochem.* **36**, 77–84 (2002).

366. OKUNO, K. *et al.* Substrate Specificity at the P1´ Site of Escherichia coli OmpT under Denaturing Conditions. *Biosci. Biotechnol. Biochem.* **66**, 127–134 (2002).

367. McCarter, J. D. *et al.* Substrate Specificity of the Escherichia coli Outer Membrane Protease OmpT. *J. Bacteriol.* **186**, 5919–5925 (2004).

368. Kramer, R. A. *et al.* Identification of essential acidic residues of outer membrane protease OmpT supports a novel active site. *FEBS Lett.* **505**, 426–430 (2001).

369. Baaden, M. & Sansom, M. S. P. OmpT: Molecular Dynamics Simulations of an Outer Membrane Enzyme. *Biophys. J.* **87**, 2942–2953 (2004).

370. Neri, M., Anselmi, C., Carnevale, V., Vargiu, A. V. & Carloni, P. Molecular dynamics simulations of outer-membrane protease T from E. coli based on a hybrid coarse-grained/atomistic potential. *J. Phys. Condens. Matter* **18**, S347 (2006).

371. Neri, M. *et al.* Microseconds Dynamics Simulations of the Outer-Membrane Protease T. *Biophys. J.* **94**, 71–78 (2008).

372. Thomassin, J.-L., Brannon, J. R., Kaiser, J., Gruenheid, S. & Le Moual, H. Enterohemorrhagic and enteropathogenic Escherichia coli evolved different strategies to resist antimicrobial peptides. *Gut Microbes* **3**, 556–561 (2012).

373. Gant Kanegusuku, A., Stankovic, I. N., Cote-Hammarlof, P. A., Yong, P. H. & White-Ziegler, C. A. A Shift to Human Body Temperature (37°C) Rapidly Reprograms Multiple Adaptive Responses in Escherichia coli That Would Facilitate Niche Survival and Colonization. *J. Bacteriol.* **203**, 10.1128/jb.00363-21 (2021).

374. Desloges, I. *et al.* Identification and characterization of OmpT-like proteases in uropathogenic Escherichia coli clinical isolates. *MicrobiologyOpen* **8**, e915 (2019).

375. Cho, Y. H. *et al.* Omptin Proteases of Enterobacterales Show Conserved Regulation by the PhoPQ Two-Component System but Exhibit Divergent Protection from Antimicrobial Host Peptides and Complement. *Infect. Immun.* **91**, e00518-22 (2022).

376. Urashima, A., Sanou, A., Yen, H. & Tobe, T. Enterohaemorrhagic Escherichia coli produces outer membrane vesicles as an active defence system against antimicrobial peptide LL-37. *Cell. Microbiol.* **19**, e12758 (2017).

377. He, X. L. *et al.* Role of uropathogenic Escherichia coli outer membrane protein T in pathogenesis of urinary tract infection. *Pathog. Dis.* **73**, ftv006 (2015).

378. Boroumand, M., Sharifi, A., Ghatei, M. A. & Sadrinasab, M. Evaluation of Biofilm Formation and Virulence Genes and Association with Antibiotic Resistance Patterns of Uropathogenic Escherichia coli Strains in Southwestern Iran. *Jundishapur J. Microbiol.* **14**, (2021).

379. Hejair, H. M. A. *et al.* Role of outer membrane protein T in pathogenicity of avian pathogenic Escherichia coli. *Res. Vet. Sci.* **115**, 109–116 (2017).

380. Duché, D., Issouf, M. & Lloubès, R. Immunity Protein Protects Colicin E2 from OmpT Protease. *J. Biochem. (Tokyo)* **145**, 95–101 (2009).

381. Ferguson, A. D. *et al.* Active transport of an antibiotic rifamycin derivative by the outer-membrane protein FhuA. *Struct. Lond. Engl. 1993* **9**, 707–716 (2001).

382. Kaufmann, A., Stierhof, Y. D. & Henning, U. New outer membrane-associated protease of Escherichia coli K-12. *J. Bacteriol.* **176**, 359–367 (1994).

383. Hwang, B.-Y. *et al.* Substrate Specificity of the Escherichia coli Outer Membrane Protease OmpP. *J. Bacteriol.* **189**, 522–530 (2007).

384. Agarkov, A., Chauhan, S., Lory, P. J., Gilbertson, S. R. & Motin, V. L. Substrate specificity and screening of the integral membrane protease Pla. *Bioorg. Med. Chem. Lett.* **18**, 427–431 (2008).

385. Bianucci, R. *et al.* Technical note: A rapid diagnostic test detects plague in ancient human remains: An example of the interaction between archeological and biological approaches (southeastern France, 16th–18th centuries). *Am. J. Phys. Anthropol.* **136**, 361–367 (2008).

386. Riehm, J. M. *et al.* Detection of Yersinia pestis using real-time PCR in patients with suspected bubonic plague. *Mol. Cell. Probes* **25**, 8–12 (2011).

387. Suomalainen, M. *et al.* Using Every Trick in the Book: The Pla Surface Protease of Yersinia pestis. in *The Genus Yersinia: From Genomics to Function* (eds. Perry, R. D. & Fetherston, J. D.) 268–278 (Springer, 2007). doi:10.1007/978-0-387-72124-8_24.

388. Yun, T. H., Cott, J. E., Tapping, R. I., Slauch, J. M. & Morrissey, J. H. Proteolytic inactivation of tissue factor pathway inhibitor by bacterial omptins. *Blood* **113**, 1139–1148 (2009).

389. Suomalainen, M. *et al.* Temperature-Induced Changes in the Lipopolysaccharide of Yersinia pestis Affect Plasminogen Activation by the Pla Surface Protease. *Infect. Immun.* **78**, 2644–2652 (2010).

390. Kukkonen, M. & Korhonen, T. K. The omptin family of enterobacterial surface proteases/adhesins: from housekeeping in Escherichia coli to systemic spread of Yersinia pestis. *Int. J. Med. Microbiol.* **294**, 7–14 (2004).

391. Brannon, J. R. *et al.* Inhibition of Outer Membrane Proteases of the Omptin Family by Aprotinin. *Infect. Immun.* **83**, 2300–2311 (2015).

392. Brannon, J. R., Thomassin, J.-L., Gruenheid, S. & Le Moual, H. Antimicrobial Peptide Conformation as a Structural Determinant of Omptin Protease Specificity. *J. Bacteriol.* **197**, 3583–3591 (2015).

393. Johnson, J. R. *et al.* Evidence of Commonality between Canine and Human Extraintestinal Pathogenic Escherichia coli Strains That ExpresspapG Allele III. *Infect. Immun.* **68**, 3327–3336 (2000).

394. De Carli, S. *et al.* Virulence gene content in Escherichia coli isolates from poultry flocks with clinical signs of colibacillosis in Brazil. *Poult. Sci.* **94**, 2635–2640 (2015).

395. Tan, C. *et al.* Serotypes and virulence genes of extraintestinal pathogenic Escherichia coli isolates from diseased pigs in China. *Vet. J.* **192**, 483–488 (2012).

396. Tong, C. *et al.* Development of an Indirect Elisa for Detection of E. COLI Antibodies in Cow Serum using A Recombinant Ompt as Antigen. *J. Immunoassay Immunochem.* **35**, 241–255 (2014).

397. Haiko, J., Laakkonen, L., Westerlund-Wikström, B. & Korhonen, T. K. Molecular adaptation of a plant-bacterium outer membrane protease towards plague virulence factor Pla. *BMC Evol. Biol.* **11**, 1–12 (2011).

398. Brok, R. G. P. M. *et al.* Molecular characterization of enterobacterial pldA genes encoding outer membrane phospholipase A. *J. Bacteriol.* **176**, 861–870 (1994).

399. de Jonge, E. F., Vogrinec, L., van Boxtel, R. & Tommassen, J. Inactivation of the Mla system and outer-membrane phospholipase A results in disrupted outer-membrane lipid asymmetry and hypervesiculation in Bordetella pertussis. *Curr. Res. Microb. Sci.* **3**, (2022).

400. Boekema, E. J. *et al.* A 7.4-Å projection structure of outer membrane phospholipase A from Escherichia coli by electron crystallography. *J. Struct. Biol.* **123**, 67–71 (1998).

401. Tan, A. E. & Fleming, K. G. Outer membrane phospholipase A dimer stability does not correlate to occluded surface area. *Biochemistry* **47**, 12095–12103 (2008).

402. Dekker, N., Tommassen, J., Lustig, A., Rosenbusch, J. P. & Verheij, H. M. Dimerization regulates the enzymatic activity of Escherichia coli outer membrane phospholipase A. *J. Biol. Chem.* **272**, 3179–3184 (1997).

403. Ubarretxena-Belandia, I. *et al.* Outer membrane phospholipase A is dimeric in phospholipid bilayers: A cross-linking and fluorescence resonance energy transfer study. *Biochemistry* **38**, 7398–7405 (1999).

404. Snijder, H. J. *et al.* Structural investigations of calcium binding and its role in activity and activation of outer membrane phospholipase A from Escherichia coli. *J. Mol. Biol.* **309**, 477–489 (2001).

405. Brok, R. G. P. M., Dekker, N., Gerrits, N., Verheij, H. M. & Tommassen, J. A Conserved Histidine Residue of Escherichia Coli Outer-Membrane Phospholipase A is Important for Activity. *Eur. J. Biochem.* **234**, 934–938 (1995).

406. Kingma, R. L. *et al.* Unusual catalytic triad of Escherichia coli outer membrane phospholipase A. *Biochemistry* **39**, 10017–10022 (2000).

407. Snijder, H. J. *et al.* Structural investigations of the active-site mutant Asn156A1a of outer membrane phospholipase A: Function of the Asn-His interaction in the catalytic triad. *Protein Sci.* **10**, 1962–1969 (2001).

408. Stanley, A. M., Chuawong, P., Hendrickson, T. L. & Fleming, K. G. Energetics of outer membrane phospholipase A (OMPLA) dimerization. *J. Mol. Biol.* **358**, 120–131 (2006).

409. Horrevoets, A. J. G., Hackeng, T. M., Verheij, H. M., Dijkman, R. & De Haas, G. H. Kinetic characterization of Escherichia coli outer membrane phospholipase A using mixed detergent-lipid micelles. *Biochemistry* **28**, 1139–1147 (1989).

410. Rangl, M. *et al.* Real-time Visualization of Phospholipid Degradation by Outer Membrane Phospholipase A using High-Speed Atomic Force Microscopy. *J. Mol. Biol.* **429**, 977–986 (2017).

411. Stanley, A. M., Treubrodt, A. M., Chuawong, P., Hendrickson, T. L. & Fleming, K. G. Lipid chain selectivity by outer membrane phospholipase A. *J. Mol. Biol.* **366**, 461–468 (2007).

412. Jia, W. *et al.* Lipid trafficking controls endotoxin acylation in outer membranes of Escherichia coli. *J. Biol. Chem.* **279**, 44966–44975 (2004).

413. Powers, M. J. & Trent, M. S. Phospholipid retention in the absence of asymmetry strengthens the outer membrane permeability barrier to last-resort antibiotics. *Proc. Natl. Acad. Sci. U. S. A.* **115**, E8518–E8527 (2018).

414. Langen, G. R., Harper, J. R., Silhavy, T. J. & Howard, S. P. Absence of the outer membrane phospholipase a suppresses the temperature-sensitive phenotype of Escherichia coli degP mutants and induces the Cpx and σEextracytoplasmic stress responses. *J. Bacteriol.* **183**, 5230–5238 (2001).

415. Weiss, J., Beckerdite-Quagliata, S. & Elsbach, P. Determinants of the action of phospholipases A on the envelope phospholipids of Escherichia coli. *J. Biol. Chem.* **254**, 11010–11014 (1979).

416. Hardaway, K. L. & Buller, C. S. Effect of ethylenediaminetetraacetate on phospholipids and outer membrane function in Escherichia coli. *J. Bacteriol.* **137**, 62–68 (1979).

417. Taketo, A. Sensitivity of Escherichia coli to viral nucleic acid. 8. Idiosyncrasy of Ca2+-dependent competence for DNA. *J. Biochem. (Tokyo)* **75**, 895–904 (1974).

418. DiRusso, C. C. & Black, P. N. Bacterial Long Chain Fatty Acid Transport: Gateway to a Fatty Acid-responsive Signaling System*. *J. Biol. Chem.* **279**, 49563–49566 (2004).

419. Homma, H., Nishijima, M., Kobayashi, T., Okuyama, H. & Nojima, S. Incorporation and metabolism of 2-acyl lysophospholipids by Escherich1a coli. *Biochim. Biophys. Acta BBA - Lipids Lipid Metab.* **663**, 1–13 (1981).

420. May, K. L. & Silhavy, T. J. The Escherichia coli Phospholipase PldA Regulates Outer Membrane Homeostasis via Lipid Signaling. *mBio* **9**, e00379-18 (2018).

421. Perumal, P., Raina, R., Narayanan, S. B. & Arockiasamy, A. Crystal structure of calcium bound outer membrane phospholipase A (OmpLA) from Salmonella typhi and in silico anti-microbial screening. 2020.01.08.898262 Preprint at https://doi.org/10.1101/2020.01.08.898262 (2020).

422. Hamidian, M. *et al.* Prevalence of putative virulence markers in campylobacter jejuni and campylobacter coli isolated from hospitalized children, raw chicken, and raw beef in Tehran, Iran. *Can. J. Microbiol.* **57**, 143–148 (2011).

423. Grant, K. A., Belandia, I. U., Dekker, N., Richardson, P. T. & Park, S. F. Molecular characterization of pldA, the structural gene for a phospholipase A from Campylobacter coli, and its contribution to cell- associated hemolysis. *Infect. Immun.* **65**, 1172–1180 (1997).

424. Wysok, B., Wojtacka, J. & Kivistö, R. Pathogenicity of Campylobacter strains of poultry and human origin from Poland. *Int. J. Food Microbiol.* **334**, (2020).

425. Ziprin, R. L. *et al.* Role of Campylobacter jejuni potential virulence genes in cecal colonization. *Avian Dis.* **45**, 549–557 (2001).

426. Cao, X. *et al.* The Unique Phospholipidome of the Enteric Pathogen Campylobacter jejuni: Lysophosholipids Are Required for Motility at Low Oxygen Availability. *J. Mol. Biol.* **432**, 5244–5258 (2020).

427. Xerry, J. & Owen, R. J. Conservation and microdiversity of the phospholipase A (pldA) gene of Helicobacter pylori infecting dyspeptics from different countries. *FEMS Immunol. Med. Microbiol.* **32**, 17–25 (2001).

428. Vollan, H. S., Tannæs, T., Caugant, D. A., Vriend, G. & Bukholm, G. Outer membrane phospholipase A's roles in Helicobacter pylori acid adaptation. *Gut Pathog.* **9**, (2017).

429. Bos, M. P. *et al.* Function of neisserial outer membrane phospholipase A in autolysis and assessment of its vaccine potential. *Infect. Immun.* **73**, 2222–2231 (2005).

430. Collado, L. & Figueras, M. J. Taxonomy, epidemiology, and clinical relevance of the genus Arcobacter. *Clin. Microbiol. Rev.* **24**, 174–192 (2011).

431. Wang, X. *et al.* The outer membrane phospholipase A is essential for membrane integrity and type III secretion in Shigella flexneri. *Open Biol.* **6**, (2016).

432. Liao, S. *et al.* Biochemical characterization of the outer membrane enzyme ompla from Riemerella anatipestifer. *J. Anim. Vet. Adv.* **11**, 4073–4078 (2012).

433. Belden, W. J. & Miller, S. I. Further characterization of the PhoP regulon: identification of new PhoP-activated virulence loci. *Infect. Immun.* **62**, 5095–5101 (1994).

434. Bishop, R. E. Structural biology of membrane-intrinsic β-barrel enzymes: Sentinels of the bacterial outer membrane. *Biochim. Biophys. Acta BBA - Biomembr.* **1778**, 1881–1896 (2008).

435. Trent, M. S., Pabich, W., Raetz, C. R. H. & Miller, S. I. A PhoP/PhoQ-induced Lipase (PagL) That Catalyzes 3-O-Deacylation of Lipid A Precursors in Membranes ofSalmonella typhimurium *. *J. Biol. Chem.* **276**, 9083–9092 (2001).

436. Ernst, R. K. *et al.* The Pseudomonas aeruginosa lipid A deacylase: Selection for expression and loss within the cystic fibrosis airway. *J. Bacteriol.* **188**, 191–201 (2006).

437. Guo, L. *et al.* Lipid A Acylation and Bacterial Resistance against Vertebrate Antimicrobial Peptides. *Cell* **95**, 189–198 (1998).

438. Murata, T., Tseng, W., Guina, T., Miller, S. I. & Nikaido, H. PhoPQ-mediated regulation produces a more robust permeability barrier in the outer membrane of Salmonella enterica serovar typhimurium. *J. Bacteriol.* **189**, 7213–7222 (2007).

439. Hwang, P. M. *et al.* Solution structure and dynamics of the outer membrane enzyme PagP by NMR. *Proc. Natl. Acad. Sci. U. S. A.* **99**, 13560–13565 (2002).

440. Cox, K. & Sansom, M. S. P. One membrane protein, two structures and six environments: a comparative molecular dynamics simulation study of the bacterial outer membrane protein PagP. *Mol. Membr. Biol.* **26**, 205–214 (2009).

441. Cuesta-Seijo, J. A. *et al.* PagP crystallized from SDS/Cosolvent reveals the route for phospholipid access to the hydrocarbon ruler. *Structure* **18**, 1210–1219 (2010).

442. Anandan, A. & Vrielink, A. Structure and function of lipid A-modifying enzymes. *Ann. N. Y. Acad. Sci.* **1459**, 19–37 (2020).

443. Ahn, V. E. *et al.* A hydrocarbon ruler measures palmitate in the enzymatic acylation of endotoxin. *EMBO J.* **23**, 2931–2941 (2004).

444. Khan, M. A. & Bishop, R. E. Molecular mechanism for lateral lipid diffusion between the outer membrane external leaflet and a beta-barrel hydrocarbon ruler. *Biochemistry* **48**, 9745–9756 (2009).

445. Iyer, B. R. & Mahalakshmi, R. Distinct Structural Elements Govern the Folding, Stability, and Catalysis in the Outer Membrane Enzyme PagP. *Biochemistry* **55**, 4960–4970 (2016).

446. Huysmans, G. H. M., Radford, S. E., Brockwell, D. J. & Baldwin, S. A. The N-terminal Helix Is a Post-assembly Clamp in the Bacterial Outer Membrane Protein PagP. *J. Mol. Biol.* **373**, 529–540 (2007).

447. Tam, C. & Missiakas, D. Changes in lipopolysaccharide structure induce the sigma(E)-dependent response of Escherichia coli. *Mol. Microbiol.* **55**, 1403–1412 (2005).

448. Chalabaev, S. *et al.* Biofilms formed by gram-negative bacteria undergo increased lipid a palmitoylation, Enhancing in vivo survival. *mBio* **5**, (2014).

449. Preston, A. *et al.* Bordetella bronchiseptica PagP is a Bvg-regulated lipid A palmitoyl transferase that is required for persistent colonization of the mouse respiratory tract. *Mol. Microbiol.* **48**, 725–736 (2003).

450. Pilione, M. R., Pishko, E. J., Preston, A., Maskell, D. J. & Harvill, E. T. pagP is required for resistance to antibody-mediated complement lysis during Bordetella bronchiseptica respiratory infection. *Infect. Immun.* **72**, 2837–2842 (2004).

451. Robey, M., O'Connell, W. & Cianciotto, N. P. Identification of Legionella pneumophila rcp, a pagP-like gene that confers resistance to cationic antimicrobial peptides and promotes intracellular infection. *Infect. Immun.* **69**, 4276–4286 (2001).

452. Thaipisuttikul, I. *et al.* A divergent Pseudomonas aeruginosa palmitoyltransferase essential for cystic fibrosis-specific lipid A. *Mol. Microbiol.* **91**, 158–174 (2014).

453. Hittle, L. E., Jones, J. W., Hajjar, A. M., Ernst, R. K. & Preston, A. Bordetella parapertussis PagP mediates the addition of two palmitates to the lipopolysaccharide lipid A. *J. Bacteriol.* **197**, 572–580 (2015).

454. Kawasaki, K., Ernst, R. K. & Miller, S. I. Deacylation and palmitoylation of lipid A by Salmonellae outer membrane enzymes modulate host signaling through Toll-like receptor 4. *J. Endotoxin Res.* **10**, 439–444 (2004).

455. Kawasaki, K., Ernst, R. K. & Miller, S. I. Purification and characterization of deacylated and/or palmitoylated lipid A species unique to Salmonella enterica serovar Typhimurium. *J. Endotoxin Res.* **11**, 57–61 (2005).

456. Kawasaki, K., Ernst, R. K. & Miller, S. I. 3-O-deacylation of lipid A by PagL, a PhoP/PhoQ-regulated deacylase of Salmonella typhimurium, modulates signaling through Toll-like receptor 4. *J. Biol. Chem.* **279**, 20044–20048 (2004).

457. Raetz, C. R. Regulated covalent modifications of lipid A. *J. Endotoxin Res.* **7**, 73–78 (2001).

458. Chandler, C. E. *et al.* Early evolutionary loss of the lipid A modifying enzyme PagP resulting in innate immune evasion in Yersinia pestis. *Proc. Natl. Acad. Sci. U. S. A.* **117**, 22984–22991 (2020).

459. Boll, J. M. *et al.* Reinforcing lipid a acylation on the cell surface of acinetobacter baumannii promotes cationic antimicrobial peptide resistance and desiccation survival. *mBio* **6**, 1–11 (2015).

460. Hews, C. L., Cho, T., Rowley, G. & Raivio, T. L. Maintaining Integrity Under Stress: Envelope Stress Response Regulation of Pathogenesis in Gram-Negative Bacteria. *Front. Cell. Infect. Microbiol.* **9**, 313 (2019).

461. Mitchell, A. M. & Silhavy, T. J. Envelope stress responses: balancing damage repair and toxicity. *Nat. Rev. Microbiol.* **17**, 417–428 (2019).

462. Leblanc, S. K. D., Oates, C. W. & Raivio, T. L. Characterization of the induction and cellular role of the BaeSR two-component envelope stress response of Escherichia coli. *J. Bacteriol.* **193**, 3367–3375 (2011).

463. Joly, N. *et al.* Managing membrane stress: the phage shock protein (Psp) response, from molecular mechanisms to physiology. *FEMS Microbiol. Rev.* **34**, 797–827 (2010).

464. Walsh, N. P., Alba, B. M., Bose, B., Gross, C. A. & Sauer, R. T. OMP peptide signals initiate the envelope-stress response by activating DegS protease via relief of inhibition mediated by its PDZ domain. *Cell* **113**, 61–71 (2003).

465. Lima, S., Guo, M. S., Chaba, R., Gross, C. A. & Sauer, R. T. Dual molecular signals mediate the bacterial response to outer-membrane stress. *Science* **340**, 837–841 (2013).

466. Rhodius, V. A., Suh, W. C., Nonaka, G., West, J. & Gross, C. A. Conserved and Variable Functions of the σE Stress Response in Related Genomes. *PLOS Biol.* **4**, e2 (2005).

467. Dartigalongue, C., Missiakas, D. & Raina, S. Characterization of the Escherichia coli sigma E regulon. *J. Biol. Chem.* **276**, 20866–20875 (2001).

468. Konovalova, A. *et al.* Inhibitor of intramembrane protease RseP blocks the σE response causing lethal accumulation of unfolded outer membrane proteins. *Proc. Natl. Acad. Sci. U. S. A.* **115**, E6614–E6621 (2018).

469. De Las Peñas, A., Connolly, L. & Gross, C. A. The sigmaE-mediated response to extracytoplasmic stress in Escherichia coli is transduced by RseA and RseB, two negative regulators of sigmaE. *Mol. Microbiol.* **24**, 373–385 (1997).

470. McEwen, J. & Silverman, P. Chromosomal mutations of Escherichia coli that alter expression of conjugative plasmid functions. *Proc. Natl. Acad. Sci.* **77**, 513–517 (1980).

471. Raivio, T. L. Everything old is new again: an update on current research on the Cpx envelope stress response. *Biochim. Biophys. Acta* **1843**, 1529–1541 (2014).

472. May, K. L., Lehman, K. M., Mitchell, A. M. & Grabowicz, M. A Stress Response Monitoring Lipoprotein Trafficking to the Outer Membrane. *mBio* **10**, e00618-19 (2019).

473. Snyder, W. B., Davis, L. J., Danese, P. N., Cosma, C. L. & Silhavy, T. J. Overproduction of NlpE, a new outer membrane lipoprotein, suppresses the toxicity of periplasmic LacZ by activation of the Cpx signal transduction pathway. *J. Bacteriol.* **177**, 4216–4223 (1995).

474. Acosta, N., Pukatzki, S. & Raivio, T. L. The Cpx system regulates virulence gene expression in Vibrio cholerae. *Infect. Immun.* **83**, 2396–2408 (2015).

475. Debnath, I. *et al.* The Cpx Stress Response System Potentiates the Fitness and Virulence of Uropathogenic Escherichia coli. *Infect. Immun.* **81**, 1450–1459 (2013).

476. Weatherspoon-Griffin, N. *et al.* The CpxR/CpxA two-component system up-regulates two tat-dependent peptidoglycan amidases to confer bacterial resistance to antimicrobial peptide. *J. Biol. Chem.* **286**, 5529–5539 (2011).

477. Cho, S.-H., Dekoninck, K. & Collet, J.-F. Envelope-Stress Sensing Mechanism of Rcs and Cpx Signaling Pathways in Gram-Negative Bacteria. *J. Microbiol. Seoul Korea* **61**, 317–329 (2023).

478. Dekoninck, K. *et al.* Defining the function of OmpA in the Rcs stress response. *eLife* **9**, e60861 (2020).

479. Konovalova, A., Perlman, D. H., Cowles, C. E. & Silhavy, T. J. Transmembrane domain of surface-exposed outer membrane lipoprotein RcsF is threaded through the lumen of β-barrel proteins. *Proc. Natl. Acad. Sci.* **111**, E4350–E4358 (2014).

480. Rodríguez-Alonso, R. *et al.* Structural insight into the formation of lipoprotein-β-barrel complexes. *Nat. Chem. Biol.* **16**, 1019–1025 (2020).

481. Khalid, S., Piggot, T. J. & Samsudin, F. Atomistic and Coarse Grain Simulations of the Cell Envelope of Gram-Negative Bacteria: What Have We Learned? *Acc. Chem. Res.* **52**, 180–188 (2019).

482. Schindler, M., Osborn, M. J. & Koppel, D. E. Lateral diffusion of lipopolysaccharide in the outer membrane of Salmonella typhimurium. *Nature* **285**, 261–263 (1980).

483. Ranjit, C. & Noll, K. M. Distension of the toga of Thermotoga maritima involves continued growth of the outer envelope as cells enter the stationary phase. *FEMS Microbiol. Lett.* **363**, fnw218 (2016).

484. Rachel, R., Engel, A. M., Huber, R., Stetter, K.-O. & Baumeister, W. A porin-type protein is the main constituent of the cell envelope of the ancestral eubacterium Thermotoga maritima. *FEBS Lett.* **262**, 64–68 (1990).

485. Sahonero-Canavesi, D. X. *et al.* Changes in the Distribution of Membrane Lipids during Growth of Thermotoga maritima at Different Temperatures: Indications for the Potential Mechanism of Biosynthesis of Ether-Bound Diabolic Acid (Membrane-Spanning) Lipids. *Appl. Environ. Microbiol.* **88**, e0176321 (2022).

486. Sexton, D. L. *et al.* The cell envelope of Thermotogae suggests a mechanism for outer membrane biogenesis. 2022.09.22.508938 Preprint at https://doi.org/10.1101/2022.09.22.508938 (2022).

487. Gammoudi, I. *et al.* Morphological and nanostructural surface changes in Escherichia coli over time, monitored by atomic force microscopy. *Colloids Surf. B Biointerfaces* **141**, 355–364 (2016).

488. Schabert, F. A. & Engel, A. Reproducible acquisition of Escherichia coli porin surface topographs by atomic force microscopy. *Biophys. J.* **67**, 2394–2403 (1994).

489. Arunmanee, W. *et al.* Gram-negative trimeric porins have specific LPS binding sites that are essential for porin biogenesis. *Proc. Natl. Acad. Sci.* **113**, E5034–E5043 (2016).

490. Jarosławski, S., Duquesne, K., Sturgis, J. N. & Scheuring, S. High-resolution architecture of the outer membrane of the Gram-negative bacteria Roseobacter denitrificans. *Mol. Microbiol.* **74**, 1211–1222 (2009).

491. Webby, M. N. *et al.* Lipids mediate supramolecular outer membrane protein assembly in bacteria. *Sci. Adv.* **8**, eadc9566 (2022).

492. Bohrer, C. H. & Xiao, J. Complex Diffusion in Bacteria. *Adv. Exp. Med. Biol.* **1267**, 15–43 (2020).

493. Verhoeven, G. S., Dogterom, M. & den Blaauwen, T. Absence of long-range diffusion of OmpA in E. coli is not caused by its peptidoglycan binding domain. *BMC Microbiol.* **13**, 66 (2013).

494. Snyder, S., Kim, D. & McIntosh, T. J. Lipopolysaccharide Bilayer Structure: Effect of Chemotype, Core Mutations, Divalent Cations, and Temperature. *Biochemistry* **38**, 10758–10767 (1999).

495. Fleming, P. J., Freites, J. A., Moon, C. P., Tobias, D. J. & Fleming, K. G. Outer membrane phospholipase A in phospholipid bilayers: A model system for concerted computational and experimental investigations of amino acid side chain partitioning into lipid bilayers. *Biochim. Biophys. Acta - Biomembr.* **1818**, 126–134 (2012).

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

497. Kleanthous, C., Rassam, P. & Baumann, C. G. Protein-protein interactions and the spatiotemporal dynamics of bacterial outer membrane proteins. *Curr. Opin. Struct. Biol.* **35**, 109–115 (2015).

498. Toyofuku, M., Schild, S., Kaparakis-Liaskos, M. & Eberl, L. Composition and functions of bacterial membrane vesicles. *Nat. Rev. Microbiol.* **21**, 415–430 (2023).

499. Schwechheimer, C. & Kuehn, M. J. Outer-membrane vesicles from Gram-negative bacteria: biogenesis and functions. *Nat. Rev. Microbiol.* **13**, 605–619 (2015).

500. Doyle, M. T. & Bernstein, H. D. BamA forms a translocation channel for polypeptide export across the bacterial outer membrane. *Mol. Cell* **81**, 2000-2012.e3 (2021).

501. Mamou, G. *et al.* Peptidoglycan maturation controls outer membrane protein assembly. *Nature* **606**, 953–959 (2022).

502. Bryant, J. A. *et al.* Bam complex associated proteins in Escherichia coli are functionally linked to peptidoglycan biosynthesis, membrane fluidity and DNA replication. 2023.07.05.547807 Preprint at https://doi.org/10.1101/2023.07.05.547807 (2023).

503. Brandon, L. D. *et al.* IcsA, a polarly localized autotransporter with an atypical signal peptide, uses the Sec apparatus for secretion, although the Sec apparatus is circumferentially distributed. *Mol. Microbiol.* **50**, 45–60 (2003).

504. Nyenhuis, D. A., Nilaweera, T. D. & Cafiso, D. S. Native Cell Environment Constrains Loop Structure in the Escherichia coli Cobalamin Transporter BtuB. *Biophys. J.* **119**, 1550–1557 (2020).

505. Balusek, C. & Gumbart, J. C. Role of the Native Outer-Membrane Environment on the Transporter BtuB. *Biophys. J.* **111**, 1409–1417 (2016).

506. Nyenhuis, D. A. *et al.* Evidence for the Supramolecular Organization of a Bacterial Outer-Membrane Protein from In Vivo Pulse Electron Paramagnetic Resonance Spectroscopy. *J. Am. Chem. Soc.* **142**, 10715–10722 (2020).

507. Berg, B. van den, Silale, A., Baslé, A., Mader, S. L. & Khalid, S. Structural basis for host recognition and superinfection exclusion by bacteriophage T5. 2022.06.28.497910 Preprint at https://doi.org/10.1101/2022.06.28.497910 (2022).

508. Kucharska, I., Liang, B., Ursini, N. & Tamm, L. K. Molecular Interactions of Lipopolysaccharide with an Outer Membrane Protein from Pseudomonas aeruginosa Probed by Solution NMR. *Biochemistry* **55**, 5061–5072 (2016).

509. Shearer, J., Jefferies, D. & Khalid, S. Outer Membrane Proteins OmpA, FhuA, OmpF, EstA, BtuB, and OmpX Have Unique Lipopolysaccharide Fingerprints. *J. Chem. Theory Comput.* **15**, 2608–2619 (2019).

510. *CDC AR Threats Report 2019*. https://www.cdc.gov/drugresistance/biggest-threats.html.

511. Rice, L. B. Federal funding for the study of antimicrobial resistance in nosocomial pathogens: no ESKAPE. *J. Infect. Dis.* **197**, 1079–1081 (2008).

512. Murray, C. J. L. *et al.* Global burden of bacterial antimicrobial resistance in 2019: a systematic analysis. *The Lancet* **399**, 629–655 (2022).

513. Nikaido, H. How do exported proteins and antibiotics bypass the periplasm in Gram-negative bacterial cells? *Trends Microbiol.* **8**, 481–483 (2000).

514. Darby, E. M. *et al.* Molecular mechanisms of antibiotic resistance revisited. *Nat. Rev. Microbiol.* **21**, 280–295 (2023).

515. Kapoor, G., Saigal, S. & Elongavan, A. Action and resistance mechanisms of antibiotics: A guide for clinicians. *J. Anaesthesiol. Clin. Pharmacol.* **33**, 300 (2017).

516. Mohapatra, S. S., Dwibedy, S. K. & Padhy, I. Polymyxins, the last-resort antibiotics: Mode of action, resistance emergence, and potential solutions. *J. Biosci.* **46**, 85 (2021).

517. Imai, Y. *et al.* A new antibiotic selectively kills Gram-negative pathogens. *Nature* **576**, 459–464 (2019).

518. Hart, E. M. *et al.* A small-molecule inhibitor of BamA impervious to efflux and the outer membrane permeability barrier. *Proc. Natl. Acad. Sci. U. S. A.* **116**, 21748–21757 (2019).

519. Storek, K. M. *et al.* Monoclonal antibody targeting the β-barrel assembly machine of Escherichia coli is bactericidal. *Proc. Natl. Acad. Sci. U. S. A.* **115**, 3692–3697 (2018).

520. Miller, R. D. *et al.* Computational identification of a systemic antibiotic for gram-negative bacteria. *Nat. Microbiol.* **7**, 1661–1672 (2022).

521. Haysom, S. F. *et al.* Darobactin B Stabilises a Lateral-Closed Conformation of the BAM Complex in E. coli Cells. *Angew. Chem.* **135**, e202218783 (2023).

522. Seyfert, C. E. *et al.* Darobactins Exhibiting Superior Antibiotic Activity by Cryo-EM Structure Guided Biosynthetic Engineering. *Angew. Chem. Int. Ed Engl.* **62**, e202214094 (2023).

523. Zmora, N., Suez, J. & Elinav, E. You are what you eat: diet, health and the gut microbiota. *Nat. Rev. Gastroenterol. Hepatol.* **16**, 35–56 (2019).

524. Pickard, J. M., Zeng, M. Y., Caruso, R. & Núñez, G. Gut microbiota: Role in pathogen colonization, immune responses, and inflammatory disease. *Immunol. Rev.* **279**, 70–89 (2017).

525. Eckburg, P. B. *et al.* Diversity of the human intestinal microbial flora. *Science* **308**, 1635–1638 (2005).

526. Radjabzadeh, D. *et al.* Diversity, compositional and functional differences between gut microbiota of children and adults. *Sci. Rep.* **10**, 1040 (2020).

527. Wardman, J. F., Bains, R. K., Rahfeld, P. & Withers, S. G. Carbohydrate-active enzymes (CAZymes) in the gut microbiome. *Nat. Rev. Microbiol.* **20**, 542–556 (2022).

528. LeBlanc, J. G. *et al.* Bacteria as vitamin suppliers to their host: a gut microbiota perspective. *Curr. Opin. Biotechnol.* **24**, 160–168 (2013).

529. Jin, L. Z., Marquardt, R. R. & Zhao, X. A strain of Enterococcus faecium (18C23) inhibits adhesion of enterotoxigenic Escherichia coli K88 to porcine small intestine mucus. *Appl. Environ. Microbiol.* **66**, 4200–4204 (2000).

530. Yadav, H. *et al.* Unveiling the role of gut-brain axis in regulating neurodegenerative diseases: A comprehensive review. *Life Sci.* **330**, 122022 (2023).

531. Ellis, S. R. *et al.* The Skin and Gut Microbiome and Its Role in Common Dermatologic Conditions. *Microorganisms* **7**, 550 (2019).

532. Almeida, A. *et al.* A unified catalog of 204,938 reference genomes from the human gut microbiome. *Nat. Biotechnol.* **39**, 105–114 (2021).

533. Koppel, N., Maini Rekdal, V. & Balskus, E. P. Chemical transformation of xenobiotics by the human gut microbiota. *Science* **356**, eaag2770 (2017).

534. Zimmermann, M., Zimmermann-Kogadeeva, M., Wegmann, R. & Goodman, A. L. Mapping human microbiome drug metabolism by gut bacteria and their genes. *Nature* **570**, 462–467 (2019).

535. M, P. & Mc, A. The global human gut microbiome: genes, lifestyles, and diet. *Trends Mol. Med.* (2023) doi:10.1016/j.molmed.2023.07.002.

536. Denisov, I. G., Grinkova, Y. V., Lazarides, A. A. & Sligar, S. G. Directed self-assembly of monodisperse phospholipid bilayer Nanodiscs with controlled size. *J. Am. Chem. Soc.* **126**, 3477–3487 (2004).

537. Grinter, R. *et al.* Structure of the bacterial plant-ferredoxin receptor FusA. *Nat. Commun.* **7**, 13308 (2016).

538. Zivanov, J., Nakane, T. & Scheres, S. H. W. Estimation of high-order aberrations and anisotropic magnification from cryo-EM data sets in RELION-3.1. *IUCrJ* **7**, 253–267 (2020).

539. Kimanius, D., Dong, L., Sharov, G., Nakane, T. & Scheres, S. H. W. New tools for automated cryo-EM single-particle analysis in RELION-4.0. *Biochem. J.* **478**, 4169–4185 (2021).

540. Rohou, A. & Grigorieff, N. CTFFIND4: Fast and accurate defocus estimation from electron micrographs. *J. Struct. Biol.* **192**, 216–221 (2015).

541. Wagner, T. *et al.* SPHIRE-crYOLO is a fast and accurate fully automated particle picker for cryo-EM. *Commun. Biol.* **2**, 218 (2019).

542. Bepler, T. *et al.* Positive-unlabeled convolutional neural networks for particle picking in cryo-electron micrographs. *Nat. Methods* **16**, 1153–1160 (2019).

543. Punjani, A., Zhang, H. & Fleet, D. J. Non-uniform refinement: adaptive regularization improves single-particle cryo-EM reconstruction. *Nat. Methods* **17**, 1214–1221 (2020).

544. Croll, T. I. ISOLDE: a physically realistic environment for model building into low-resolution electron-density maps. *Acta Crystallogr. Sect. Struct. Biol.* **74**, 519–530 (2018).

545. Afonine, P. V. *et al.* Real-space refinement in PHENIX for cryo-EM and crystallography. *Acta Crystallogr. Sect. Struct. Biol.* **74**, 531–544 (2018).

546. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 486–501 (2010).

547. MolProbity: all-atom structure validation for macromolecular crystallography - PubMed. https://pubmed.ncbi.nlm.nih.gov/20057044/.

548. Goddard, T. D. *et al.* UCSF ChimeraX: Meeting modern challenges in visualization and analysis. *Protein Sci. Publ. Protein Soc.* **27**, 14–25 (2018).

549. Van Der Spoel, D. *et al.* GROMACS: fast, flexible, and free. *J. Comput. Chem.* **26**, 1701–1718 (2005).

550. Marrink, S. J., Risselada, H. J., Yefimov, S., Tieleman, D. P. & de Vries, A. H. The MARTINI force field: coarse grained model for biomolecular simulations. *J. Phys. Chem. B* **111**, 7812–7824 (2007).

551. de Jong, D. H. *et al.* Improved Parameters for the Martini Coarse-Grained Protein Force Field. *J. Chem. Theory Comput.* **9**, 687–697 (2013).

552. Wassenaar, T. A., Ingólfsson, H. I., Böckmann, R. A., Tieleman, D. P. & Marrink, S. J. Computational Lipidomics with insane: A Versatile Tool for Generating Custom Membranes for Molecular Simulations. *J. Chem. Theory Comput.* **11**, 2144–2155 (2015).

553. Parrinello, M. & Rahman, A. Polymorphic transitions in single crystals: A new molecular dynamics method. *J. Appl. Phys.* **52**, 7182–7190 (1981).

554. Bussi, G., Donadio, D. & Parrinello, M. Canonical sampling through velocity rescaling. *J. Chem. Phys.* **126**, 014101 (2007).

555. Hess, B., Bekker, H., Berendsen, H. J. C. & Fraaije, J. G. E. M. LINCS: A linear constraint solver for molecular simulations. *J. Comput. Chem.* **18**, 1463–1472 (1997).

556. Song, W. *et al.* PyLipID: A Python Package for Analysis of Protein-Lipid Interactions from Molecular Dynamics Simulations. *J. Chem. Theory Comput.* **18**, 1188–1201 (2022).

557. Humphrey, W., Dalke, A. & Schulten, K. VMD: visual molecular dynamics. *J. Mol. Graph.* **14**, 33–38, 27–28 (1996).

558. Rothman, J. E. & Lenard, J. Membrane asymmetry. *Science* **195**, 743–753 (1977).

559. van Meer, G. Cellular lipidomics. *EMBO J.* **24**, 3159–3165 (2005).

560. Ikeda, M., Kihara, A. & Igarashi, Y. Lipid asymmetry of the eukaryotic plasma membrane: functions and related enzymes. *Biol. Pharm. Bull.* **29**, 1542–1546 (2006).

561. Malinverni, J. C. & Silhavy, T. J. An ABC transport system that maintains lipid asymmetry in the gram-negative outer membrane. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 8009–8014 (2009).

562. Cheng, H.-T., Megha, null & London, E. Preparation and properties of asymmetric vesicles that mimic cell membranes: effect upon lipid raft formation and transmembrane helix orientation. *J. Biol. Chem.* **284**, 6079–6092 (2009).

563. St Clair, J. W., Kakuda, S. & London, E. Induction of Ordered Lipid Raft Domain Formation by Loss of Lipid Asymmetry. *Biophys. J.* **119**, 483–492 (2020).

564. Verherstraeten, S. *et al.* Perfringolysin O: The Underrated Clostridium perfringens Toxin? *Toxins* **7**, 1702–1721 (2015).

565. Reshetnyak, Y. K., Andreev, O. A., Segala, M., Markin, V. S. & Engelman, D. M. Energetics of peptide (pHLIP) binding to and folding across a lipid bilayer membrane. *Proc. Natl. Acad. Sci. U. S. A.* **105**, 15340–15345 (2008).

566. Peterson, J. H., Plummer, A. M., Fleming, K. G. & Bernstein, H. D. Selective pressure for rapid membrane integration constrains the sequence of bacterial outer membrane proteins. *Mol. Microbiol.* **106**, 777–792 (2017).

567. Caffrey, M. & Hogan, J. LIPIDAT: a database of lipid phase transition temperatures and enthalpy changes. DMPC data subset analysis. *Chem. Phys. Lipids* **61**, 1–109 (1992).

568. Chen, T. & Guestrin, C. XGBoost: A Scalable Tree Boosting System. in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* 785–794 (Association for Computing Machinery, 2016). doi:10.1145/2939672.2939785.

569. Smith, M. C., Crist, R. M., Clogston, J. D. & McNeil, S. E. Zeta potential: a case study of cationic, anionic, and neutral liposomes. *Anal. Bioanal. Chem.* **409**, 5779–5787 (2017).

570. Soema, P. C., Willems, G.-J., Jiskoot, W., Amorij, J.-P. & Kersten, G. F. Predicting the influence of liposomal lipid composition on liposome size, zeta potential and liposome-induced dendritic cell maturation using a design of experiments approach. *Eur. J. Pharm. Biopharm. Off. J. Arbeitsgemeinschaft Pharm. Verfahrenstechnik EV* **94**, 427–435 (2015).

571. Sęk, A., Perczyk, P., Wydro, P., Gruszecki, W. I. & Szcześ, A. Effect of trace amounts of ionic surfactants on the zeta potential of DPPC liposomes. *Chem. Phys. Lipids* **235**, 105059 (2021).

572. Svirina, A. & Terterov, I. Electrostatic effects in saturation of membrane binding of cationic cell-penetrating peptide. *Eur. Biophys. J. EBJ* **50**, 15–23 (2021).

573. Makino, K. *et al.* Temperature- and ionic strength-induced conformational changes in the lipid head group region of liposomes as suggested by zeta potential data. *Biophys. Chem.* **41**, 175–183 (1991).

574. Silvander, M., Hansson, P. & Edwards, K. Liposomal Surface Potential and Bilayer Packing As Affected by PEG−Lipid Inclusion. *Langmuir* **16**, 3696–3702 (2000).

575. Tatulian, S. A. Effect of lipid phase transition on the binding of anions to dimyristoylphosphatidylcholine liposomes. *Biochim. Biophys. Acta* **736**, 189–195 (1983).

576. Lairion, F. & Disalvo, E. A. Effect of dipole potential variations on the surface charge potential of lipid membranes. *J. Phys. Chem. B* **113**, 1607–1614 (2009).

577. Le, Q.-C., Ropers, M.-H., Terrisse, H. & Humbert, B. Interactions between phospholipids and titanium dioxide particles. *Colloids Surf. B Biointerfaces* **123**, 150–157 (2014).

578. Luzardo, M. del C., Peltzer, G. & Disalvo, E. A. Surface Potential of Lipid Interfaces Formed by Mixtures of Phosphatidylcholine of Different Chain Lengths. *Langmuir* **14**, 5858–5862 (1998).

579. Fatouros, D. G. & Antimisiaris, S. G. Effect of amphiphilic drugs on the stability and zeta-potential of their liposome formulations: a study with prednisolone, diazepam, and griseofulvin. *J. Colloid Interface Sci.* **251**, 271–277 (2002).

580. Morini, M. A. *et al.* Influence of temperature, anions and size distribution on the zeta potential of DMPC, DPPC and DMPE lipid vesicles. *Colloids Surf. B Biointerfaces* **131**, 54–58 (2015).

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

582. Danoff, E. J. & Fleming, K. G. The soluble, periplasmic domain of OmpA folds as an independent unit and displays chaperone activity by reducing the self-association propensity of the unfolded OmpA transmembrane β-barrel. *Biophys. Chem.* **159**, 194–204 (2011).

583. Ridolfi, A. *et al.* Stiffness of Fluid and Gel Phase Lipid Nanovesicles: Weighting the Contributions of Membrane Bending Modulus and Luminal Pressurization. *Langmuir ACS J. Surf. Colloids* **37**, 12027–12037 (2021).

584. Takechi-Haraya, Y., Goda, Y. & Sakai-Kato, K. Atomic Force Microscopy Study on the Stiffness of Nanosized Liposomes Containing Charged Lipids. *Langmuir ACS J. Surf. Colloids* **34**, 7805–7812 (2018).

585. Pocanschi, C. L., Patel, G. J., Marsh, D. & Kleinschmidt, J. H. Curvature elasticity and refolding of OmpA in large unilamellar vesicles. *Biophys. J.* **91**, L75-77 (2006).

586. Takechi-Haraya, Y. *et al.* Atomic Force Microscopic Analysis of the Effect of Lipid Composition on Liposome Membrane Rigidity. *Langmuir ACS J. Surf. Colloids* **32**, 6074–6082 (2016).

587. Et-Thakafy, O. *et al.* Mechanical Properties of Membranes Composed of Gel-Phase or Fluid-Phase Phospholipids Probed on Liposomes by Atomic Force Spectroscopy. *Langmuir ACS J. Surf. Colloids* **33**, 5117–5126 (2017).

588. Rodionova, N. A., Tatulian, S. A., Surrey, T., Jähnig, F. & Tamm, L. K. Characterization of two membrane-bound forms of OmpA. *Biochemistry* **34**, 1921–1929 (1995).

589. Jackups, R. & Liang, J. Interstrand pairing patterns in beta-barrel membrane proteins: the positive-outside rule, aromatic rescue, and strand registration prediction. *J. Mol. Biol.* **354**, 979–993 (2005).

590. Lomize, M. A., Lomize, A. L., Pogozheva, I. D. & Mosberg, H. I. OPM: orientations of proteins in membranes database. *Bioinforma. Oxf. Engl.* **22**, 623–625 (2006).

591. Jumper, J. *et al.* Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589 (2021).

592. Varadi, M. *et al.* AlphaFold Protein Structure Database: massively expanding the structural coverage of protein-sequence space with high-accuracy models. *Nucleic Acids Res.* **50**, D439–D444 (2022).

593. Tsirigos, K. D., Bagos, P. G. & Hamodrakas, S. J. OMPdb: a database of {beta}-barrel outer membrane proteins from Gram-negative bacteria. *Nucleic Acids Res.* **39**, D324-331 (2011).

594. Brüning, B., Stehle, R., Falus, P. & Farago, B. Influence of charge density on bilayer bending rigidity in lipid vesicles: a combined dynamic light scattering and neutron spin-echo study. *Eur. Phys. J. E Soft Matter* **36**, 77 (2013).

595. Liu, P., Zabala-Ferrera, O. & Beltramo, P. J. Fabrication and electromechanical characterization of freestanding asymmetric membranes. *Biophys. J.* **120**, 1755–1764 (2021).

596. Maktabi, S., Schertzer, J. W. & Chiarot, P. R. Dewetting-induced formation and mechanical properties of synthetic bacterial outer membrane models (GUVs) with controlled inner-leaflet lipid composition. *Soft Matter* **15**, 3938–3948 (2019).

597. Lundstedt, E., Kahne, D. & Ruiz, N. Assembly and Maintenance of Lipids at the Bacterial Outer Membrane. *Chem. Rev.* **121**, 5098–5123 (2021).

598. Zhou, H.-X. & Pang, X. Electrostatic Interactions in Protein Structure, Folding, Binding, and Condensation. *Chem. Rev.* **118**, 1691–1741 (2018).

599. Zaydman, M. A. *et al.* Kv7.1 ion channels require a lipid to couple voltage sensing to pore opening. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 13180–13185 (2013).

600. Liu, B. *et al.* Structure and genetics of Escherichia coli O antigens. *FEMS Microbiol. Rev.* **44**, 655–683 (2020).

601. Abellon-Ruiz, J. Forward or backward, that is the question: phospholipid trafficking by the Mla system. *Emerg. Top. Life Sci.* **7**, 125–135 (2023).

602. Maldonado, R. F., Sá-Correia, I. & Valvano, M. A. Lipopolysaccharide modification in Gram-negative bacteria during chronic infection. *FEMS Microbiol. Rev.* **40**, 480–493 (2016).

603. Simpson, B. W. & Trent, M. S. Pushing the envelope: LPS modifications and their consequences. *Nat. Rev. Microbiol.* **17**, 403–416 (2019).

604. Elhenawy, W. *et al.* LPS Remodeling Triggers Formation of Outer Membrane Vesicles in Salmonella. *mBio* **7**, e00940-16 (2016).

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

606. Patel, D. S. *et al.* Dynamics and Interactions of OmpF and LPS: Influence on Pore Accessibility and Ion Permeability. *Biophys. J.* **110**, 930–938 (2016).

607. Ellis, T. N. & Kuehn, M. J. Virulence and immunomodulatory roles of bacterial outer membrane vesicles. *Microbiol. Mol. Biol. Rev. MMBR* **74**, 81–94 (2010).

608. Xu, C., Soyfoo, D. M., Wu, Y. & Xu, S. Virulence of Helicobacter pylori outer membrane proteins: an updated review. *Eur. J. Clin. Microbiol. Infect. Dis. Off. Publ. Eur. Soc. Clin. Microbiol.* **39**, 1821–1830 (2020).

609. Rodrigues, I. C., Rodrigues, S. C., Duarte, F. V., Costa, P. M. da & Costa, P. M. da. The Role of Outer Membrane Proteins in UPEC Antimicrobial Resistance: A Systematic Review. *Membranes* **12**, 981 (2022).

610. Hritonenko, V. & Stathopoulos, C. Omptin proteins: an expanding family of outer membrane proteases in Gram-negative Enterobacteriaceae (Review). *Mol. Membr. Biol.* **24**, 395–406 (2007).

611. Sinsinbar, G. *et al.* Role of Lipopolysaccharide in Protecting OmpT from Autoproteolysis during In Vitro Refolding. *Biomolecules* **10**, 922 (2020).

612. Kramer, R. A. *et al.* Lipopolysaccharide regions involved in the activation of Escherichia coli outer membrane protease OmpT. *Eur. J. Biochem.* **269**, 1746–1752 (2002).

613. Eren, E. & Berg, B. van den. Structural Basis for Activation of an Integral Membrane Protease by Lipopolysaccharide. *J. Biol. Chem.* **287**, 23971–23976 (2012).

614. Eren, E., Murphy, M., Goguen, J. & Berg, B. van den. An Active Site Water Network in the Plasminogen Activator Pla from Yersinia pestis. *Structure* **18**, 809–818 (2010).

615. Kingma, R. L. & Egmond, M. R. Activation of a covalent outer membrane phospholipase A dimer. *Eur. J. Biochem.* **269**, 2178–2185 (2002).

616. Istivan, T. S. & Coloe, P. J. Phospholipase A in Gram-negative bacteria and its role in pathogenesis. *Microbiology* **152**, 1263–1274 (2006).

617. Nygaard, R., Kim, J. & Mancia, F. Cryo-electron microscopy analysis of small membrane proteins. *Curr. Opin. Struct. Biol.* **64**, 26–33 (2020).

618. Madej, M. *et al.* Structural and functional insights into oligopeptide acquisition by the RagAB transporter from Porphyromonas gingivalis. *Nat. Microbiol.* **5**, 1016–1025 (2020).

619. Abellon-Ruiz, J. *et al.* BtuB TonB-dependent transporters and BtuG surface lipoproteins form stable complexes for vitamin B12 uptake in gut Bacteroides. 2022.11.17.516869 Preprint at https://doi.org/10.1101/2022.11.17.516869 (2022).

620. Oulghazi, S. *et al.* Pectobacterium brasiliense: Genomics, Host Range and Disease Management. *Microorganisms* **9**, 106 (2021).

621. Grinter, R., Milner, J. & Walker, D. Ferredoxin containing bacteriocins suggest a novel mechanism of iron uptake in Pectobacterium spp. *PloS One* **7**, e33033 (2012).

622. Shahmiri, M. *et al.* Membrane Core-Specific Antimicrobial Action of Cathelicidin LL-37 Peptide Switches Between Pore and Nanofibre Formation. *Sci. Rep.* **6**, 38184 (2016).

623. Haiko, J., Suomalainen, M., Ojala, T., Lähteenmäki, K. & Korhonen, T. K. Invited review: Breaking barriers — attack on innate immune defences by omptin surface proteases of enterobacterial pathogens. *Innate Immun.* **15**, 67–80 (2009).

624. Biou, V. Lipid-membrane protein interaction visualised by cryo-EM: A review. *Biochim. Biophys. Acta Biomembr.* **1865**, 184068 (2023).

625. Vorobieva, A. A. *et al.* De novo design of transmembrane β barrels. *Science* **371**, eabc8182 (2021).

626. Zeno, W. F., Day, K. J., Gordon, V. D. & Stachowiak, J. C. Principles and Applications of Biological Membrane Organization. *Annu. Rev. Biophys.* **49**, 19–39 (2020).

627. Yamashita, H. *et al.* Single-Molecule Imaging on Living Bacterial Cell Surface by High-Speed AFM. *J. Mol. Biol.* **422**, 300–309 (2012).

628. Casuso, I. *et al.* Characterization of the motion of membrane proteins using high-speed atomic force microscopy. *Nat. Nanotechnol.* **7**, 525–529 (2012).

629. Feodorova, V. A. *et al.* New Promising Targets for Synthetic Omptin-Based Peptide Vaccine against Gram-Negative Pathogens. *Vaccines* **7**, 36 (2019).

630. Ebbensgaard, A., Mordhorst, H., Aarestrup, F. M. & Hansen, E. B. The Role of Outer Membrane Proteins and Lipopolysaccharides for the Sensitivity of Escherichia coli to Antimicrobial Peptides. *Front. Microbiol.* **9**, (2018).

631. Kingma, R. L., Snijder, H. J., Dijkstra, B. W., Dekker, N. & Egmond, M. R. Functional importance of calcium binding sites in outer membrane phospholipase A. *Biochim. Biophys. Acta BBA - Biomembr.* **1561**, 230–237 (2002).

632. Bishop, R. E. The lipid A palmitoyltransferase PagP: molecular mechanisms and role in bacterial pathogenesis. *Mol. Microbiol.* **57**, 900–912 (2005).

633. Baarda, B. I., Zielke, R. A., Holm, A. K. & Sikora, A. E. Comprehensive Bioinformatic Assessments of the Variability of Neisseria gonorrhoeae Vaccine Candidates. *mSphere* **6**, e00977-20 (2021).

634. Wu, E. L. *et al.* E. coli outer membrane and interactions with OmpLA. *Biophys. J.* **106**, 2493–2502 (2014).

635. Ercan, B., Low, W.-Y., Liu, X. & Chng, S.-S. Characterization of Interactions and Phospholipid Transfer between Substrate Binding Proteins of the OmpC-Mla System. *Biochemistry* **58**, 114–119 (2019).

636. Vogt, J. & Schulz, G. E. The structure of the outer membrane protein OmpX from Escherichia coli reveals possible mechanisms of virulence. *Struct. Lond. Engl. 1993* **7**, 1301–1309 (1999).

637. Snijder, H. J. *et al.* Structural evidence for dimerization-regulated activation of an integral membrane phospholipase. *Nature* **401**, 717–721 (1999).

638. Jeanteur, D. *et al.* Structural and functional alterations of a colicin-resistant mutant of OmpF porin from Escherichia coli. *Proc. Natl. Acad. Sci. U. S. A.* **91**, 10675–10679 (1994).

639. Chimento, D. P., Mohanty, A. K., Kadner, R. J. & Wiener, M. C. Substrate-induced transmembrane signaling in the cobalamin transporter BtuB. *Nat. Struct. Biol.* **10**, 394–401 (2003).

640. Zidovetzki, R. & Levitan, I. Use of cyclodextrins to manipulate plasma membrane cholesterol content: Evidence, misconceptions and control strategies. *Biochim. Biophys. Acta BBA - Biomembr.* **1768**, 1311–1324 (2007).

641. Huber, A. H., Kampf, J. P., Kwan, T., Zhu, B. & Kleinfeld, A. M. Fatty acid-specific fluorescent probes and their use in resolving mixtures of unbound free fatty acids in equilibrium with albumin. *Biochemistry* **45**, 14263–14274 (2006).

642. Curry, S., Mandelkow, H., Brick, P. & Franks, N. Crystal structure of human serum albumin complexed with fatty acid reveals an asymmetric distribution of binding sites. *Nat. Struct. Biol.* **5**, 827–835 (1998).

643. Sekula, B., Zielinski, K. & Bujacz, A. Crystallographic studies of the complexes of bovine and equine serum albumin with 3,5-diiodosalicylic acid. *Int. J. Biol. Macromol.* **60**, 316–324 (2013).

644. Togashi, D. M. & Ryder, A. G. A fluorescence analysis of ANS bound to bovine serum albumin: binding properties revisited by using energy transfer. *J. Fluoresc.* **18**, 519–526 (2008).

645. Evans, R. *et al.* Protein complex prediction with AlphaFold-Multimer. 2021.10.04.463034 Preprint at https://doi.org/10.1101/2021.10.04.463034 (2021).

646. Elfmann, C. & Stülke, J. PAE viewer: a webserver for the interactive visualization of the predicted aligned error for multimer structure predictions and crosslinks. *Nucleic Acids Res.* **51**, W404–W410 (2023).

647. Topitsch, A., Schwede, T. & Pereira, J. Outer membrane β-barrel structure prediction through the lens of AlphaFold2. *Proteins* (2023) doi:10.1002/prot.26552.

648. Alhindi, T. *et al.* Protein interaction evolution from promiscuity to specificity with reduced flexibility in an increasingly complex network. *Sci. Rep.* **7**, 44948 (2017).

649. Andreani, J. & Guerois, R. Evolution of protein interactions: From interactomes to interfaces. *Arch. Biochem. Biophys.* **554**, 65–75 (2014).

650. Lai, E.-M., Nair, U., Phadke, N. D. & Maddock, J. R. Proteomic screening and identification of differentially distributed membrane proteins in Escherichia coli. *Mol. Microbiol.* **52**, 1029–1044 (2004).

651. Stenberg, F. *et al.* Protein Complexes of the Escherichia coli Cell Envelope *. *J. Biol. Chem.* **280**, 34409–34419 (2005).

652. Ormsby, M. J. & Davies, R. L. Diversification of OmpA and OmpF of Yersinia ruckeri is independent of the underlying species phylogeny and evidence of virulence-related selection. *Sci. Rep.* **11**, 3493 (2021).

653. Montezano, D., Bernstein, R., Copeland, M. M. & Slusky, J. S. G. General features of transmembrane beta barrels from a large database. *Proc. Natl. Acad. Sci.* **120**, e2220762120 (2023).

654. Giordano, N. P., Cian, M. B. & Dalebroux, Z. D. Outer Membrane Lipid Secretion and the Innate Immune Response to Gram-Negative Bacteria. *Infect. Immun.* **88**, e00920-19 (2020).

655. Bonnington, K. E. & Kuehn, M. J. Outer Membrane Vesicle Production Facilitates LPS Remodeling and Outer Membrane Maintenance in Salmonella during Environmental Transitions. *mBio* **7**, e01532-16 (2016).

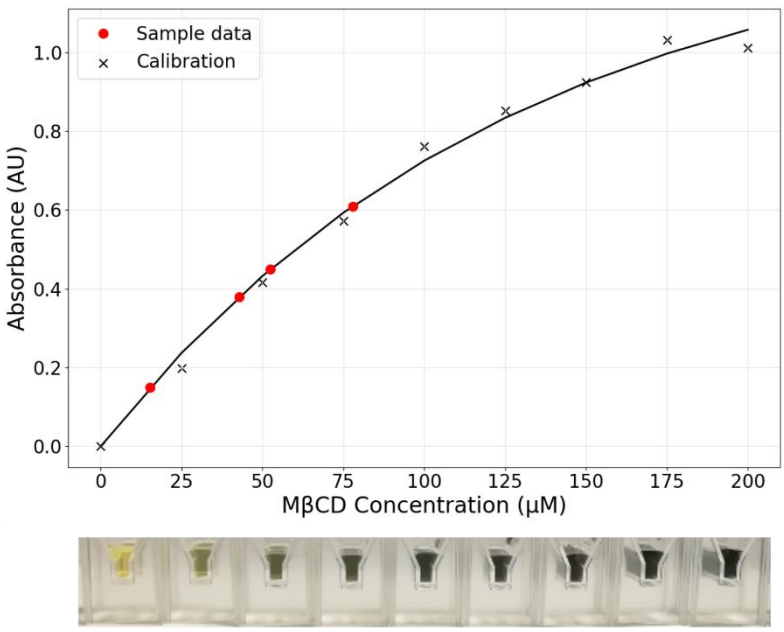
656. Ying, Y.-L. *et al.* Nanopore-based technologies beyond DNA sequencing. *Nat. Nanotechnol.* **17**, 1136–1146 (2022).

657. Lorent, J. H. *et al.* Plasma membranes are asymmetric in lipid unsaturation, packing and protein shape. *Nat. Chem. Biol.* **16**, 644–652 (2020).

658. Hong, H., Choi, H.-K. & Yoon, T.-Y. Untangling the complexity of membrane protein folding. *Curr. Opin. Struct. Biol.* **72**, 237–247 (2022).

Appendix A: Supplementary Figures

a



b

Sample	Initial concentration (mM)	Final Concentration (μM)
1	67	21
2	67	44
3	67	52
4	67	77

Figure A1: MβCD is effectively removed by ultracentrifugation. (a) The sugar anthrone assay can colorimetrically detect MβCD at μM concentrations (black crosses: calibration, red circles: sample data). Calibrant samples are shown below. **(b)** Example measurements of MβCD concentrations before and after ultracentrifugation (red circles in panel (b)). Typical reductions are >1000x, reducing free MβCD, and hence its associated lipid, to insignificant quantities for the assays described. Four replicates with an identical initial concentration of 67,000 μM are shown.

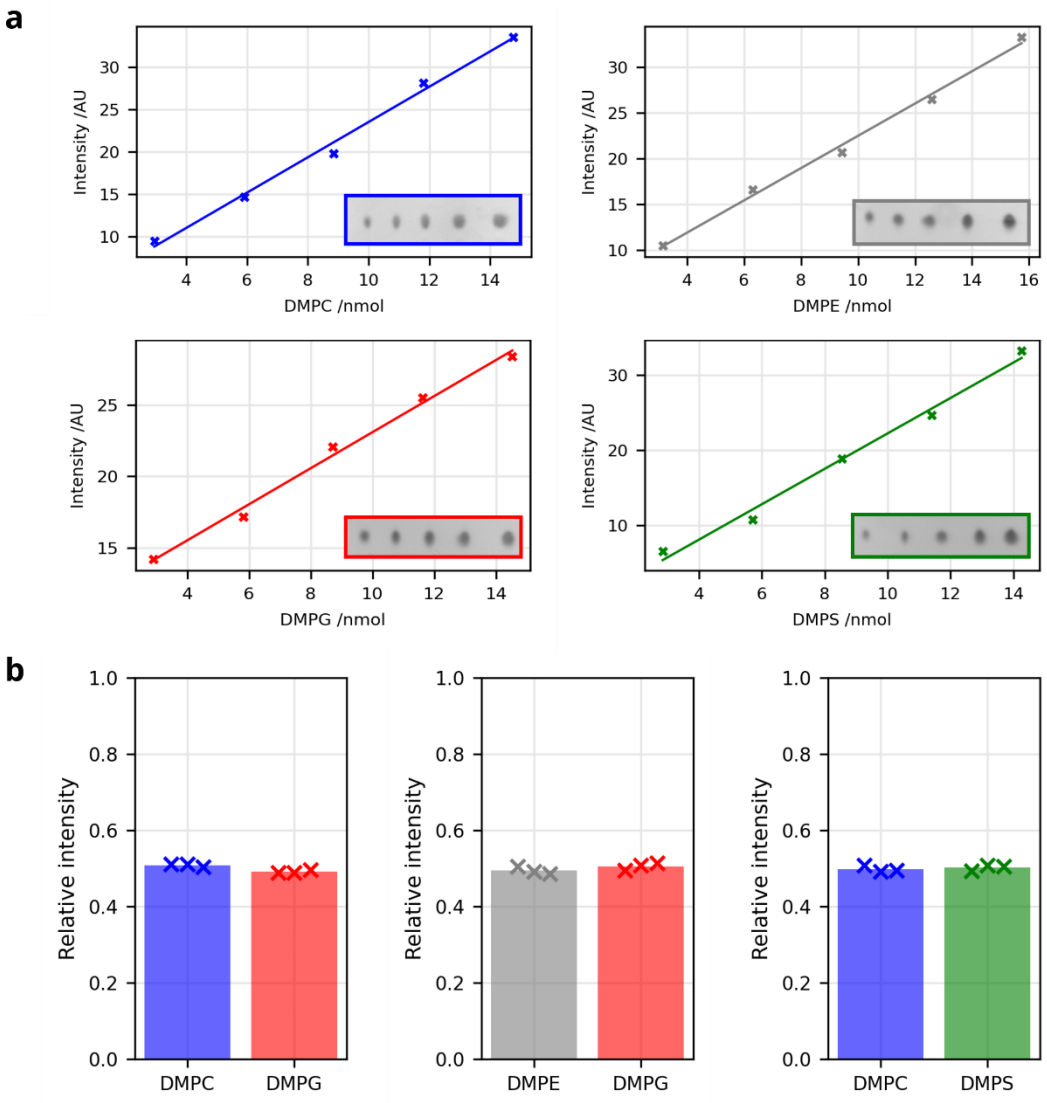


Figure A2: Thin layer chromatography (TLC) can be used to determine relative lipid ratios. (a) DMPC (blue), DMPG (red), DMPE (grey) and DMPS (green) lipid staining depth is directly and linearly proportional to amount of lipid loaded. One example from three replicates shown. An example TLC plate is shown inset in each. **(b)** DMPC-DMPG, DMPE-DMPG and DMPC-DMPS lipids stain to equivalent depths for the same molar lipid amounts (10 nmol of each lipid loaded).

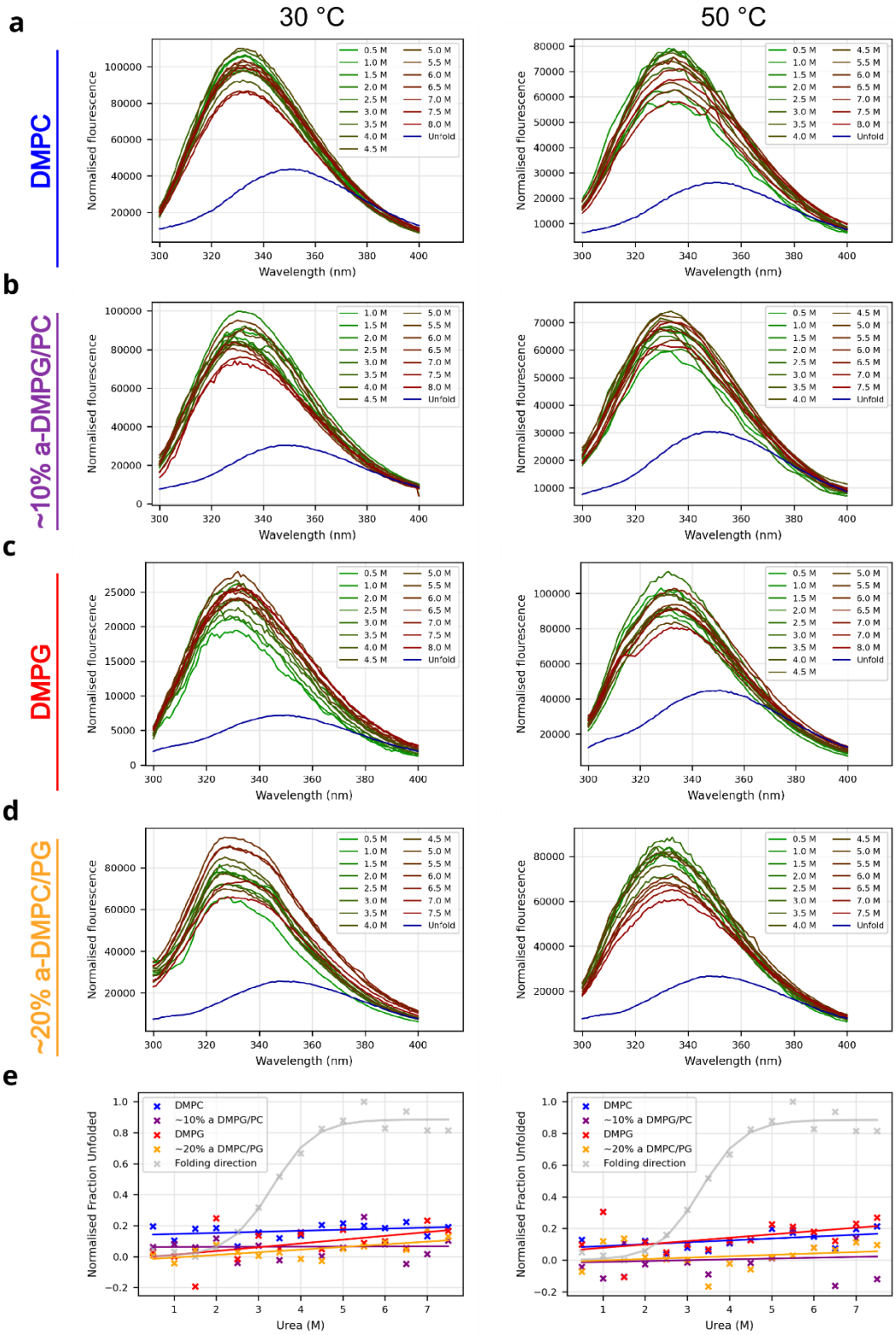


Figure A3: (Legend overleaf)

Figure A3: Urea-unfolding curves of OmpA measured by tryptophan fluorescence. (a-d) Native OmpA folded into LUVs of different lipid composition and organisation does not unfold following incubation at 30 °C (left) or 50 °C (right) in different concentrations of urea overnight. The slight reduction in intensity indicates that the liposomes have started to aggregate during the overnight reaction. The spectrum of unfolded OmpA in 7.5 M urea in the absence of lipid is shown for comparison. (e) The intensity ratio (335/350 nm) for each condition shows no evidence for unfolding of natively folded and membrane embedded OmpA at all urea conditions measured. By contrast, OmpA folds into 10:90 DMPC:DMPG LUVs with a midpoint of ~ 3.5M urea (grey) (folding was measured at 30 °C, both panels).

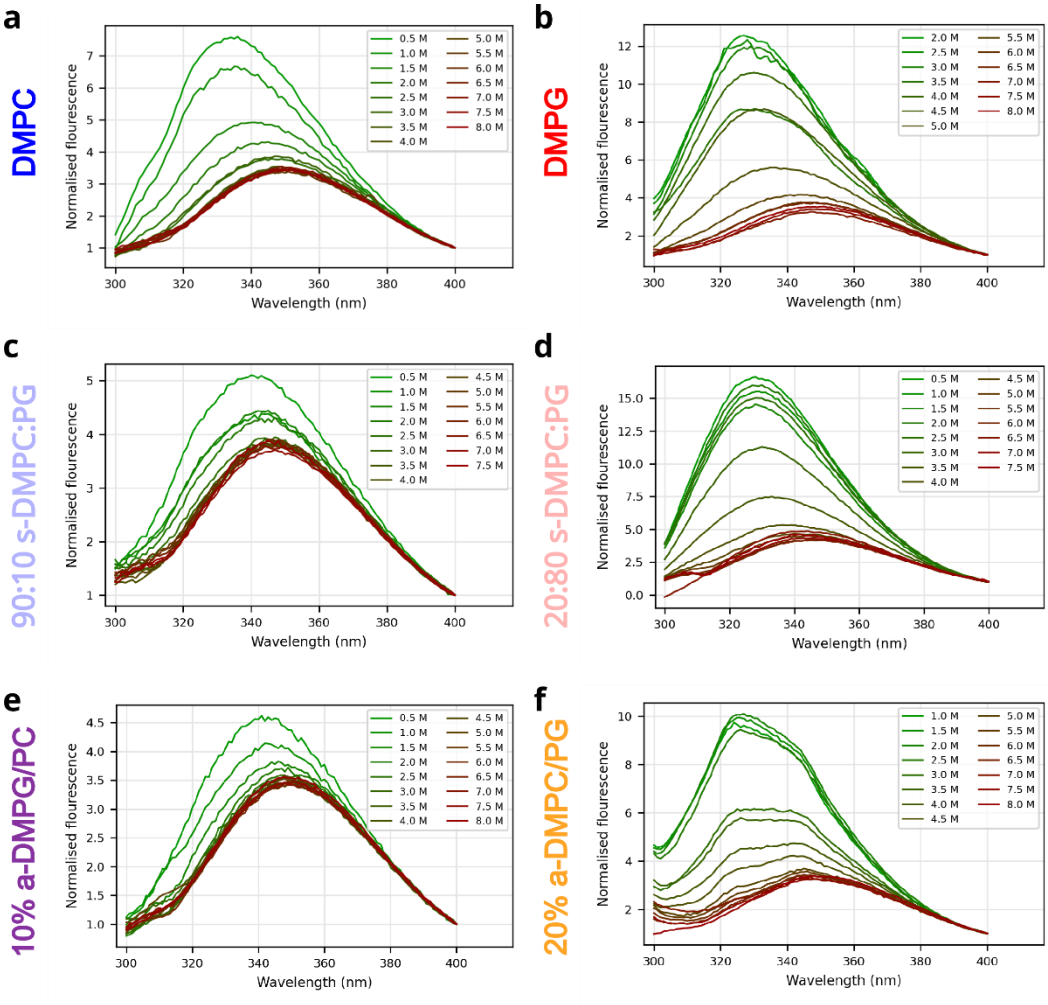


Figure A4: Urea-folding raw curves for OmpA in different liposome conditions measured by tryptophan fluorescence (c.f. Fig. 3.10d-e). (a-f) Tryptophan fluorescence emission spectra of OmpA folded into LUVS of different liposome composition in 0.5 – 8 M urea.

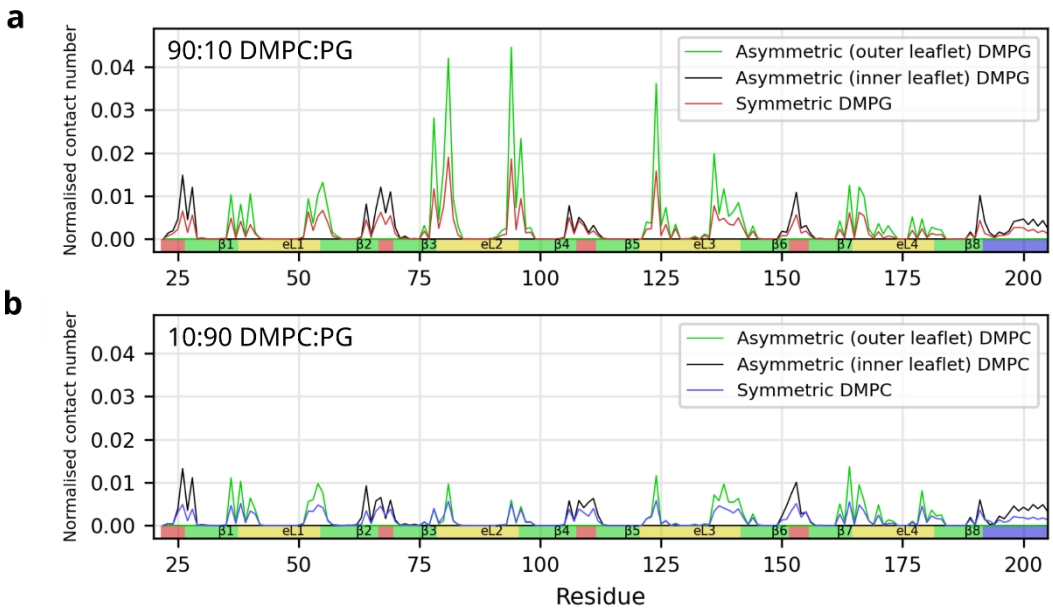


Figure A5: Lipids in the outer leaflet interact with specific residues in OMP loops in simulations of symmetric and asymmetric membranes. Normalised lipid-protein contact counts (number of interactions between each type of lipid and each residue in the transmembrane domain of OmpA normalised by lipid concentration and simulation frame number) for a total of 10% symmetrically or asymmetrically distributed lipid, as indicated in the legend for **(a)** DMPG (in DMPC base membranes) and **(b)** DMPC (in DMPG base membranes). The total lipid composition, indicated on each panel, was the same in both symmetric and asymmetric membranes, only protein-lipid interaction data for the 10% supplemented lipid is shown. The data show that DMPG interacts with OmpA's R81, K94 and R124 in the outer leaflet of asymmetric and symmetric membranes.

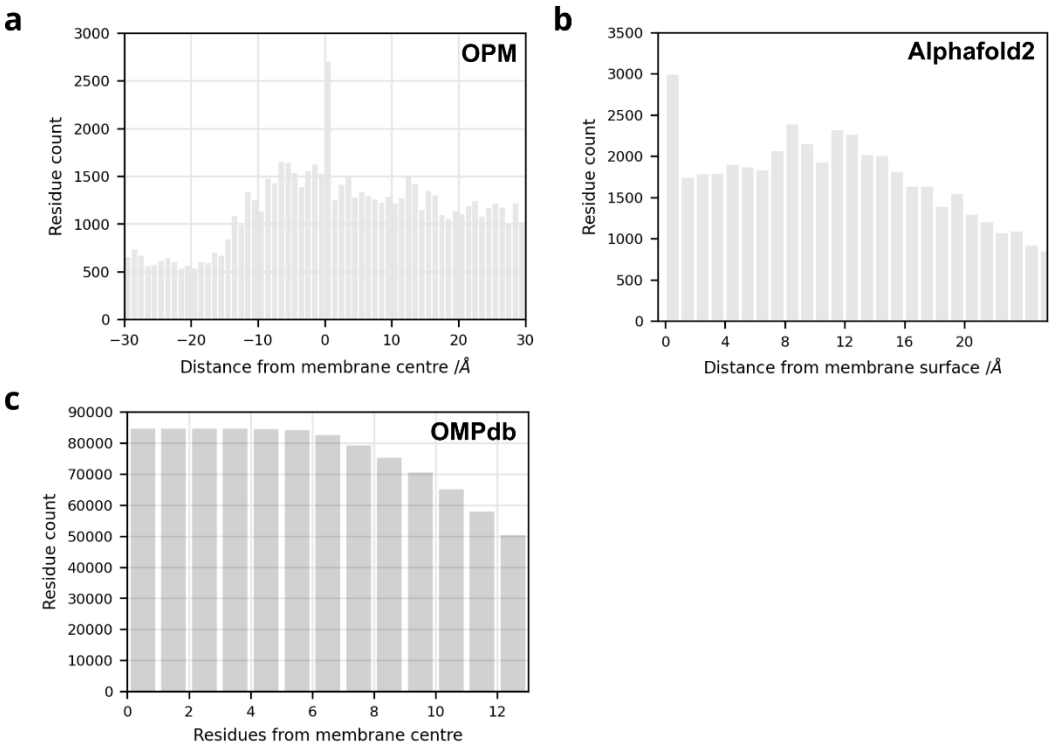


Figure A6: **Number of residues per z-axis slab for (a) OPM, (b) Alphafold2 and (c) OMPdb residue enrichment analysis**. Residue count used to calculate enrichments in each z-axis slab (per Å for OPM and Alphafold2 datasets, per residue for OMPdb).

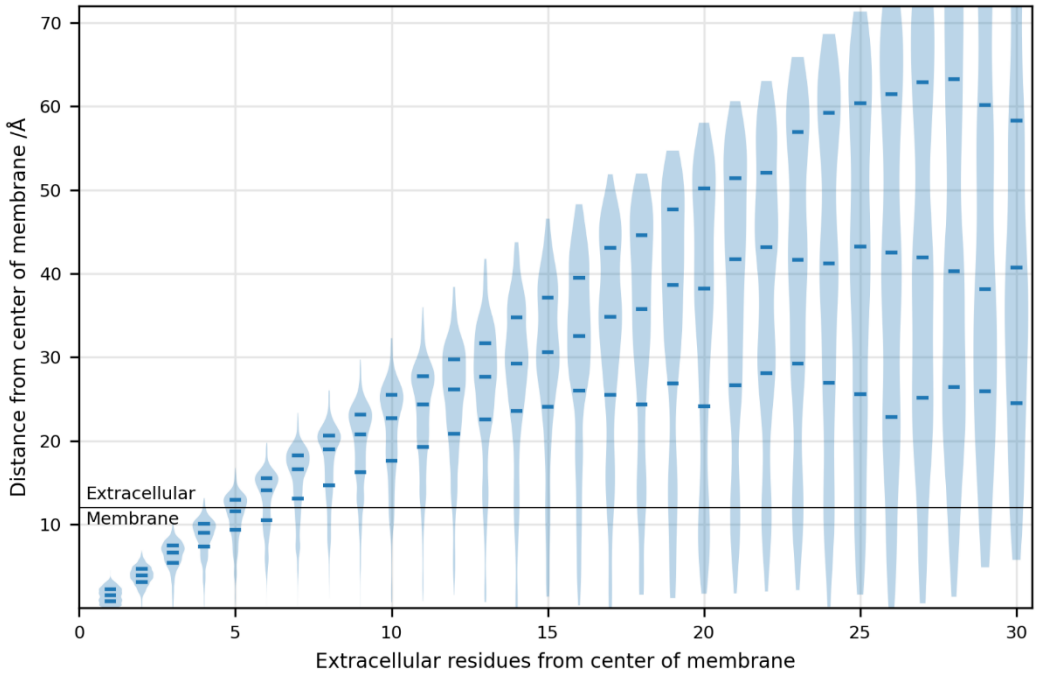


Figure A7: Calibration curve to estimate the distance from the membrane of an OMP sequence, calculated from non-redundant OMPs in the OPM database. Violin plots indicating the average distance from the centre of the membrane for the Cα of residues in extracellular loops in the OPM dataset (n = 75), showing the average and first and third quartile. Loops were truncated at their midpoints, and each part considered separately

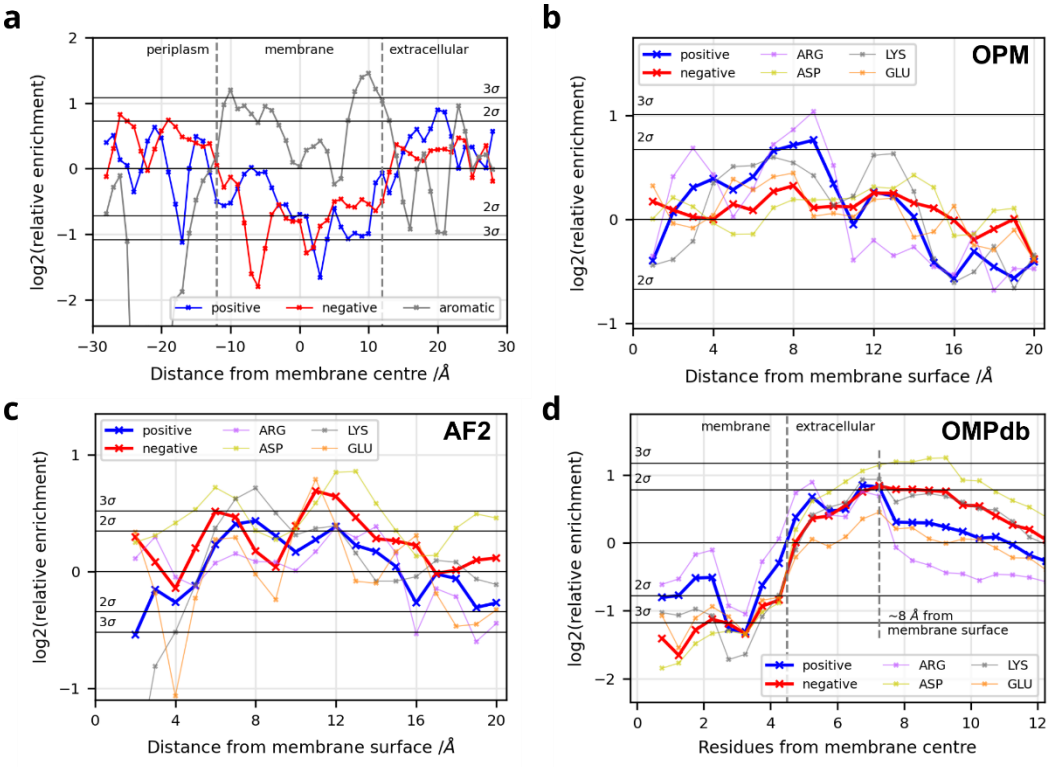


Figure A8: **All charged residue enrichments relative to membrane proximity.** Average enrichment of positively and negatively charged residues, and the underlying enrichment of arginine/lysine and aspartic/glutamic acid relative to random chance for the **(a-b)** OPM, **(c)** Alphafold2 (AF2) and **(d)** OMPdb datasets.

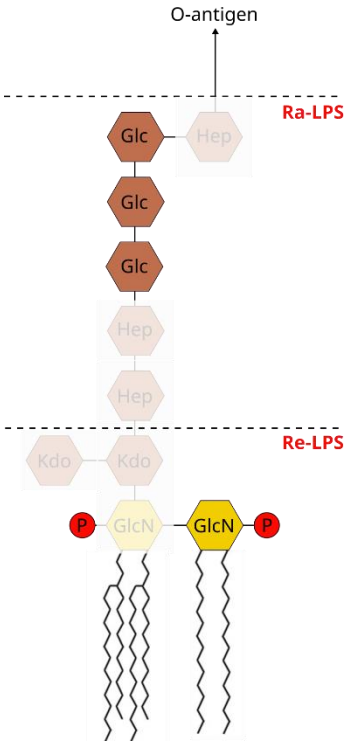


Figure A9: **LPS sugar reactivity to the anthrone assay.** Un-grayed out sugars are expected to react. Hexose and pentose sugars (i.e. not heptoses of GlcN) generally react under acid in the presence of anthrone to form furfural. The inner GlcN sugar is non-reactive to anthrone due to the scar from the acid cleavage of its GlcN and/or Kdo bond.

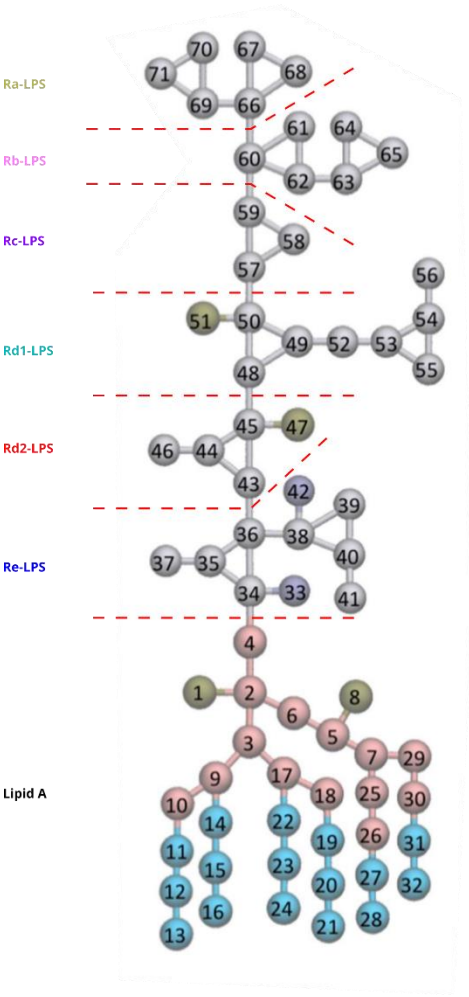


Figure A10: CG-LPS chemotype bead designation. CG-LPS model splitting into different chemotypes of LPS, and bead designation for each group.

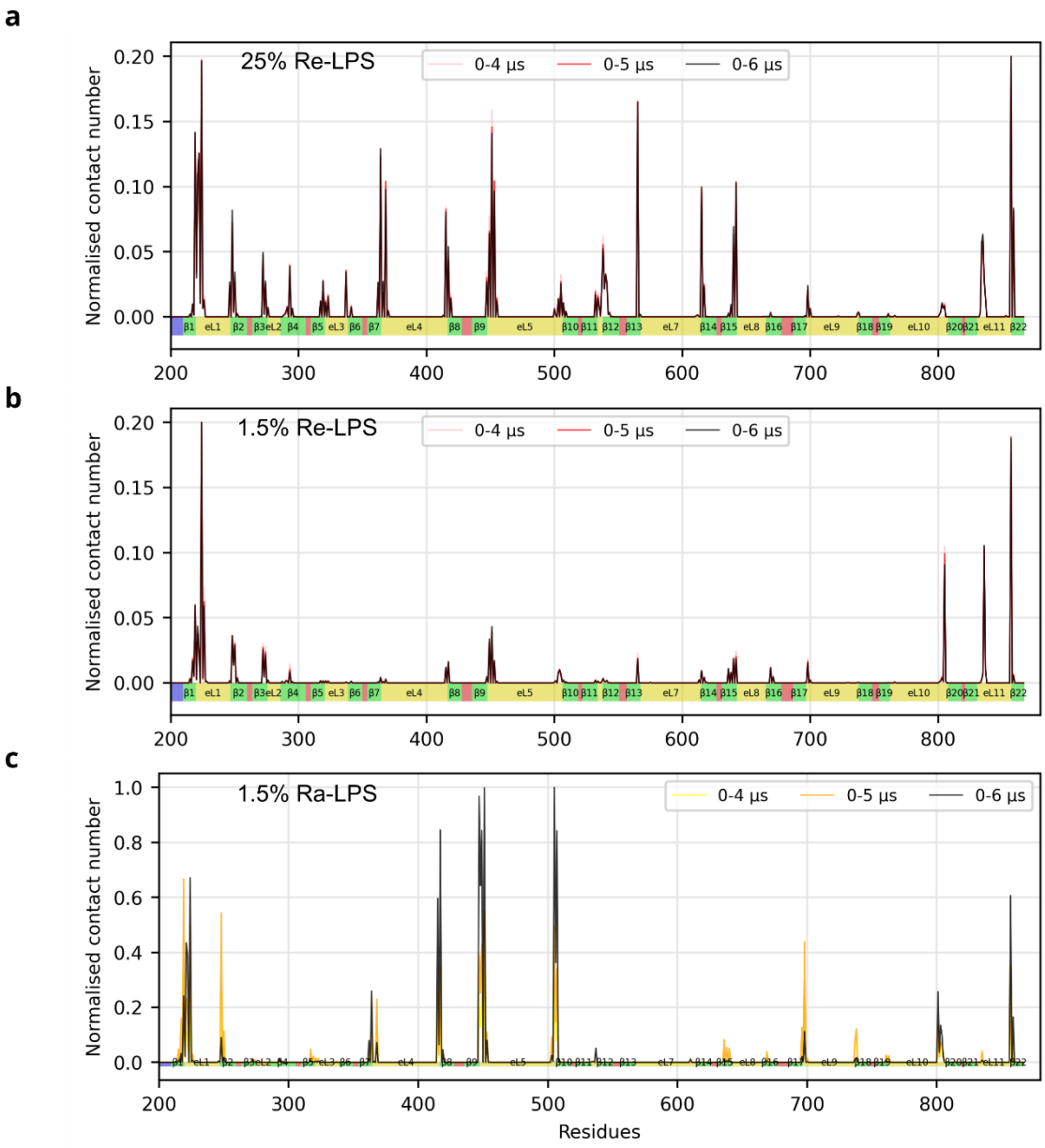


Figure A11: Re-LPS and Ra-LPS contacts with FusA largely converge after 6 µs. The normalised lipid-protein interactions (number of interactions between each type of lipid and each protein residue normalised by lipid concentration and simulation frame number) are approximately the same for the last 3 µs of simulations with outer leaflets containing **(a)** 25% Re-LPS, **(b)** 1.5% Re-LPS, and **(c)** 1.5% Ra-LPS, indicating contact convergence and system equilibration.

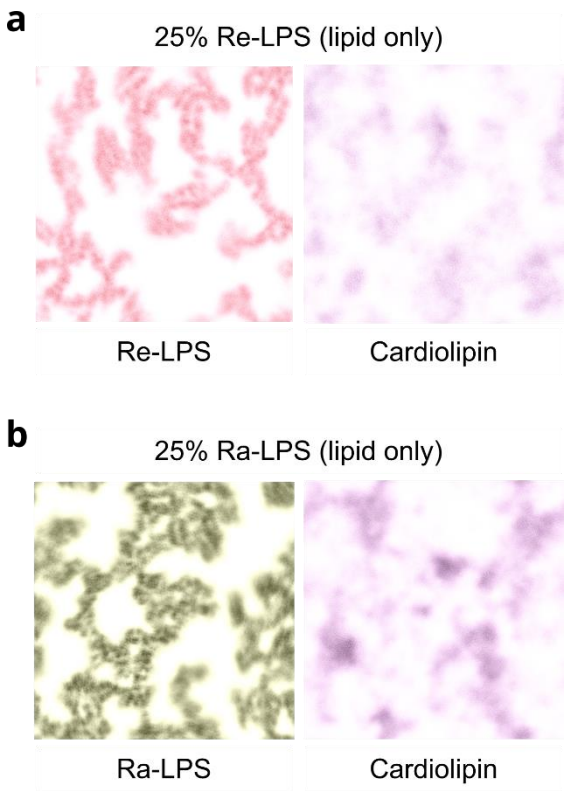


Figure A12: Cardiolipin is excluded from under LPS. Cardiolipin (purple, right) is preferentially excluded from existing under LPS patches of either **(a)** Re-LPS (red, left) or **(b)** Ra-LPS (yellow, left). Increasing colour intensity indicates higher concentrations of lipid. An inverted pattern between the LPS and cardiolipin pattern is observed in both **(a)** and **(b)**.

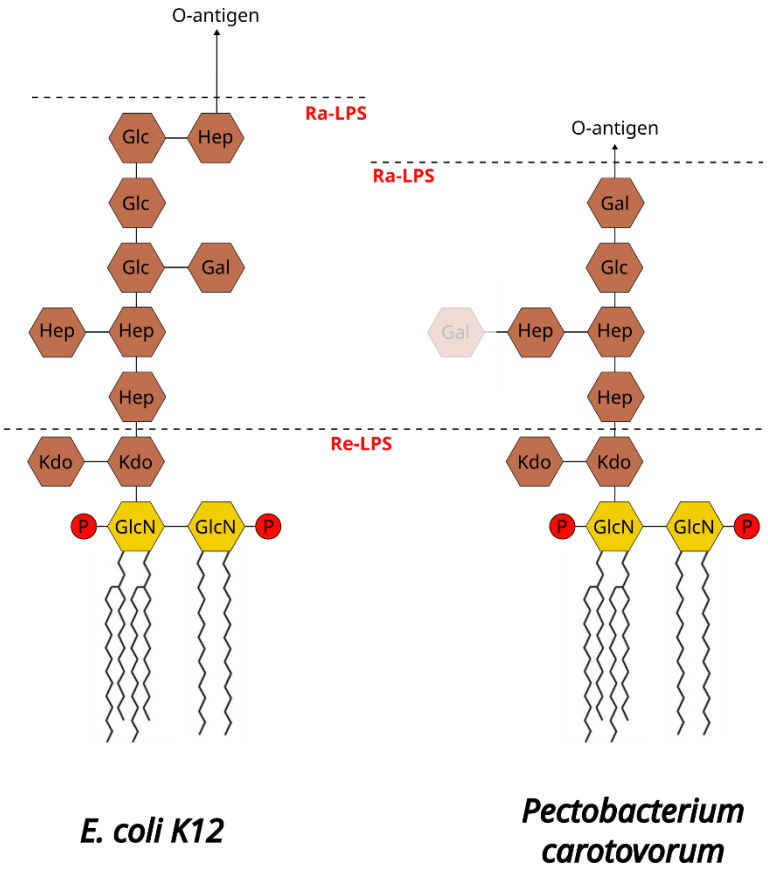


Figure A13: Comparison of K12 *E. coli* and *Pectobacterium carotovorum* LPS. The half-coloured Gal in *Pectobacterium* was identified substoichiometrically. (GlcN: D-glucosamine, Kdo: 2-keto-3-deoxyoctulosonate, Hep: heptose, Glc: glucose, Gal: galactose).

Figure A14: CryoEM processing data for FusA:Ra-LPS. Image processing workflow of Fusa:Re-LPS. Micrographs **(a)** were picked using crYOLO's general model, and the resulting protein-like 2D classes **(b)** were used to train a crYOLO model, particles from which were 2D classified and split into view groups to train additional crYOLO/topaz models, yielding higher particle counts for each view **(c)**. Initial model generation **(d)** and 3D refinements **(e)** yielded broadly correct but low-resolution models. These models were used to run multiple rounds of 3D classification against all the data **(f)**. The optimised particle stack was subjected to five cycles of cryoSPARCs Non-Uniform refinement and polishing/CTF refinement in RELION, yielding the final maps **(g)**, with local resolution estimates **(h)**. **(I)** The FSC curves of the model. The masked/corrected maps do not fully drop to zero, likely due to the model resolution practically reaching Nyquist.

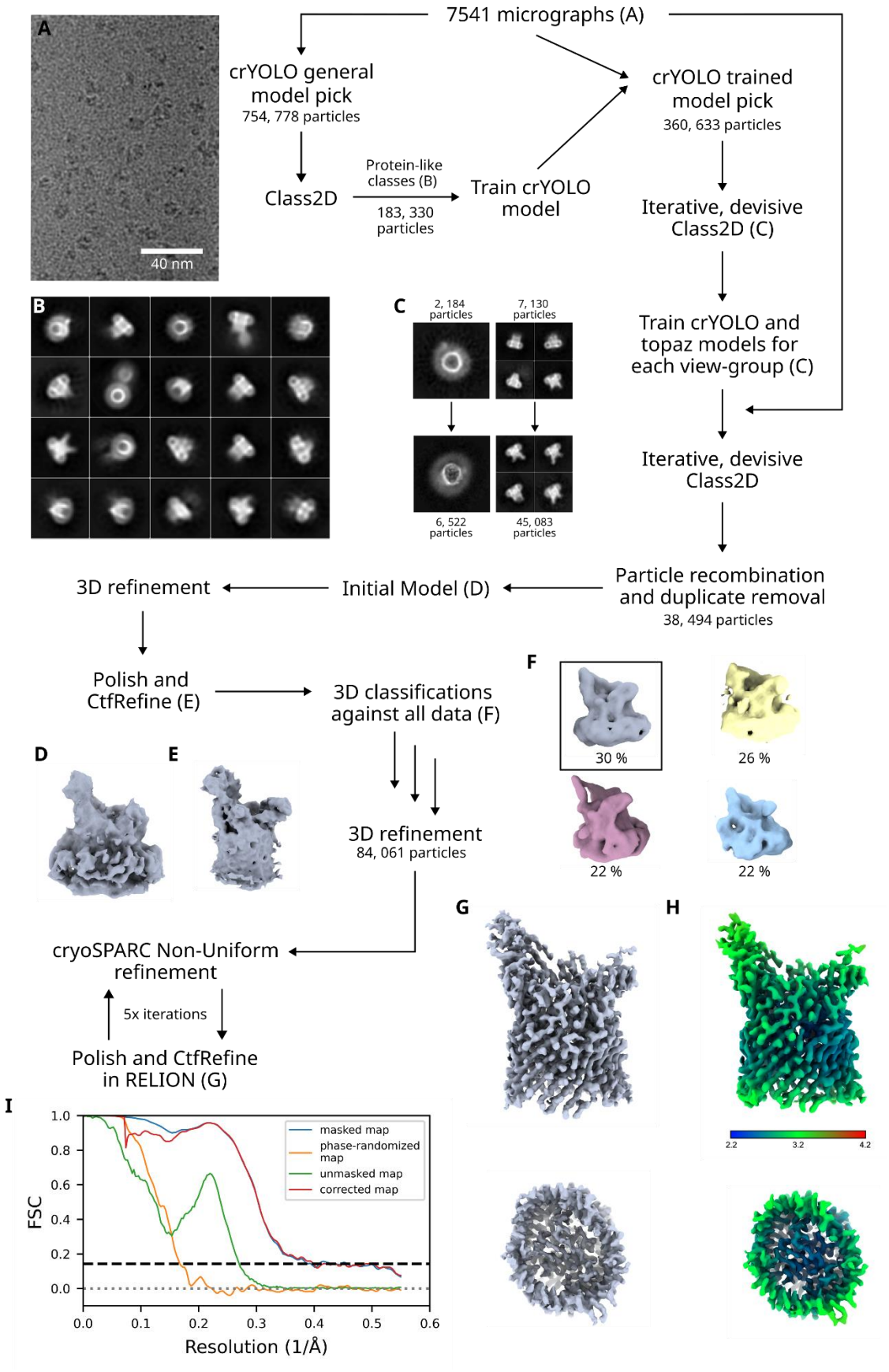


Figure A14 (Legend overleaf)

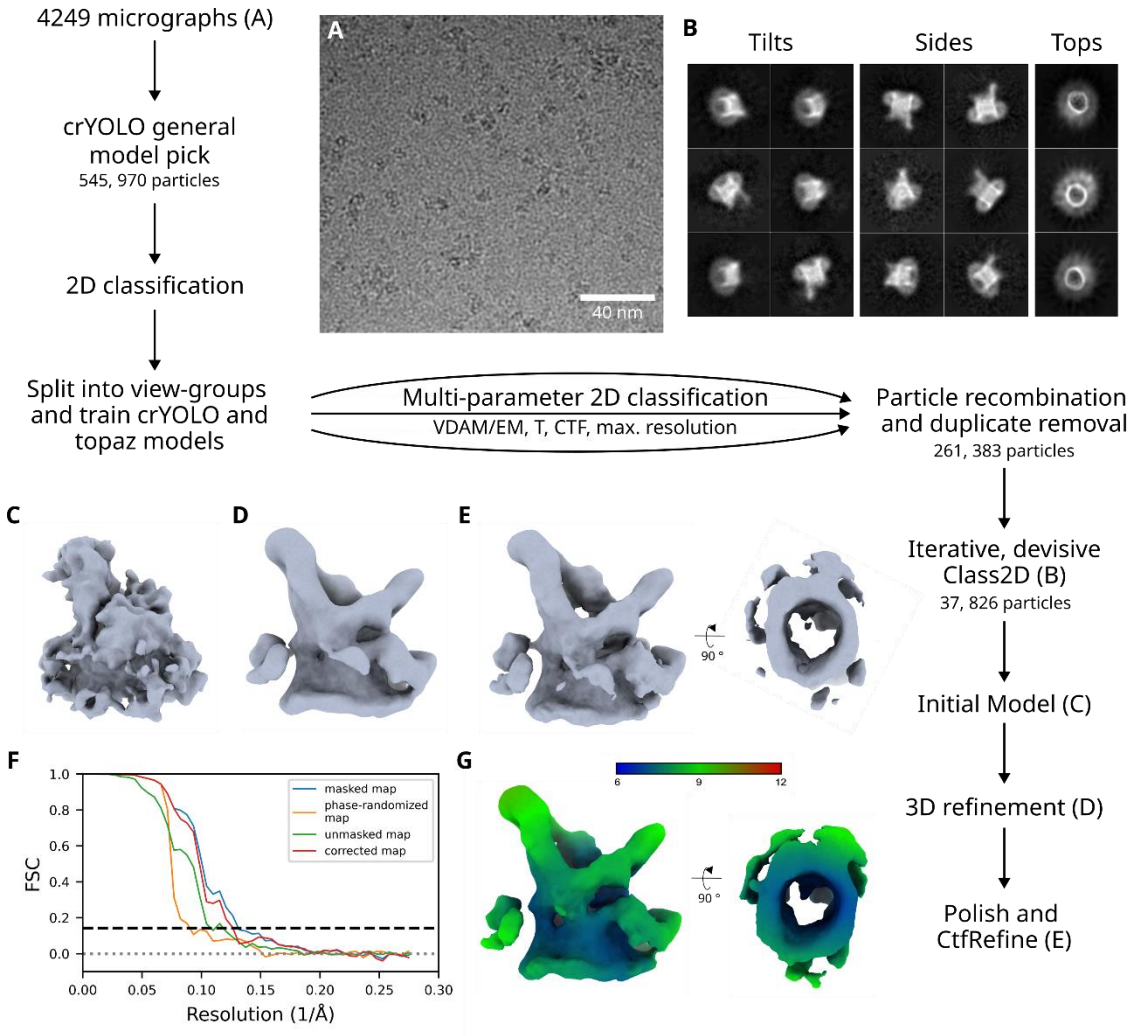


Figure A15: CryoEM processing data for FusA:Re-LPS. Image processing workflow of Fusa:Re-LPS. Micrographs **(a)** were picked using crYOLO's general model, and the resulting 2D classes were split into views and used to train new models. Following parallel multi-parameter 2D classification, good particles were combined and additional 2D classification performed, resulting in excellent classes **(b)**. Initial model generation **(c)** yielded an approximately correct reconstruction **(c)**, improved via 3D refinement **(d)**, as well as partilcle polishing and CTF refinement **(e)**. Together this resulted in a model ~8Å resolution, shown in the **(f)** FSC curves and **(g)** local resolution maps.

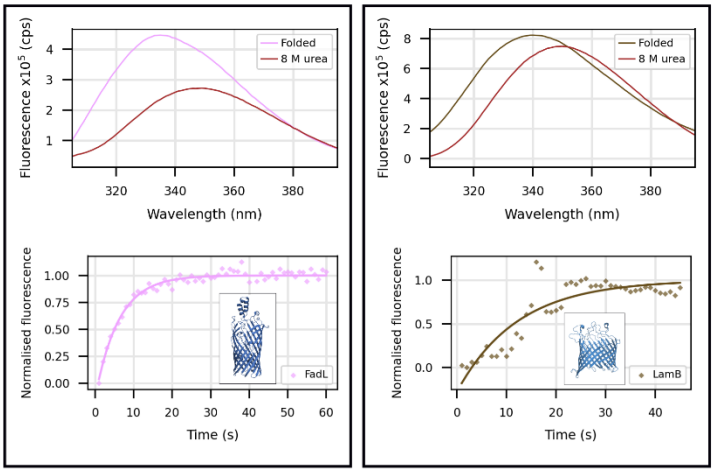


Figure A16: FadL/LamB DMPG folding into DMPG in 0.5M urea and at 23.5 °C. For both of the OMPs the folding kinetics (lower panels, followed by tryptophan fluorescence) and the folded vs unfolded (8 M urea) intrinsic fluorescence spectra are shown. Each protein shows clear, rapid folding kinetics and the expected unfolded-folded spectral transition.

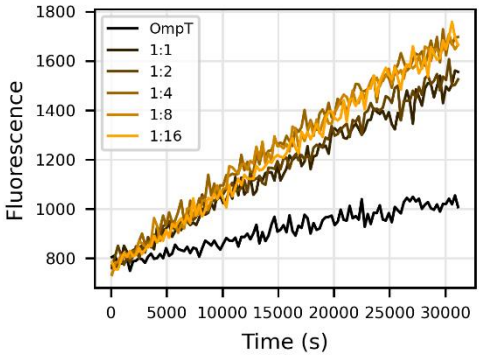


Figure A17: OmpA rate enhancement of OmpT is saturated at 1:1, even at very high OmpT:OmpA ratios. OmpA mediated rate enhancement of OmpT up to 1:16 OmpT:OmpA ratios.

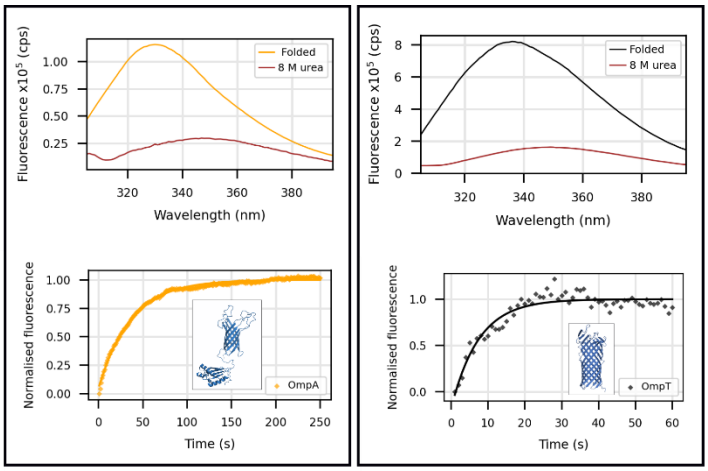


Figure A18: OmpA and OmpT fold well into DLPC membranes at 0.5 M and 30 °C. For both OMPs the folding kinetics (lower panels, followed by tryptophan fluorescence) and the folded vs unfolded (8 M urea) intrinsic fluorescence spectra are shown. Each protein shows clear, rapid folding kinetics and the expected unfolded-folded spectral transition.

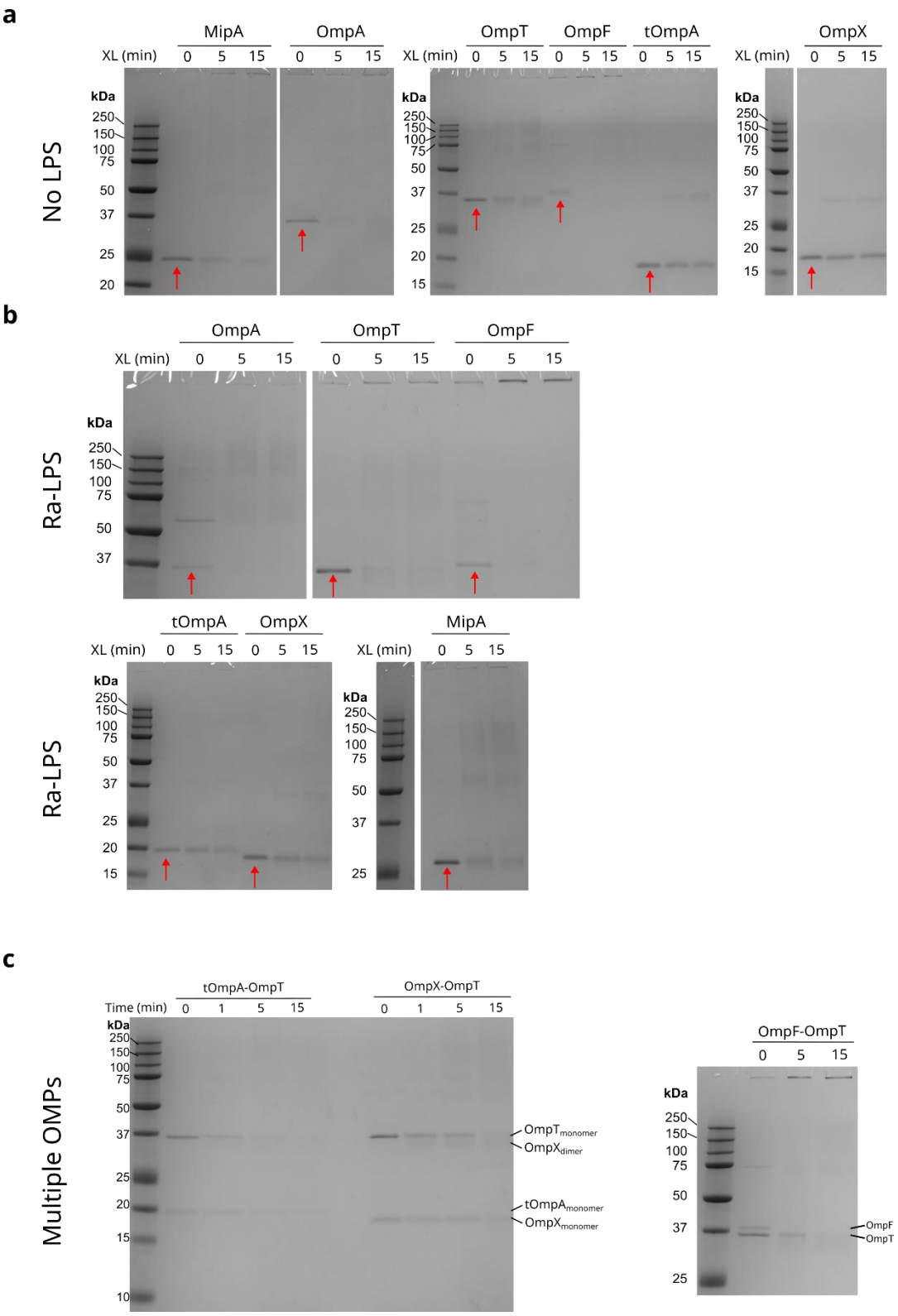


Figure A19: Sample raw crosslinking gels for (a) DMPG, (b) DMPG:Ra-LPS, and (c) multiple OMPs. Note the presence of heavily crosslinked material that did not enter the gel in some lanes, for example OmpF.

Appendix B: Supplementary Tables

Supplementary Table 1: CryoEM data collection, refinement, validation and model building statistics for FusA:Ra-LPS (collection 8)

	FusA:Ra-LPS
Data collection and processing	
Magnification	130k
Voltage (kV)	300
Electron exposure (e–/Å^2)	47.1
Defocus range (μm)	-0.9 to -3.0
Pixel size (Å)	0.91
Symmetry imposed	C1
Initial particle images (no.)	754 778
Final particle images (no.)	84 061
Map resolution (Å)	2.4
0.143 FSC threshold	
Map resolution range (Å)	2.2-4.2
Refinement	
Initial model used (PDB code)	4ZGV
Model resolution (Å)	3.2
Map sharpening *B* factor (Å^2)	Local sharpening
Model composition	
Non-hydrogen atoms	
Protein residues	809
Ligands	4
B factors (Å^2)	
Protein	0
R.M.S. deviations	
Bond lengths (Å)	0.003
Bond angles (°)	0.595
Validation	
MolProbity score	1.60
Clashscore	5.05
Favored rotamers (%)	98.15
Poor rotamers (%)	0
Ramachandran plot	
Favored (%)	95.17
Allowed (%)	4.83
Disallowed (%)	0

Table A2: CryoEM data collection statistics for low-resolution reconstructions

	Collection 1 K2 detector Gatan filter (20 eV)	Collection 2 K2 detector Gatan filter (20 eV)	Collection 3 F4 detector	Collection 4 F4 detector DDM-FusA	Collection 5 F4 detector Low defocus	Collection 6 F4i detector Selectris (5 eV)	Collection 7 F4i detector Selectris (5 eV) 10x Re-LPS
Data collection and processing							
Magnification	130k	130k	130k	96k	96k	130k	130k
Voltage (kV)	300	300	300	300	300	300	300
Electron exposure (e–/Å^2)	62.0	66.12	40.0	40.0	40.0	45.11	47.1
Defocus range (μm)	-1.0 to -2.5	-1.0 to -2.5	-1.0 to -3.0	-1.0 to -3.0	-0.25 to -2.0	-0.9 to -3.0	-0.9 to -3.0
Pixel size (Å)	1.07	1.07	0.91	0.83	0.83	0.91	0.91
Symmetry imposed	C1	C1	C1	C1	C1	C1	C1
Micrograph number	892	3235	3584	1336	1846	7511	4249
Initial particle images (no.)	72 358	286 338	279 457	377 986	236 914	711 757	545 970
Picking model training particle images (no.)	37 132	5 804	54 706	17 369	28 157	-	-
Model picked particles (no.)	361 633	417 907	896 632	377 986	214 571	-	-
Final particle images (no.)	10 623	31 357	76 383	31 872	32 329	74 343	37 826
Map resolution (Å) 0.143 FSC threshold	13	8.6	9.7	8.6	18	8.3	7.2