MONASH University

The assembly of fimbrial ushers, inverse autotransporters and LptDE into bacterial outer membranes

Christopher James Stubenrauch

BSc (Hons)

A thesis submitted for the degree of Doctor of Philosophy at

Monash University in 2016

Faculty of Medicine, Nursing and Health Sciences

Department of Microbiology

Declaration

This thesis contains no material which has been accepted for the award of any other degree or diploma at any university or equivalent institution and that, to the best of my knowledge and belief, this thesis contains no material previously published or written by another person, except where due reference is made in the text of the thesis.

Signature: 

Print Name: Christopher James Stubenrauch

Date: 11 August 2016

Abstract

The translocation and assembly module (TAM) is capable of assembling outer membrane proteins, but its function has remained enigmatic in that: (i) only the AIDA-like autotransporter family had been shown to require the TAM, and (ii) outer membrane protein biogenesis is generally considered to be catalysed only by the β-barrel assembly machinery (BAM) complex. To assess whether the TAM plays a broader role in β-barrel assembly, I developed a pulse chase assay to characterise the specific contributions of the TAM and the BAM complex for a range of outer membrane protein substrates.

Working on the hypothesis that the TAM assembles more complicated outer membrane proteins that diverge from the classical β-barrel architecture, five candidate TAM substrates were analysed for their requirement of the TAM: FimD (a fimbrial usher), intimin (an inverse autotransporter), TolC, PhoE, and LptDE. While the PhoE data was inconclusive, the four remaining candidate substrates were found to have varying requirements for the TAM. Furthermore, six related fimbrial ushers and a second inverse autotransporter were also confirmed to require the TAM for efficient biogenesis.

FimD contains an unusual extracellular loop that is readily proteolytically degraded on addition of exogenous proteinase K. Because distinct proteolytic fragmentation "fingerprints" were generated depending on the presence or absence of the TAM, this provided a tool to dissect the contribution of the TAM toward FimD biogenesis. The TAM was subsequently found to initiate β-barrel assembly from the C-terminal end of the substrate, and while the BAM complex

was also capable of assembling FimD, it was significantly less efficient because it initiated β-barrel assembly from a central region of the protein. This data was interpreted to mean that the BAM complex is subsidiary to the TAM during fimbrial usher biogenesis.

In contrast, although assembly of the LptDE complex was also shown to be assisted by the TAM, considering both the BAM and LptDE complexes are essential (but the TAM is not), the TAM likely plays a minor role in LptDE assembly that is subsidiary to the BAM complex. As such, the hypothesis that only the BAM complex is necessary for β-barrel biogenesis should be extended to include the contribution that TAM makes - depending on the substrate - as either the major or minor outer membrane protein insertase.

Table of Contents

Preface

Some of the work undertaken toward this degree has already been published.

Stubenrauch, C., Belousoff, M. J., Hay, I. D., Shen, H.-H., Lillington, J., Tuck, K. L., Peters, K. M., Phan, M.-D., Lo, A. W., Schembri, M. A., Strugnell, R. A., Waksman, G. & Lithgow, T. (2016). Effective assembly of fimbriae in *Escherichia coli* depends on the translocation and assembly module nanomachine. *Nat Microbiol* **1**, 16064.

Heinz, E.*, **Stubenrauch, C. J.***, Grinter, R., Croft, N. P., Purcell, A. W., Strugnell, R. A., Dougan, G. & Lithgow, T. (2016). Conserved features in the structure, mechanism, and biogenesis of the inverse autotransporter protein family. *Genome Biol Evol* **8**, 1690-1705.

*Joint first authors

All experimental work published in the above papers was performed by me, except for the following:

M.J. Belousoff in Stubenrauch *et al.* (2016):

-Performed some pulse chase analyses of TolC and PhoE biogenesis corresponding to one of the biological replicates for each strain.

-Performed densitometry on TolC and PhoE autoradiographs and determined all observed rate constant.

I. D. Hay in Stubenrauch *et al.* (2016):

-Performed the functional fimbrial assays (Figure 3c, therein).

K. M. Peters, M-D Phan and A. W. Lo in Stubenrauch *et al.* (2016):

-Performed the yeast agglutination assays (Figure 3b, therein).

Monash Biomedical Proteomics Facility in Stubenrauch *et al.* (2016).

-Analysed FimD (and its fragments) by LC-MS/MS, excluding the corresponding SDS-PAGE analysis of FimD biogenesis and subsequent in-gel tryptic digestion, which were performed by me.

E. Heinz in Heinz *et al.* (2016):

-All bioinformatic analyses

R. Grinter in Heinz *et al.* (2016):

-Primary sequence alignments and generation of structural homology models

N. P. Croft in Heinz *et al.* (2016):

-Analysed FdeC fragments by LC-MS/MS, excluding the corresponding SDS-PAGE analysis of FdeC biogenesis, which was performed by me.

Acknowledgments

Firstly, I would like to thank my supervisors: Prof. Trevor Lithgow, and secondary supervisors Dr. Matthew Belousoff and Dr. Eva Heinz. Trevor has been a constant stream of positive energy and has such a way with words that whenever I'd lost my confidence, I could always find it lurking in his office. Matt was my main lab mentor and was a constant source of cheerfulness. He has such a distinctive laugh that you could always tell when he came into the lab. When I first started in the Lithgow lab, the project I was working on seemed to be a dead-end; Matt B. (and Trevor) conceptualised a new project, which formed the basis for my current thesis, to which I am forever grateful. Eva was the bioinformatics genius that our lab sadly (for us) donated to the Sanger Institute. Her work ethic (something like 14-16-hour days) was inspirational, and the discussions we had about bioinformatics always left me thinking I was just a little bit smarter every time.

I came to Trevor's lab when he was part of the Biochemistry department at Monash, and I would like to think my undergraduate degree in Microbiology was somehow the tipping point that prompted Trevor to transfer into the Microbiology department. Sadly, I have yet to convert him from the clan of Carlton Blues to the mighty Essendon Bombers faithful – little does he know that the constant imagery from my fan merchandise during past and future lab meetings is slowly chipping away at the logic he first used when deciding to not support Essendon. Fortunately, I met my secondary supervisor Matt before I knew that he was a Hawthorn Hawks supporter; had I have known before-hand… I actually first met Matt during my undergraduate degree. I wonder if he remembers me from then? He was my Chemistry demonstrator when Dr. Kellie Tuck had just started

lecturing for and convening a 2nd-year Chemistry unit. Matt used to love scaring the undergraduate students: he sometimes brought weird safety equipment in (swimming goggles, etc) to explain how horrible some of the "actually-not-so-bad" chemicals were. When Eva first moved to the Sanger, I "misheard" Trevor during lab meeting one day and thought he said that "Eva" had been in a horrible accident. Following several confusing communications with her, it was clear I had made a fool of myself. Thanks for not rubbing it in too much.

I would like to thank my Honours supervisors: Prof. Julian Rood and Dr. Trudi Bannam. I did the undergraduate MIC3990 research project in the Rood lab before attempting to join the Victorian police force after graduation. At the time, there was a massive push for people to join the police and I thought I would enjoy helping the community. In the end, I did not pass and fortunately, when I came crawling back, Julian took it in his stride and immediately explained that they said "no" because I was "too smart" for them. To this day, that is what I tell everyone when I tell them I tried for the police. Julian was a strict Microbiologist and taught me the importance of biological replicates and fulfilling molecular Koch's postulates through gene complementation. Trudi was my lab mentor, who was sadly donated to Qiagen. She was a very experienced supervisor that taught me, with utmost patience, how to work efficiently and effectively in a lab.

I would like to thank the Lithgow lab members both past and present. Even since I began my Ph. D., several members have left the lab to continue on into greatness, and I believe it is the "Lithgow" lab culture and publication record that has allowed them to do this. Firstly, I would like to thank my fellow Ph. D. students: (now) Dr.

Victoria Hewitt, (now) Dr. Rhys Dunstan, Matthew Stellato, Sachith (Dilshan) Gunasinghe and Von Vergel Torres. Victoria was the Lithgow lab golden child before she headed for greatness overseas. I find it interesting that Von decided to then dub me as the "next Victoria" or the "new Lithgow lab golden child", and although he probably means I'm pedantic and somewhat annoying, I take it as a great compliment to be compared to someone like Victoria. Rhys D. was my office desk buddy, and for a while we shared a four-person office space between us in great luxury. Although he has since moved into a new office space (for Post-docs), I'll always remember the awesome chats we had about music and sport and have to confess I was super jealous of the data generated for his thesis. Without incriminating myself, I would like to say that Von has been an excellent source of memes generated from our Lithgow lab. He has yet to find some of them on our Lithgow lab photo wall, including those generated from when he "accidentally" told us he looked like the fashion consultant Gok Wan or Mowgli from the Jungle Book.

I would like to thank the post-docs: Dr. Rhys Grinter, Dr. Iain Hay, Dr. Denisse Leyton, Dr Hsin-Hui Shen, Dr. Miguel Shingu-Vazquez, Dr. Takuya Shiota and Dr. Chaille Webb. I remember thinking when Rhys G. came, "why are there two Matt's, and now two Rhys's, but only one Chris?" I thought my name was popular, and I'm sure Von said something tongue-in-cheek like "there can only be one Chris". Since Iain came to our lab, we have had many humourous and intelligent conversations. I asked him, how he would like to be acknowledged herein and from his response… Iain David Hay is my mighty saviour, without him I would be nowhere… was that okay? Dr. Denisse Leyton was donated to the ANU, and was my co-conspirator in

organising the Monash SOBS footy tipping competitions. She always made me laugh, especially from the voices she put on when presenting her work; I get that it was a professional tone, but I could never adjust from your jovial personality to your serious one. Thank you Shio and Hsin-Hui for all of your patient explanations when I struggled to understand your specialist techniques. You guys are brilliant scientists, and it has been a privilege to work next to Shio in the lab. Thank you to Miguel for his humour, I feel like you were the most mischievous one of the group and in your absence I believe I have done you proud. Chaille for your humour and shared love for musicals. I know you were one of the only ones that put up with my seemingly random lab-day music decisions, from rock to musical to classical to heavy metal. Naturally, it was also a privilege to work opposite you in the lab.

I would also like to thank my collaborators that contributed to my publications. Firstly, a big thank you to Denisse for her help with the pulse chase assay. She had perfected the technique for her protein of choice and graciously passed down the technique so that I could optimise it for my favourite proteins. Thank you to Hsin-Hui for your AFM expertise; thank you to Iain for the fimbrial binding assays and thank you to the Schembri lab. Thank you Oded Kleinfeld for your help with the tryptic digests and, together with Rob Goode, your analysis of the FimD LC-MS/MS data. Thank you also to Nathan Croft and the Purcell lab for performing the tryptic digests and analysis of FdeC LC-MS/MS data. Also, and obviously, thank you to my supervisors for their help in putting together those manuscripts.

Thank you also to Gilu Abraham, Rebecca Bamert, Tricia Lo, Kher Shing Tan, and Nora Tenis. Without people like you to support the Lithgow, Traven and Naderer

labs, the labs would have probably imploded already. You have always been helpful, but I don't know if I have thanked you enough! Especially the cakes and delicious goodies Bec always brings, Bec's assistance and explanation of protein purification, and Kher Shing's assistance with everything and anything around the lab. Thank you also to Sam Palframan and Julie Nguyen, the unofficial Lithgow lab members (from other labs). You have always been fun and a joy to work with, sit near and chat to.

Last, but not least, I would like to thank my friends and family, especially my wife. On my science journey, Caitlin has always lovingly supported and encouraged me and now that I am almost finished, she appears to be more excited than I am. We got married during my candidature, and she has always been my rock, despite sometimes picking up that same rock and throwing it at me. Whenever we meet new people and are asked "what do you do?", I encourage Caitlin to answer, to determine if my explanations to her were good enough. She typically compares me to a chef: "He cooks a broth, he follows recipes, he runs them through a gel", but after prompting, she usually gets it right: "He looks at how proteins cross membranes". Likewise, I explain what she does: among the many roles she performs as an occupational therapist, she assesses a patient's ability to plan and prepare meals, but not for the "free" lunch… she also plays the Nintendo Wii to improve balance and motor-control in patients, naturally in two-player-mode, which is why the games may need to be practised at home… is that a fair assessment? Thank you for all of your support and love along the way; hopefully, I can be as good a support for you now that you've decided to complete your Masters of Healthcare Leadership.

Thank you to my parents, Bruce and Cathrine; my brothers, Josh and Nathan, and; my favourite in-laws: Ben, Madeleine and Marijke. Thank you for your love and support; every Tuesday or Friday I have an excellent break at one of your houses where we can chat about anything and everything. The movie nights and AFL games we attend are always fun, even if my team is slightly struggling to find the ball. Thank you to Maddy for the fun weekend brunches we have, and when you can explain all the things wrong with the world so simply. Thank you to Meggan for the fun nights we have when you sleep over; you really put the "egg in Man".

Thank you to my friends, Glenn, Wee-man, Gerry, Grindlay, Squeak, Squire, Nolan, Rossco and Kitch. I would be remiss to exclude their rocks and my friends, especially Sarah, Steph, Rhian, Brooke, Jess, Chloe and Lauren. I really do enjoy my breaks from science when we all hang out. A big thank you especially to best man Glenn, pretty much the organiser of the group; we're all looking forward to your new house and the resumption of boys' nights there. As requested, here is an even bigger thank you for Wee-man and Steph: although we don't play futsal together anymore Wee-man, our weekend catch-ups are always fun, even if the former professional cheerers (Steph and Caitlin) are there too. Thank you everyone for your genuine interest in my thesis and your generosity in sharing your houses and cooking, even if it's as simple as calling the local pizza shop. I love you all!

List of Figures

List of Tables

Abbreviations

Abbreviations may be defined in text in plural, but are shown here in singular.

ACP	acyl carrier protein
ADP	adenosine diphosphate
ATP	adenosine triphosphate
BAM	β-barrel assembly machinery
Big	Bacterial immunoglobulin-like
CDP	cytidine diphosphate
CTD	C-terminal domain
CTD1	C-terminal domain 1
CTD2	C-terminal domain 2
CTP	cytidine triphosphate
CL	cardiolipin
CoA	coenzyme A
CSH	Cold Spring Harbor
DAP	*meso*-diaminopimelic acid
FAS2	type 2 fatty acid biosynthesis
G3P	glycerol-3-phosphate
gDNA	genomic DNA
Hep	L-glycero-D-manno-heptose
HRP	horse-radish peroxidase
IM	inner membrane
Kdo	keto-deoxyoctulosonate
LC-MS/MS	liquid chromatography-tandem mass spectrometry
LOS	lipooligosaccharide

Lpp	Braun's lipoprotein
LPS	lipopolysaccharide
MCS	multiple cloning site
MES	2-(*N*-morpholino)ethanesulphonic acid
MIC	minimum inhibitory concentration
NAG	β-*N*-acetylglucosamine
NAM	β-*N*-acetylmuramic acid
NTD	N-terminal domain
OD_{280}	optical density at 280 nm
OD_{600}	optical density at 600 nm
OM	outer membrane
OMP	outer membrane protein
PA	phosphatidic acid
PAGE	polyacrylamide gel electrophoresis
PAI	pathogenicity-associated island
PDB	protein data bank
PE	phosphatidylethanolamine
PG	phosphatidylglycerol
P_i	inorganic phosphate
PK	proteinase K
PPIase	peptidyl-prolyl *cis-trans* isomerase
POTRA	polypeptide transport associated
RNC	ribosome-nascent chain
SDS	sodium dodecyl sulphate
SN	semi-native

SPase	signal peptidase
SPATE	serine protease autotransporter of *Enterobacteriaceae*
SRP	signal recognition particle
T1SS	type 1 secretion system
T2SS	type 2 secretion system
T3SS	type 3 secretion system
T4SS	type 4 secretion system
T5SS	type 5 secretion system
T6SS	type 6 secretion system
TAM	translocation and assembly module
Tat	twin-arginine translocation
TEV	tobacco etch virus
UDP	uridine diphosphate

Chapter 1 – Cell envelope biogenesis

1.1 Introduction

Escherichia coli is the paradigm Gram-negative bacterium used as the preferred model for protein purification, genetic modification and studying the biology of phages or bacteria (Cronan, 2001). All bacterial proteins are synthesised within the cytoplasm and constitute about 60-70 % of the dry weight of the cell (Emmerling *et al.*, 2002; Fischer & Sauer, 2003; Taymaz-Nikerel *et al.*, 2010). Roughly 70 % of proteins remain within the cytoplasm where they play diverse roles, including catalysis of fatty acid biosynthesis, DNA replication, RNA transcription and protein translation (Bernsel & Daley, 2009; Elofsson & von Heijne, 2007). The remaining 30 % of the proteome is localised to the cell envelope or secreted into the extracellular milieu (Figure 1.1.1).

Our understanding of bacterial cell biology is constantly evolving with the technology and techniques employed to understand it. Indeed, models for cell wall architecture and protein biogenesis are constantly being refined as new data contradicts accepted paradigms. The traditional belief that the cell wall adopts a uniform structure is now being overturned in favour of a non-uniform, but highly organised architecture (Nguyen *et al.*, 2015). What was once thought to be rare examples of surface-exposed lipoproteins is now known to be much more widespread (Konovalova & Silhavy, 2015; Wilson & Bernstein, 2016). The need for the β-barrel assembly machinery (BAM) complex in assembling all integral outer membrane proteins (OMPs) has recently been challenged (Dunstan *et al.*, 2015). In keeping with this evolution of our knowledge, this thesis will investigate the importance of the translocation and assembly module (TAM) - discovered by

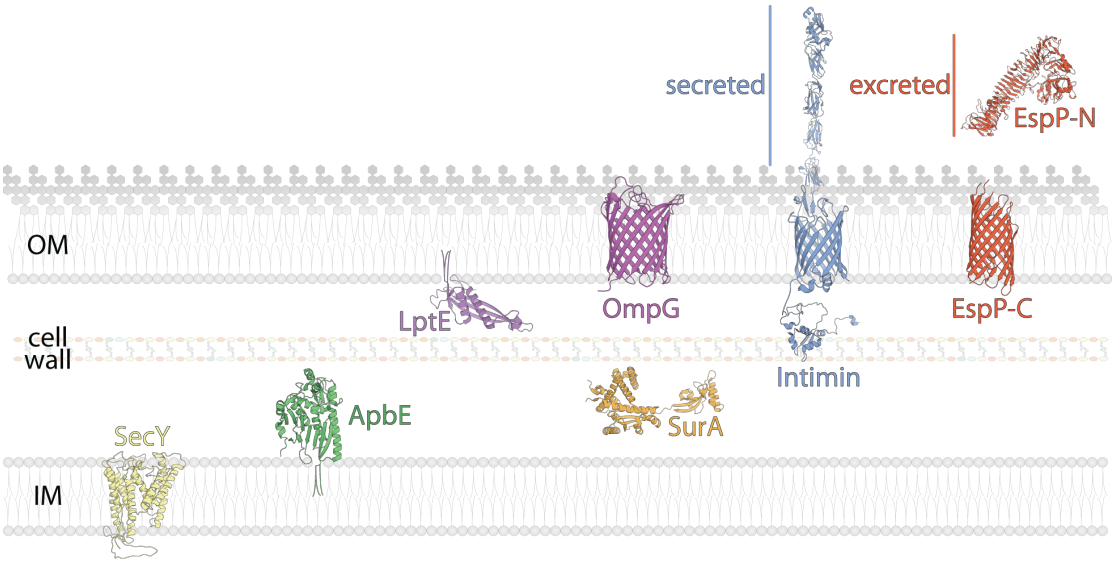


Figure 1.1.1 Extracytoplasmic proteins.

Ribbon diagrams of extracytoplasmic proteins in various sub- or extra-cellular locations. Integral inner membrane (IM) proteins, such as SecY (PDB: 3DIN), contain at least one transmembrane α-helix. Lipoproteins, such as ApbE (PDB: 2O18) and LptE (PDB: 4N4R), contain a lipid moiety embedded within the periplasmic leaflet of either membrane, and a globular polypeptide region that usually resides within the periplasm. Periplasmic proteins, such as SurA (PDB: 1M5Y), are globular and, as their name suggests, reside within the periplasm. Integral outer membrane (OM) proteins, such as OmpG (PDB: 2X9K), typically adopt a β-barrel architecture. Proteins that are localised extracellularly but remain associated with the cell are considered to be "secreted" proteins. Intimin, itself a β-barrel protein (PDB: 4E1S) anchored to the cell wall (via its LysM domain, PDB: 2MPW), secretes its passenger domain (PDB: 1F02 and a Phyre2 homology model of residues 455-653 modelled after the same domain from invasin PDB: 1CWV) into the extracellular milieu for host cell attachment (Leo *et al.*, 2012). Extracellular proteins that do not remain associated with the cell are instead referred to as "excreted" proteins. EspP contains a C-terminal β-barrel (EspP-C, PDB: 2QOM) domain and an N-terminal protease (EspP-N, PDB: 3SZE) region, and following an autocatalytic proteolysis event, EspP-N is excreted through EspP-C into the extracellular milieu (Leo *et al.*, 2012).

Selkrig *et al.* (2012) - in assembling β-barrels in *E. coli*.

1.2 Phospholipid biogenesis

Phospholipids are the major constituents of biological membranes. They are amphiphilic molecules comprised of two fatty acyl chains (lipophilic region) and a phosphate-containing "head group" (hydrophilic region) connected via a glycerol backbone. The fluidity of the membrane is controlled mainly by the saturation of the fatty acid "tail" groups, where saturated fatty acids increase membrane rigidity and unsaturated fatty acids increase membrane fluidity. This aspect is especially important during thermoregulation, where cooler temperatures require a more fluid membrane and conversely, warmer temperatures require a more rigid membrane (Casadei *et al.*, 2002; Morein *et al.*, 1996). For example, the saturated:unsaturated fatty acid ratios reported by Morein *et al.* (1996), were found to be 1:1.78 at 17 °C, 1:1.18 at 27 °C and 1:0.94 at 37 °C, based on the fatty acid averages using a 2:1 weighting for inner membrane (IM) and outer membrane (OM) content, respectively.

In practice, there may be hundreds of phospholipid species within a bacterial membrane (Dowhan, 1997), but the most common *E. coli* phospholipids are phosphatidylethanolamine (PE), phosphatidylglycerol (PG) and cardiolipin (CL) harbouring two of the three fatty acids: palmitic acid (C16:0), palmitoleic acid (C16:1) or vaccenic acid (C18:1) (Lugtenberg & Peters, 1976; Morein *et al.*, 1996) (Figure 1.2.1). The diversity among phospholipids, at least in *Gammaproteobacteria*, is in part due to their ability to incorporate exogenous fatty acids during phospholipid biosynthesis, which are imported by the FadL-

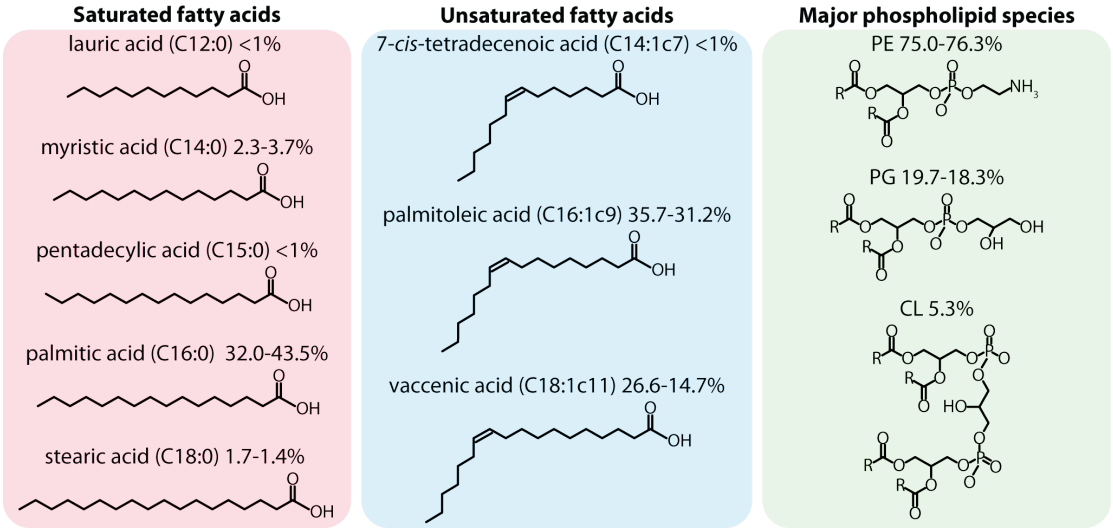


Figure 1.2.1 Composition of phospholipids.

The chemical structure of the most common saturated (red) and unsaturated (blue) fatty acids and phospholipid head groups (green) that comprise *E. coli* phospholipids are shown. Relative amounts given are taken from (Morein *et al.*, 1996).

FadD pathway (Parsons & Rock, 2013; Rock & Cronan, 1996) (Figure 1.2.2). However, the bulk of the phospholipid fatty acid component is comprised of endogenous fatty acids, which are synthesised via the type 2 fatty acid biosynthesis (FAS2) pathway (Parsons & Rock, 2013).

The FAS2 pathway makes use of an acyl carrier protein (ACP) (Figure 1.2.3) that chaperones a growing fatty acyl chain between successive enzymes in a series of elongation steps. With each cycle of elongation, two carbons are added to the acyl chain of the acyl-ACP intermediate. When the chain-length reaches ten carbons, the acyl chain may remain saturated or desaturate into an ω-7 fatty acid (Parsons & Rock, 2013; Yao & Rock, 2013). In either case, chain elongation will continue and with each successive iteration, acyl-ACPs become poorer substrates for FAS2 enzymes and better substrates for phospholipid synthesis enzymes (Parsons & Rock, 2013; Yao & Rock, 2013).

Phospholipid biogenesis begins with the synthesis of phosphatidic acid (PA), the simplest of, and precursor to, all phospholipid species (Yao & Rock, 2013). *E. coli* contains two pathways involved in PA synthesis, the two-step PlsB pathway, which is essential, and the three-step PlsX/Y pathway, which can only incorporate endogenous fatty acids into PA (Yao & Rock, 2013; Zhang & Rock, 2008) (Figure 1.2.4). Following synthesis, PA is converted into a nucleolipid that acts as the major branch point for phospholipid differentiation into negatively charged phospholipids, PG and CL, or zwitterionic phospholipid species, including PE (Figure 1.2.5). Although it is clear that phospholipids are constituents of the IM and the periplasmic leaflet of the OM, the mechanisms by which trans- and inter-

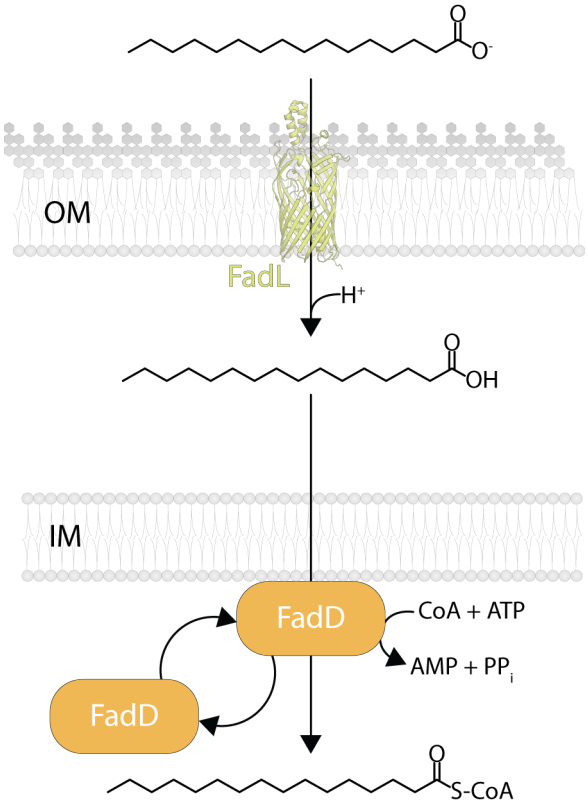


Figure 1.2.2 Exogenous fatty acid import of long chain fatty acids.

An illustration of exogenous fatty acid import is shown using palmitate as the example. Chemical structures of each fatty acid species and a ribbon diagram of FadL (PDB: 1T16) are shown. Palmitate is imported into the periplasm through FadL and subsequently protonated in the H^+-rich periplasm to form palmitic acid. Although the mechanisms by which fatty acids enter the inner membrane and flip to the cytoplasmic leaflet are unknown, passive diffusion is thought to play the major role (Kamp & Hamilton, 2006). FadD responds to the presence of an appropriate substrate (i.e. palmitic acid) by peripherally associating with the inner membrane. Following the ATP-dependent activation of its substrate by the attachment of coenzyme-A (CoA), FadD dissociates from the membrane (Yao & Rock, 2013).

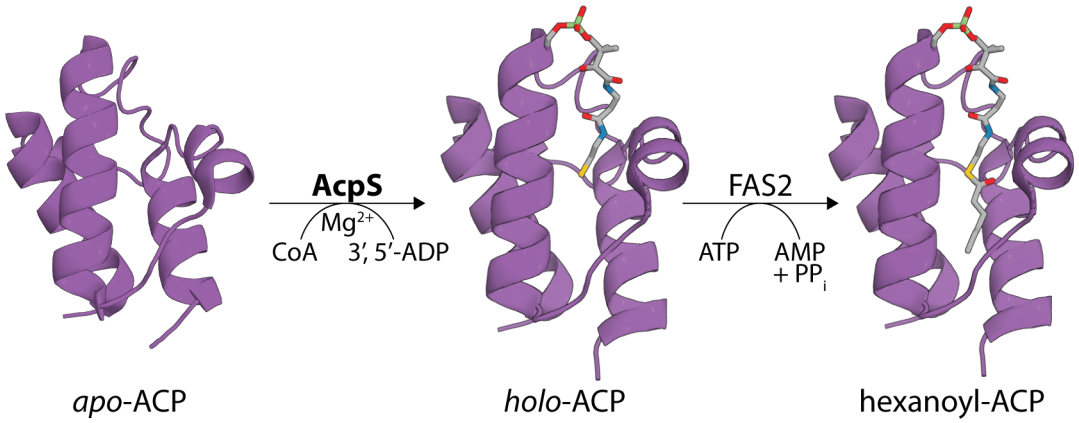


Figure 1.2.3 Acyl carrier protein maturation.

apo-ACP is converted into *holo*-ACP via the covalent attachment of a 4'-phosphopantetheine from CoA onto a conserved serine residue (Roujeinikova *et al.*, 2007). During the elongation steps of the FAS2 pathway, acetates are sequentially added to the thiol residue of the 4'-phosphopantetheine. In this example, a six-carbon acyl group is currently attached, but will undergo successive rounds of elongation. Ribbon diagrams of the *apo*-ACP (PDB: 1T8K) and *holo*-ACP (PDB: 2FAC) are shown. Ball-and-stick representations of the atoms comprising the conserved serine-36 (prosthetic group attachment site) are coloured as follows: grey carbons, red oxygens, blue nitrogens, yellow sulphur, green phosphor, and hydrogens are omitted.

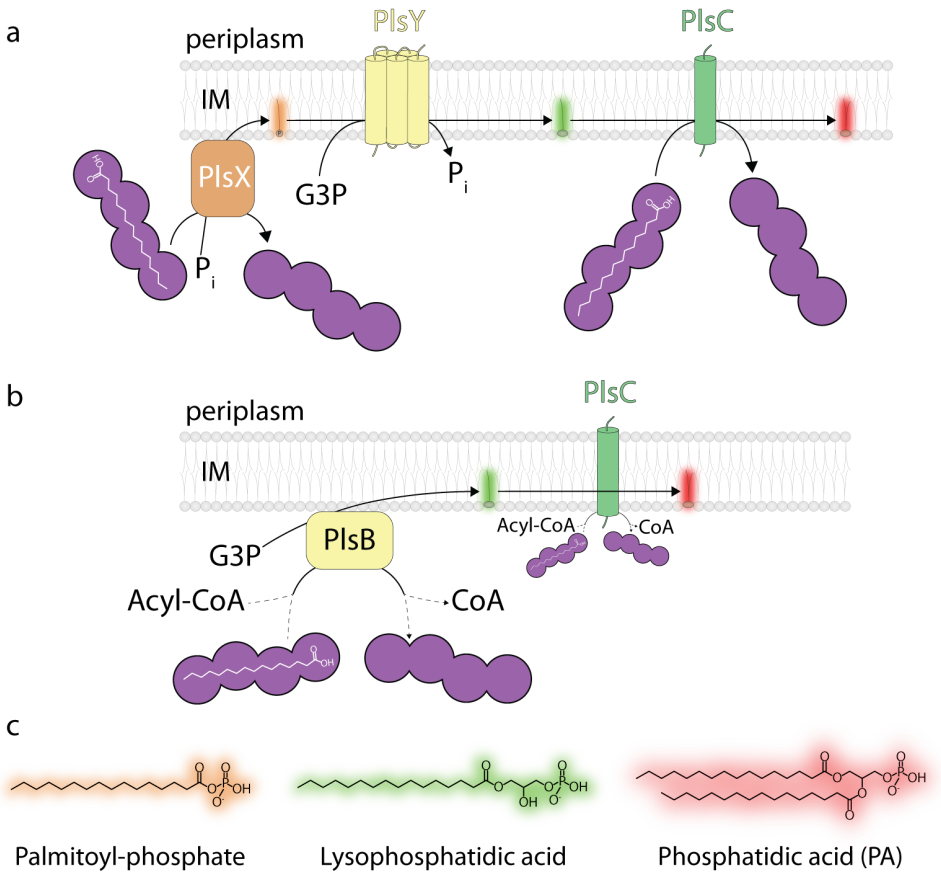


Figure 1.2.4 Phosphatidic acid synthesis.

PA synthesis may follow one of two pathways (Parsons & Rock, 2013). In both examples, palmitic acids are used to demonstrate incorporation of fatty acids moieties and enzymes that form a similar product are coloured the similarly. **a,** Many prokaryotes synthesise phosphatidic acid via a three-step mechanism that can only incorporate endogenous fatty acids. **b,** *Gammaproteobacteria* typically utilise a two-step mechanism for synthesising phosphatidic acids. The dashed lines indicate that either exogenous or endogenous fatty acids may be incorporated. **c,** The chemical structures of phosphatidic acid and its intermediates are shown, where their coloured background corresponds to the species shown in "**a**" and "**b**". G3P represents glycerol-3-phosphate and P_i represents inorganic phosphate

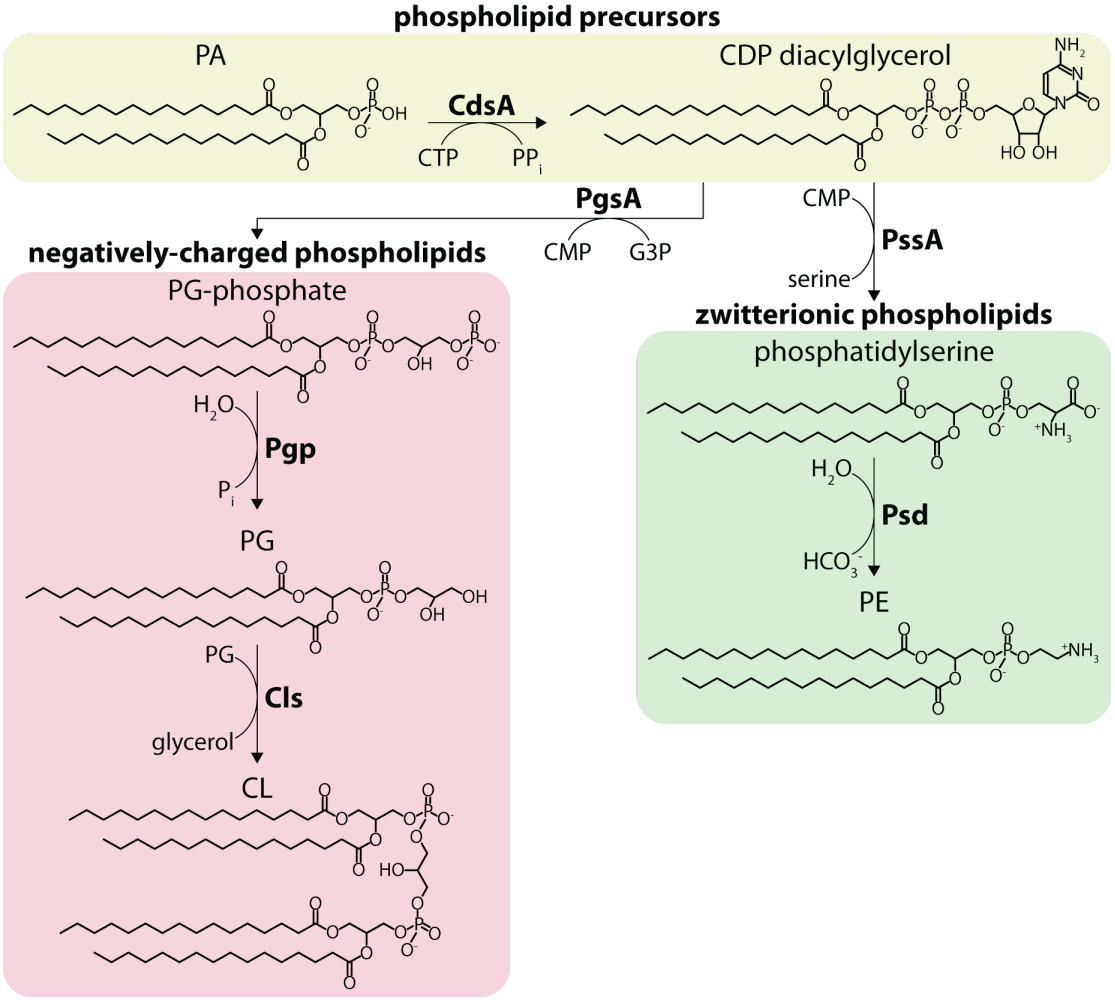


Figure 1.2.5 Phospholipid differentiation from phosphatidic acid.

Synthesis of the major and minor *E. coli* phospholipid species occurs at the cytoplasmic leaflet of the inner membrane. Precursor phospholipids (yellow) may differentiate either into zwitterionic (green) or negatively-charged (red) phospholipids as indicated (Parsons & Rock, 2013). The enzymes that catalyse a particular reaction are shown in bold typeface and all fatty acid moieties depicted are derived from palmitic acid. G3P represents glycerol-3-phosphate, P_i represents inorganic phosphate, and CMP and CTP represent cytidine mono- and tri-phosphate, respectively.

bilayer phospholipid movement occurs are largely unknown (Figure 1.2.6). Considering this movement is rapid in natural membranes (Donohue-Rolfe & Schaechter, 1980; Huijbregts *et al.*, 1998; Kubelt *et al.*, 2002) and significantly slower in synthetic, model membranes (Kornberg & McConnell, 1971), a protein component is potentially involved. Indeed, MsbA which catalyses the ATP-dependent 'flopping' of lipid A during LPS biosynthesis (see Section 1.3), is thought to participate in *E. coli* phospholipid flopping (Doerrler *et al.*, 2004; Doerrler *et al.*, 2001; Eckford & Sharom, 2010) (Figure 1.2.6a). However, whether MsbA actually plays a role in the transbilayer movement of phospholipid species has been met with several criticisms (Tefsen *et al.*, 2005), especially since phospholipid flip-flop apparently does not require ATP (Donohue-Rolfe & Schaechter, 1980; Huijbregts *et al.*, 1998). Indeed, (Kol *et al.*, 2001; Kol *et al.*, 2003a; Kol *et al.*, 2003b) demonstrated that - rather than MsbA specifically - classical transmembrane α-helices may act like a scaffold to support phospholipid flip-flop (Figure 1.2.6a-b), which also satisfies the proteinacious component requirement of transbilayer movement.

An integral inner membrane protein, PbgA, was shown to be involved in anterograde shuttling of CL in *Salmonella enterica* subsp. *enterica* serovar Typhimurium (Dalebroux *et al.*, 2015). PbgA forms a tunnel for CL transfer from the IM to the OM, but how PbgA interacts with the OM is unknown. In *E. coli*, an OM lipoprotein, YraP, was recently identified and is thought to contribute to PG antereograde trafficking (Maderbocus, 2012). If the *E. coli* PbgA homologue, YejM, functions similarly to PbgA it may be that PbgA and YraP form an intermembrane-spanning complex (Figure 1.2.6c) that together facilitate phospholipid

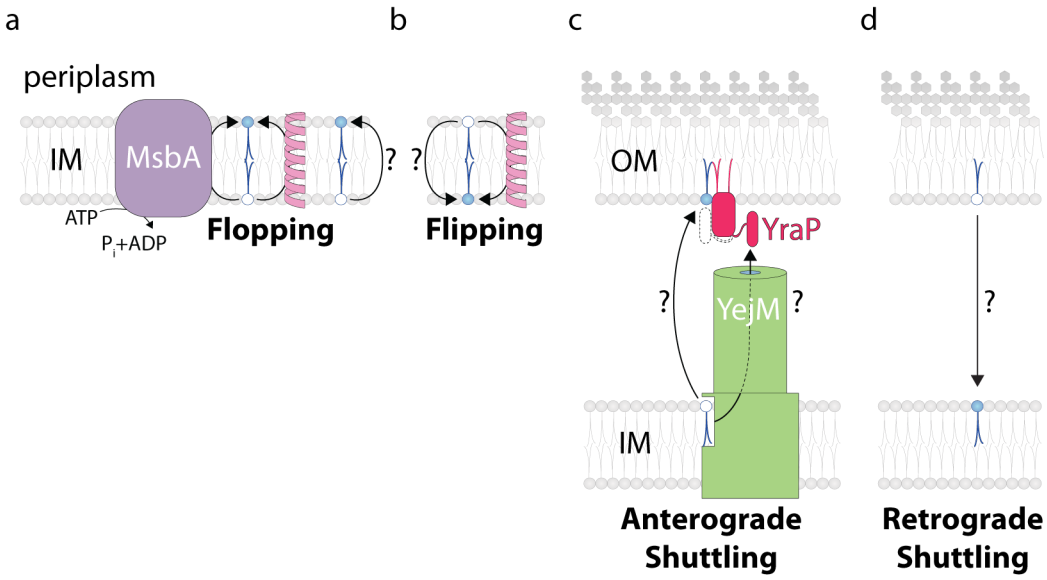


Figure 1.2.6 Phospholipid transbilayer and interbilayer movement.

In *E. coli*, phospholipids constitute three of the four membrane leaflets, but are synthesised at the cytoplasmic leaflet of the inner membrane. The mechanisms by which phospholipids transfer from one leaflet to another are unknown (as indicated by question marks), but several hypotheses have been proposed and are depicted here. The direction and type of movement is indicated in bold typeface, and the phospholipid's original position (blue outline) and new position (blue outline and fill) are shown. **a** and **b,** phospholipid transbilayer movement may be catalysed by MsbA (**a** only) or may spontaneously occur with or without the assistance of transmembrane α-helices (shown in pink) that act as a scaffold. **c** and **d,** no mechanisms for interbilayer phospholipid trafficking in *E. coli* are known, although a potential and putative YejM-YraP complex may somehow be involved (**c** only).

anterograde trafficking.

To maintain an asymmetrical OM lipid bilayer, *E. coli* employ diverse mechanisms to remove phospholipids from the outer leaflet of the OM. During OM stress, such as on exposure to EDTA or antimicrobial peptides, *E. coli* may shed LPS molecules. To restore the bilayer, phospholipids flop from the inner leaflet as a "quick fix" until newly synthesised LPS may be inserted (Nikaido, 2003). In response to the OM stress, LPS molecules are modified by the addition or removal of fatty acids or sugar moieties to improve the quality of the OM and prevent further LPS shedding (Needham & Trent, 2013; Wang & Quinn, 2010). OmpLA (also commonly named PldA) or the maintenance of OM lipid asymmetry (Mla) proteins then destroy or recycle the outer leaflet phospholipids and lysophospholipids, respectively, to ensure OM asymmetry is maintained (Chong *et al.*, 2015; Dekker, 2000; Malinverni & Silhavy, 2009) (Figure 1.2.7).

Apart from being the main constituent of the biological membranes, phospholipids act as substrates in a range of other biosynthetic pathways. To maintain osmotic homeostasis, the "head" group from PG is donated to membrane-derived oligosaccharides (Goldberg *et al.*, 1981) and the remaining diacylglycerol by-product may be recycled by DgkA (Zhang & Rock, 2008). Three fatty acid moieties, two from PG and one from PE, are donated during lipoprotein biosynthesis, generating a free "head" group and a 2-acyl-glycero-phosphoethanolamine (see Section 1.5). In *E. coli*, this lysophospholipid may be degraded into a free fatty acid by PldB or acylated by Aas to reconstitute PE (Zhang & Rock, 2008). Additionally, the LPS modifications used to alter the OM

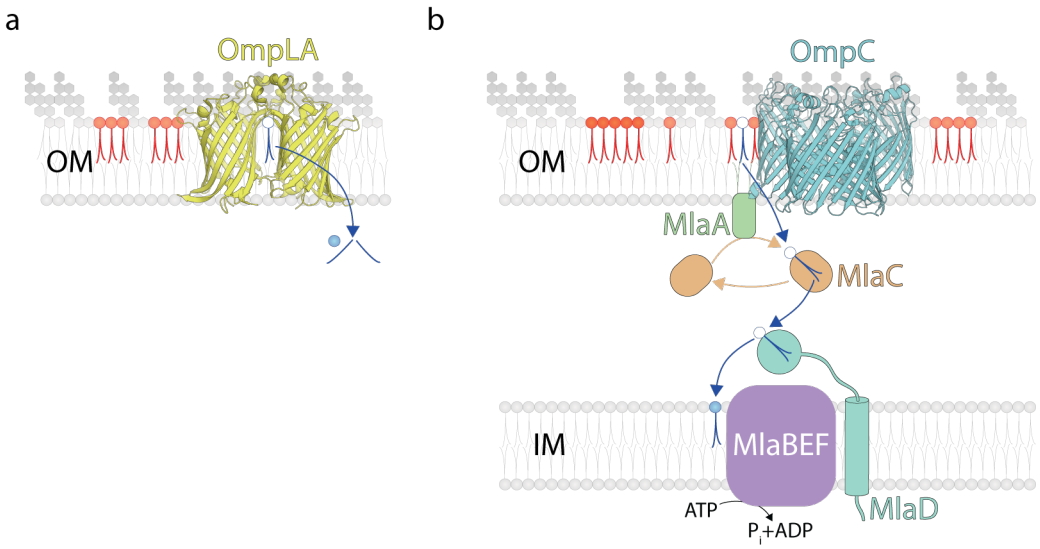


Figure 1.2.7 Maintenance of outer membrane asymmetry.

Phospholipids found in the LPS layer (red) destabilise the outer membrane and are subsequently removed. The initial and intermediate locations (blue outline) and final position (blue outline and fill) of the LPS layer phospholipid being transferred are also indicated. **a,** OmpLA normally exists as a monomer, but when phospholipids enter the LPS layer, OmpLA activates into a dimer (PDB: 1QD6) and removes the fatty acid "tail" groups (Dekker, 2000). The phospholipid "head" and "tail" groups may then dissociate back into the outer membrane and cytoplasm, and are probably recycled as though they were exogenous material. For clarity, OmpLA generation of LPS layer lysophospholipids, which are thought to have an alternate but unknown recycling pathway, are not shown. **b,** The Mla pathway recycles LPS layer phospholipids by retrograde trafficking to the inner membrane (Malinverni & Silhavy, 2009). MlaA and OmpC (PDB: 2J1N) interact in the outer membrane, and it is thought that this interaction facilitates transfer of LPS layer phospholipid to MlaC (Chong *et al.*, 2015). Although the leaflet where the phospholipid is ultimately transferred to is unknown, the periplasmic leaflet is used in this example.

quality typically require phospholipid substrates (Needham & Trent, 2013; Wang & Quinn, 2010).

1.3 Lipopolysaccharide biogenesis

LPS is the main component of the outer leaflet of the bacterial OM and one of the major hallmarks of Gram-negative bacteria. All LPS contain a similar amphiphilic architecture to phospholipids, with a lipophilic "lipid A" component and hydrophilic "core oligosaccharide" and "O-antigen" components (Figure 1.3.1). Unlike the phospholipid leaflet, the outer leaflet of the OM is incredibly rigid, due to the dense packing of the acyl groups (typically six) attached to lipid A, as well as the electrostatic interactions between divalent cations (e.g. Ca^{2+} and Mg^{2+}) and the negatively charged groups in lipid A and core oligosaccharides (e.g. phosphates) (Wu *et al.*, 2013). It acts as a structural barrier protecting against harmful substances, including hydrophobic antibiotics, detergents and short chain free fatty acids (Desbois & Smith, 2010; Zhang *et al.*, 2013). Additionally, LPS is considered to be an endotoxin based on the toxicity of the lipid A component to the host (Raetz & Whitfield, 2002).

The core oligosaccharide can be divided into two regions: the inner core and the outer core. The inner core invariably contains 3-deoxy-D-*manno*-oct-2-ulosonic acid residues (Kdo, keto-deoxyoctulosonate) and also often contains L-glycero-D-manno-heptose (Hep), whereas the outer core is more diverse and is usually comprised of neutral or basic pyranose sugars (Frirdich & Whitfield, 2005). Five *E. coli* outer core architectures are known, R1 to R4 and K-12 (Jansson *et al.*, 1981) (Figure 1.3.2), and a sixth mixed R1R4 architecture has also been reported,

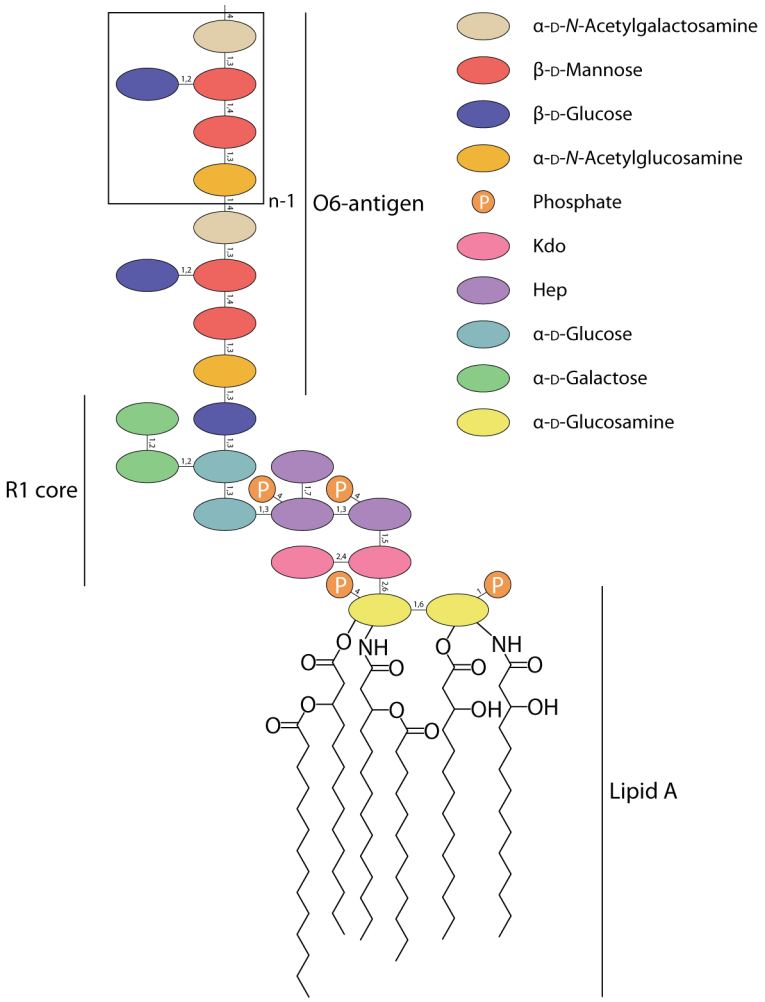


Figure 1.3.1 Structural example of *E. coli* lipopolysaccharide.

Of the 5 core oligosaccharides and almost 170 O-antigen variations, the R1 core and O6 antigen are highly prevalent within pathogenic *E. coli* (Appelmelk *et al.*, 1994; Gibbs *et al.*, 2004; Stenutz *et al.*, 2006). The number of O-antigen repeating units ('n') varies between 10 and 25 (Stenutz *et al.*, 2006) and the carbons involved in glycosidic or phosphoester bonding are numbered as shown. Cartoon depictions of the various subunits may be identified using the legend to the right, where Kdo represents 3-deoxy-D-*manno*-oct-2-ulosonic acid and Hep designates L-glycero-D-manno-heptose.

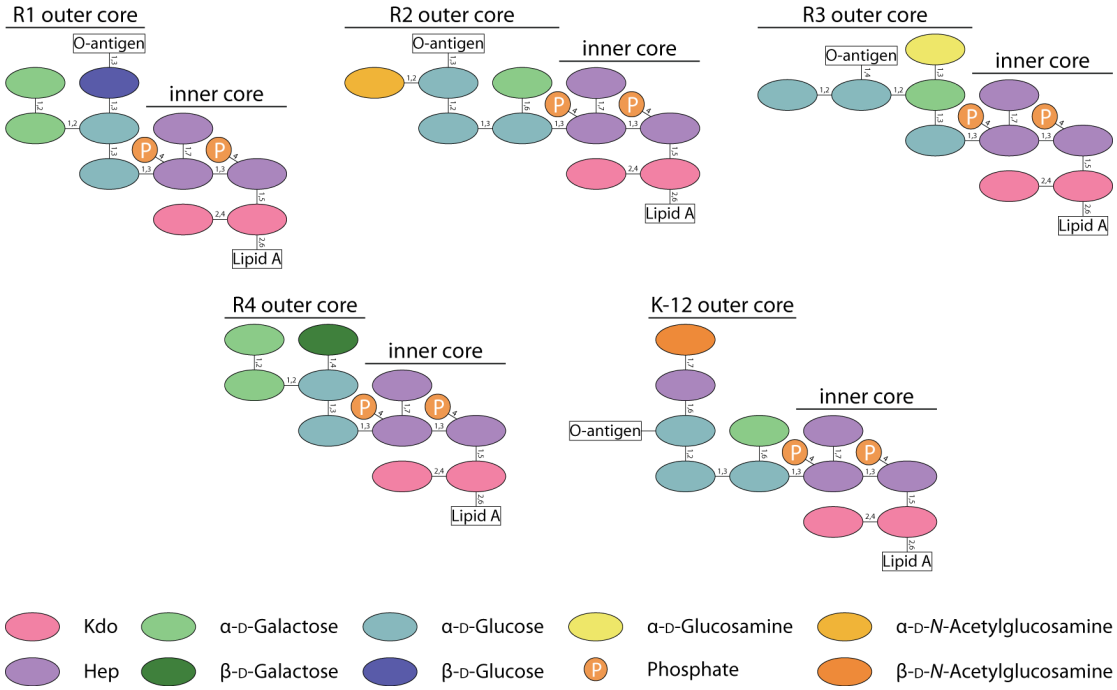


Figure 1.3.2 The five *E. coli* core oligosaccharides.

The inner and outer core are indicated, and the lipid A and known O-antigen attachment sites are also shown. The identity of each subunit is indicated in the legend below the core structures and carbons involved in glycosidic or phosphoester bonds are numbered as shown.

possibly arising from co-expression within strains harbouring both R1 and R4 core synthesis genes (Amor *et al.*, 2000; Appelmelk *et al.*, 1994).

Bacteria with an intact LPS are said to have a 'smooth' LPS phenotype, distinct from the 'rough' phenotype (missing *O*-antigen) and the 'deep rough' phenotype (missing both outer core and *O*-antigen). The term "rough" is used to indicate the increasing hydrophobicity and permeability conferred on the OM (Frirdich & Whitfield, 2005). The lipid A component is the only "essential" component of the LPS layer, where lipid A mutants are normally inviable (Raetz & Whitfield, 2002). Among very few interesting exceptions are two species of colistin-resistant *Acinetobacter*: *A. baumannii* (Henry *et al.*, 2012) and *A. nosocomialis* (Vila-Farrés *et al.*, 2015). These bacteria do not produce LPS, due to a defect in their lipid A biosynthesis pathway. This presumably causes the bacteria to produce an OM bilayer comprised entirely of phospholipid, where the upregulation of a range of cell envelope and membrane biogenesis machineries enables the bacteria to remain viable and resistant to the LPS-targeting colistin antibiotic.

Several strains within the *Neisseria* and *Haemophilus* species - particularly the human pathogens - naturally lack an O-antigen, and produce lipooligosaccharide (LOS, also known as "rough" LPS) instead of traditional LPS (Harvey *et al.*, 2001). However, these bacterial strains modify their LOS by attaching sialic acid residues to their core oligosaccharide moiety, which functions to mimic host-cell glycolipids and glycosphingolipids. As a result of the similar surface epitope, these strains are able to evade detection by the host's immune defence (Harvey *et al.*, 2001). Most laboratory strains of *E. coli*, especially *E. coli* K-12 and B strains,

typically generate LOS instead of LPS (Chart *et al.*, 2000) (Figure 1.3.3), but they do not modify their LOS to mimic host cells. Their LOS instead likely arose as an evolutionary adaptation resulting from the numerous rounds of *in vitro* culturing and genetic modification, such that these *E. coli* laboratory strains are rapidly cleared by the host immune system (Browning *et al.*, 2013; Chart *et al.*, 2000).

E. coli K-12 strains produce the eponymous K-12 core oligosaccharide (Figure 1.3.3a) and due to an *IS*5 insertion in *wbbL*, cannot synthesise their O16 antigen (Browning *et al.*, 2013; Liu & Reeves, 1994). *E. coli* B strains - such as BL21 and REL606 (and their derivatives) - produce a variant of the R1 outer core that lacks the two α-D-galactose residues (Figure 1.3.3b, cf. R1 core from Figure 1.3.2). The resulting 3-hexose ring outer core is formed because of an *IS*1 element disrupting the gene whose product is responsible for incorporating the α-D-galactose residues, *waaT* (Jeong *et al.*, 2009).

Jansson *et al.* (1981) and Jeong *et al.* (2009) reported that the *waaT* inactivation produces a 2-hexose ring outer core that also lacks the β-D-glucose residue an O-antigen would otherwise attach to. This would certainly be true if *waaV*, which is involved in the attachment of that β-D-glucose residue (Heinrichs *et al.*, 1998), was also inactivated; however, the sequence of *waaV* indicates that WaaV should be functional and polarity effects can be ruled because *waaV* is not located within the operon *waaT* is a member of (Caspi *et al.*, 2014). Instead, Prehm *et al.* (1975) demonstrated that *E. coli* BB produce a 2-hexose ring outer core and perhaps because *E. coli* B strains REL606, BB and BL21 share a common ancestor (Daegelen *et al.*, 2009), later groups may have inferred that *E. coli* B strains BL21

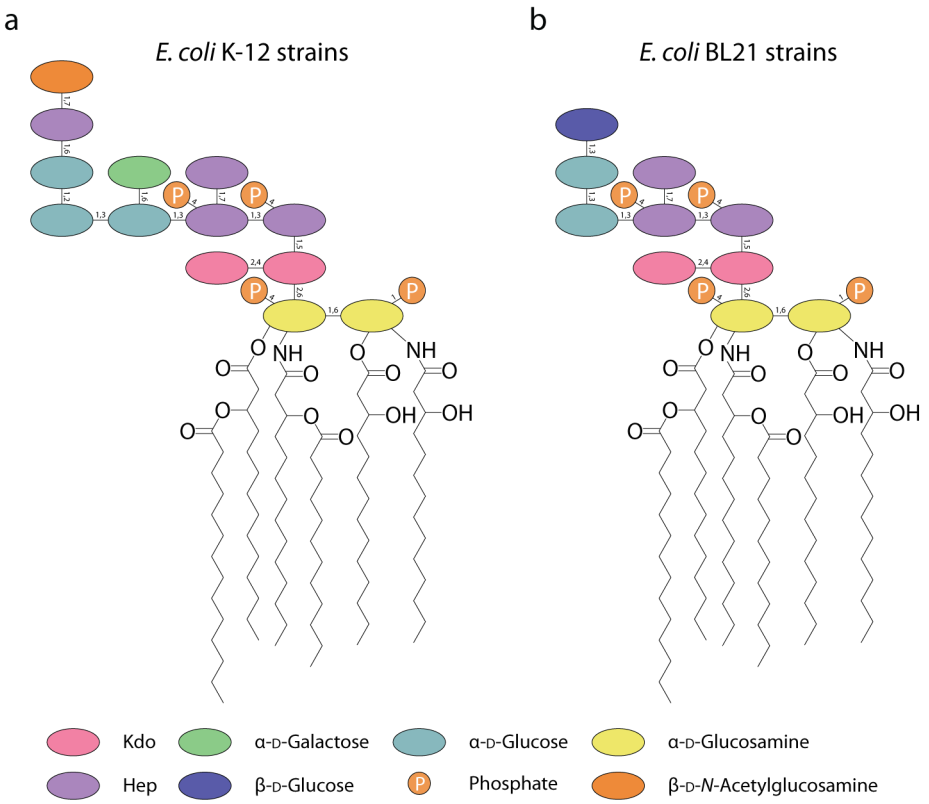


Figure 1.3.3 Lipooligosaccharide of common *E. coli* laboratory strains.

While laboratory strain *E. coli* may harbour O-antigen biosynthesis genes, at least one is non-functional and therefore these strains produce a rough LPS (LOS) variant instead. Depicted here are the LOS from *E. coli* K-12 or BL21 parent strains, where each subunit is indicated in the legend and carbons are numbered if they are involved in glycosidic or phosphoester bonding.

and REL606 produce the 2-hexose ring outer core. The reason *E. coli* B strains REL606 and BL21 do not produce O-antigen is instead due to an *IS*1 insertion that disrupts *wbbD* (Jeong *et al.*, 2009), whose product is involved in the biosynthesis of the O7-antigen repeating unit (Riley *et al.*, 2005).

Lipid A is comprised of two phosphorylated and acylated glucosamine sugars connected via a characteristic β-1,6-glycosidic bond (Raetz & Whitfield, 2002). Lipid A variability between bacterial genera is due to the length of their β-hydroxy fatty acids, although each bacterial species may modify their lipid A in response to environmental pressures by removing, adding, swapping or altering fatty acid components (Needham & Trent, 2013; Raetz *et al.*, 2007). Lipid A, along with two Kdo residues from the inner core, are synthesised via the Raetz pathway, and the remaining core oligosaccharide subunits are subsequently added by a series of glycosyltransferases to produce a LOS molecule. (Raetz & Whitfield, 2002). LOS synthesis occurs at the cytoplasmic leaflet of the IM, and once complete, MsbA flops the LOS to the periplasmic leaflet where WaaL can transfer the O-antigen to it, producing LPS (Raetz & Whitfield, 2002) (Figure 1.3.4).

There are about 170 known *E. coli* O-antigens, resulting in a diverse range of *E. coli* serotypes (Stenutz *et al.*, 2006). Like LOS synthesis, O-antigens are synthesised at the periplasmic interface of the IM, but may proceed via one of three distinct pathways: the Wzy-dependent pathway, the ATP-dependent pathway or synthase-dependent pathway (Raetz & Whitfield, 2002). The majority of *E. coli* O-antigens are assembled via the Wzy-dependent pathway (Figure 1.3.4a), although up to eight *E. coli* O-antigens are synthesised via the ATP-

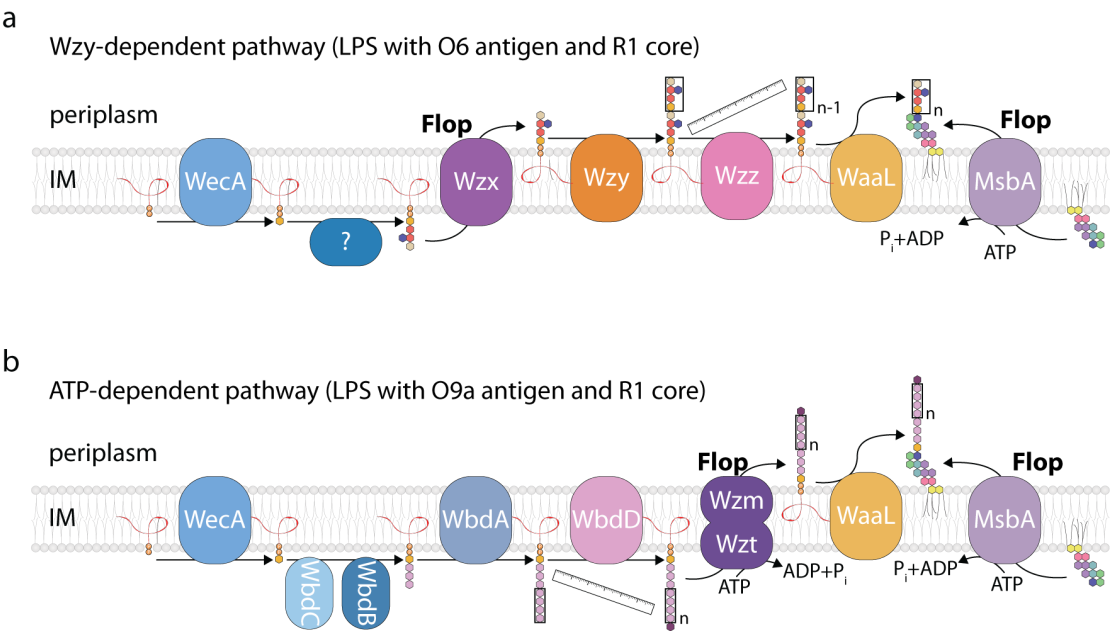


Figure 1.3.4 O-antigen biosynthesis pathways in *E. coli*.

There are three known O-antigen biosynthesis pathways, two are depicted here: **a,** the Wzy-dependent pathway and **b,** the ATP-dependent pathway. The third, synthase-dependent pathway has not been identified in *E. coli* and is therefore not depicted here. In each pathway, WecA initiates O-antigen synthesis by transferring an *N*-acetylhexosamine sugar onto the universal glycan lipid carrier, undecaprenyl diphosphate (Stenutz *et al.*, 2006). Additional sugar subunits are added by glycosyltransferases (shown in blue, the question mark indicates that they have not been identified yet). Floppases are coloured purple and specialised glycosyltransferases that transfer O-antigen repeating subunits or the entire O-antigen are coloured orange. The final number of O-antigen repeats, represented by "n", is determined by enzymes coloured pink (a ruler is also used to highlight this step), and in the case of WbdD, a terminal sugar moiety is added to prevent the incorporation of additional O-antigen. The final biosynthesis step involves the transfer of the entire O-antigen by WaaL to LOS to produce LPS, which is subsequently transferred to the Lpt machinery for OM translocation.

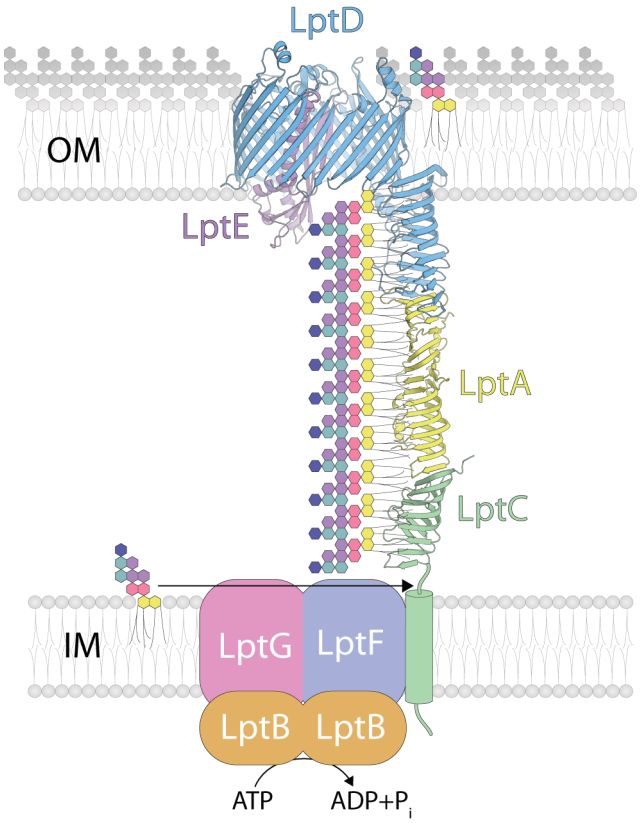


Figure 1.3.5 PEZ model for lipopolysaccharide transfer to the outer membrane.

Rather than LPS biogenesis, outer membrane biogenesis of the simpler LOS is shown for clarity. The LOS depicted is characteristic of the *E. coli* BL21 and its derivative strains used throughout this thesis. LtpA (PDB: 2R19) and the periplasmic domains of both LptC (PDB: 3MY2) and LptD each adopt a β-jellyroll architecture. Although an LptA dimer is shown, about 40 subunits of LptA bridge LptC and the LptDE (PDB: 4Q35) complex, allowing LOS to slide from the IM to the OM in an ATP-dependent manner (Okuda *et al.*, 2016; Whitfield & Trent, 2014).

dependent pathway (Raetz & Whitfield, 2002; Stenutz *et al.*, 2006) (Figure 1.3.4b). There are no known *E. coli* O-antigens that follow the synthase-dependent pathway and, to-date, only the O:54 antigen of *S. enterica* subsp. *enterica* serovar Borreze is known to assemble this way (Keenleyside *et al.*, 1994).

Among the final steps in LPS biogenesis is transfer to the LPS layer in the OM from the cytoplasmic leaflet of the IM (Whitfield & Trent, 2014), potentially followed by their modification in response to extracellular stimuli (Needham & Trent, 2013). LPS (or LOS) transfer to the OM is performed by the ATP-dependent Lpt (lipopolysaccharide transport) pathway (Whitfield & Trent, 2014). Okuda *et al.* (2016) suggested the 'PEZ model' for LPS transfer to the OM (Figure 1.3.5), where LPS transfer from the IM to the OM is thought to occur along a periplasmic β-jellyroll bridge formed by LptC, about 40 LptA subunits and LptD. As LPS molecules are added to the bridge, they are pushed toward the OM, much like PEZ candy are pushed toward the dispenser's opening. LptD contains a lateral gate that enables the LPS molecules to enter both the bilayer and the LptD lumen simultaneously (Li *et al.*, 2015b). This ensures the hydrophobic portion of LPS enters the bilayer and the hydrophilic portion enters the LptD lumen for secretion through an exit pore into the extracellular milieu.

1.4 Peptidoglycan biogenesis

The cell wall is responsible for maintaining the shape of the prokaryotic cell and consists of peptidoglycan, also commonly known as the murein sacculus ("murein" and "sacculus" are derived from the Latin words for wall and bag, respectively). In Gram-negative bacteria, the peptidoglycan is found within the

periplasmic space (Figure 1.4.1), where it envelops the cell as a continuous, mesh-like macromolecule (Nguyen *et al.*, 2015; Vollmer & Seligman, 2010). Because of its structural role, peptidoglycan has often been inadvertently characterised as a "rigid" (inflexible) structure, especially in an historical context (Braun *et al.*, 1974; Glauert & Thornley, 1969; Weidel *et al.*, 1960) or in undergraduate textbooks (Barton, 2005, p. 94; Kutter & Goldman, 2015, p. 867). Although peptidoglycan does have a degree of rigidity enabling it to play that structural role, it also has a degree of flexibility that enables it to contract or expand to protect the cell from internal turgor pressure due to osmotic shock or mechanical stress (Turner *et al.*, 2013).

The term "peptidoglycan" is derived from its chemical composition, consisting of a polysaccharide (glycan) component and a series of peptides that are derived from several unusual amino acids. The entire peptidoglycan lattice can be considered one of nature's largest and toughest macromolecules, due to its extensive covalent bonded structure (Koch & Woeste, 1992). Despite the complex three-dimensional architecture (Nguyen *et al.*, 2015; Vollmer & Seligman, 2010), peptidoglycan is comprised of a relatively uniform series of subunits, known as muropeptides. Each newly synthesised muropeptide is comprised of a β-*N*-acetylglucosamine (NAG), a β-*N*-acetylmuramic acid (NAM) and a pentapeptide connected to the lactyl group of NAM (Figure 1.4.2a).

In *E. coli*, the newly synthesised pentapeptide contains four amino acids: the L- and D-enantiomers of alanine, D-glutamic acid (where peptide linkages involve the β-amine and ε-carboxylic acid), and *meso*-diaminopimelic acid (DAP), which uses its

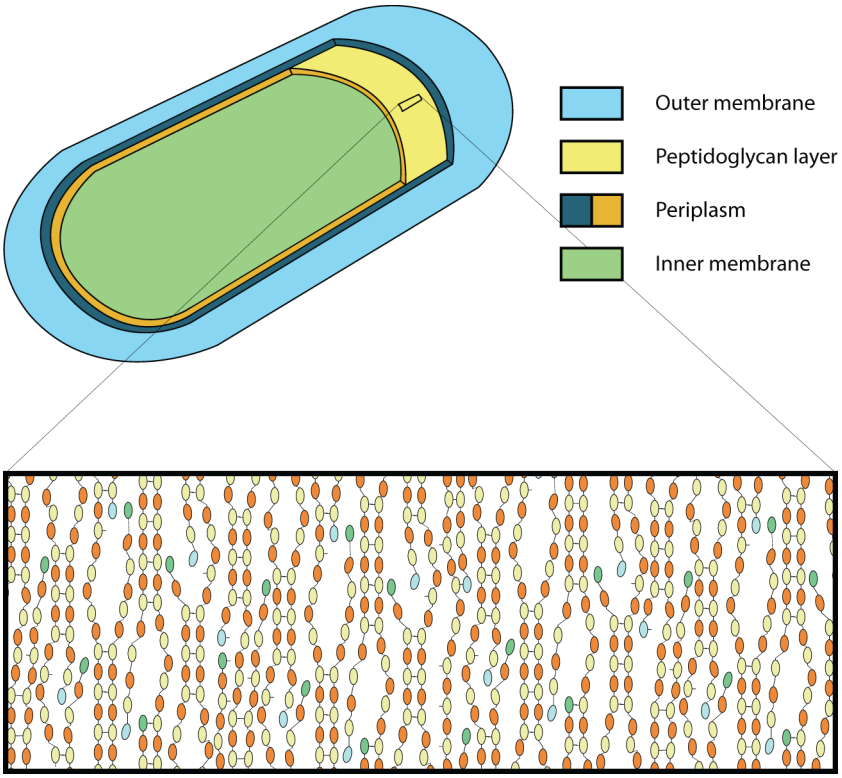


Figure 1.4.1 Peptidoglycan architecture.

A cartoon representation of an *E. coli* cell is shown with a portion of the cell envelope removed. The circumferential model for the arrangement of the layer is shown, where the glycan strands (coloured ovals) wrap circumferentially around the cell and the peptide crosslinks run parallel to the long axis of the cell (black lines running parallel to the text) (Nguyen *et al.*, 2015; Vollmer & Seligman, 2010).

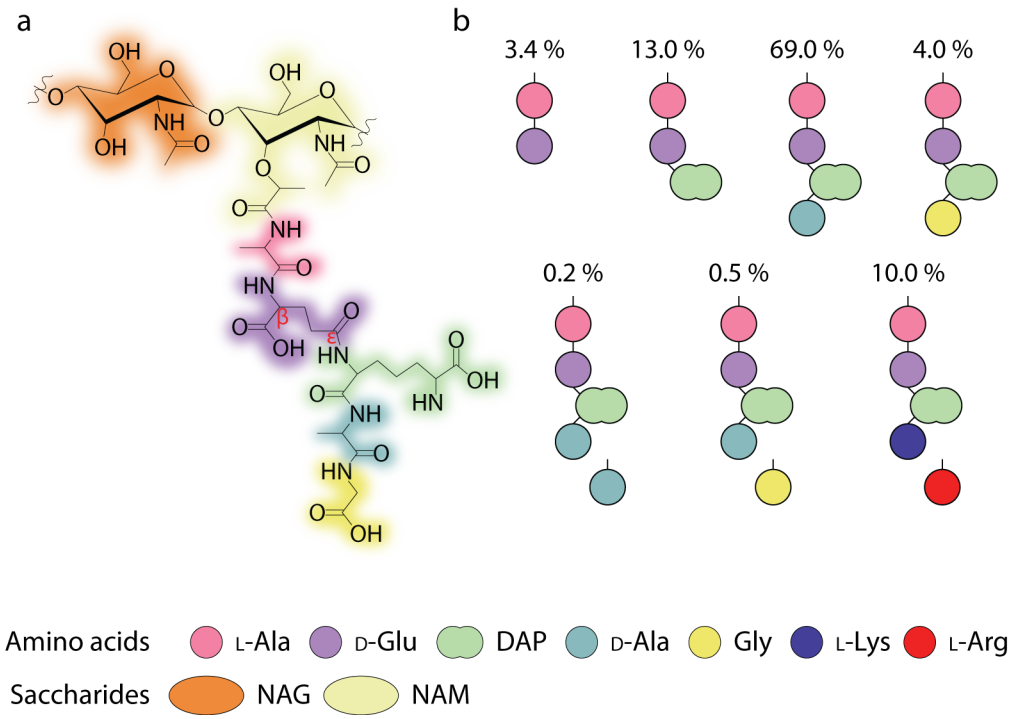


Figure 1.4.2 Muropeptide subunits.

a, Chemical structure of a newly synthesised muropeptide subunit. The β- and ε-carbons of D-Glu that are positioned within the plane of the peptide bonds are indicated in red. **b,** Cartoon depictions of the various monomeric muropeptides (glycan subunits are omitted for clarity) with their average abundance indicated as per Table 1.4.1. The peptide component The peptide component of the muropeptide is typically modified to allow for peptide crosslinkages or covalent attachment to Braun's lipoprotein.

L-centre for peptide bonding in the pentapeptide chain (Vollmer & Bertsche, 2008) (Figure 1.4.2a). Interestingly, although the L-Ala-D-Glu-DAP-D-Ala-D-Ala pentapeptide species is present in all muropeptide initially, less than 1% of muropeptides within peptidoglycan retain it (Table 1.4.1). This is because it will rapidly be degraded or modified on addition to the peptidoglycan network into one of six other muropeptides (Figure 1.4.2b, Table 1.4.1).

Each muropeptide forms up to two interactions with neighbouring muropeptides, such that peptidoglycan is considered a single macromolecule: β-1,4-glycosidic bonds between neighbouring glycan subunits and peptide crosslinks. Muropeptide chains consist of an average of about 33 disaccharide subunits (Table 1.4.2), where the terminal pair contains NAG and 1,6-anhydro-NAM subunits (i.e. NAM with an intramolecular ether linkage between carbons 1 and 6) (Figure 1.4.3). About 45.4 % of muropeptides (Table 1.4.2) then form peptide crosslinkages with neighbouring strands to ensure a continuous series of chains can wrap around the cell. The majority (86-88 %, Tables 1.4.2-1.4.3) of these interactions consist of a D,D-crosslink between a D-Ala "donor" and the D-centre of a *meso*-DAP "acceptor" (Figure 1.4.4a), and the remaining 12-14 % (Tables 1.4.2-1.4.3) form L,D-crosslinks that occur between the L-centre "donor" and D-centre "acceptor" of adjacent DAPs (Figure 1.4.4b).

Considering the diversity of the monomeric peptides, it is not surprising that there are a variety of dimeric muropeptides (Figure 1.4.4) produced during the lifespan of *E. coli* (Table 1.4.3). Indeed, despite the simplicity of the muropeptide subunit, there are up to 50 distinct muropeptide structures that may be produced. These

Table 1.4.1 The distribution of monomeric peptides depends on the strain and growth conditions.

The distribution of monomeric peptides are indicated. Only references where the author(s) determined the abundance of all known monomeric peptides were included and used to determine monomeric peptide averages. Muropeptides may contain: dipeptides (2), tripeptides with (3-Lpp) or without (3) covalently bound Lpp, tetrapeptides with (4g) or without (4) a glycine at position 4 and pentapeptides with (5g) or without (5) a glycine at position 5.

	Growth conditions[a]			Distribution of monomeric peptides[b]							
***E. coli* strain**	media	°C	phase	2	3	3-Lpp	4	4g	5	5g	**References**
W7	PB	30	Exp	4.86	14.85	6.18	68.84	4.51	0.15	0.61	Glauner (1988)
KN 126	PB	30	Exp	4.96	16.15	6.85	68.31	3.19	0.11	0.44	Glauner (1988)
KN 126	PB	37	Exp	4.01	14.63	8.81	68.71	3.22	0.13	0.49	Glauner *et al.* (1988)
KN 126	PB	42	Exp	4.59	13.98	7.45	70.05	3.27	0.20	0.46	Glauner *et al.* (1988)
KN 126	LB	30	Exp	2.42	13.84	6.72	73.27	3.41	0.16	0.18	Glauner *et al.* (1988)
KN 126	LB	30	Sta	5.15	21.37	8.92	53.95	9.76	0.13	0.72	Glauner *et al.* (1988)
MC4100	LB	37	Exp	3.45	17.34	11.69	59.73	6.26	0.27	1.26	Obermann and Holtje (1994)
P678-54 minicells	LB	37	Exp	1.69	18.47	14.39	58.14	5.65	0.86	0.81	Obermann and Holtje (1994)
GC2700	LB	37	Exp	2.1	7.3	7.8	79.9	2.4	0.2	0.3	Signoretto *et al.* (1996)
GC2700[c]	LB + Mc	37	Exp	2.5	7.0	12.5	75.0	2.5	0.2	0.4	Signoretto *et al.* (1996)
GC2702	LB	37	Exp	1.9	6.8	8.2	80.1	2.5	0.2	0.3	Signoretto *et al.* (1996)
GC2702[c]	LB + Mc	37	Exp	2.9	9.0	15.0	69.8	2.9	0.2	0.2	Signoretto *et al.* (1996)
GC5391[c]	LB	37	Exp	3.1	7.5	15.3	70.8	2.7	0.2	0.4	Signoretto *et al.* (1996)
			AVERAGES	3.4	13.0	10.0	69.0	4.0	0.2	0.5	

[a] Incubation media, temperature in °C and growth phase - exponential (Exp) or stationary (Sta) - are indicated. Refer to the relevant references for the recipes to Panassay broth (PB) and LB, and for the concentration of mecillinam (Mc). [b] Values were calculated (or recalculated) from raw data if provided as per Glauner (1988) and are expressed as the percentage of total monomeric peptides. Significant figures are as per the cited article. [c]. Cells are coccal-shaped.

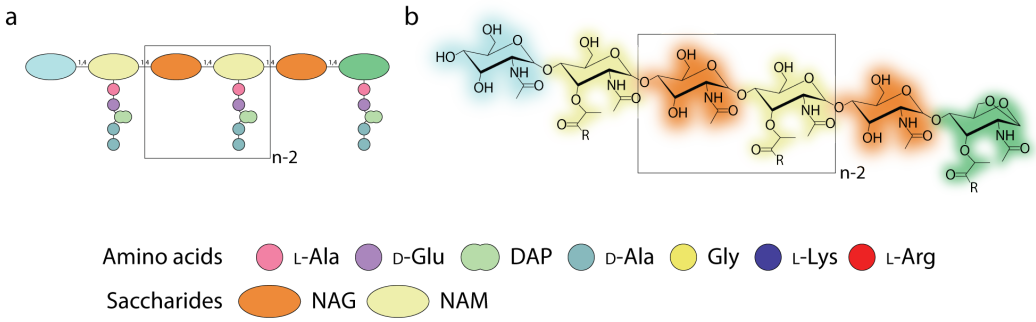


Figure 1.4.3 Glycosidic interactions between muropeptide subunits.

Each muropeptide chain has a variable number of "n" subunits and is terminated by the variant NAG-1,6-anhydro-NAM pair. On average, about 33 muropeptides constitute each chain, but depending on the bacterial strain and growth conditions, chains lengths of between 7 and 76 muropeptides have been observed as indicated in Table 1.4.2. Each glycan residue is connected to the next via a β-1,4-glycosidic bond as indicated by the cartoon depiction (**a**) and the chemical structure (**b**) of a muropeptide chain. Although one of seven peptides may be found on the NAM residue, for clarity only a pentapeptide (in **a**) or an "R" to represent any peptide (in **b**) is shown.

Table 1.4.2 Peptidoglycan composition.

Peptidoglycan composition is indicated. Only references where the author(s) determined each of the listed characteristics were included and used to determine composition averages.

	Growth conditions[a]			Sample	Oligomerisation state[c]				Crosslinkages[c,d]		Chain	Lpp	
***E. coli* strain**	media	°C	phase or rate	**taken[b]**	Mono	Di	Tri	Tetra	D,D	D,L	**length[e]**	**bound[c]**	**References**
W7	Min	37	48' doubling	All	55.3	39.1	5.2	0.4	21.7	1.6	24.3	5.6	Tuomanen and Cozens (1987)
W7	Min	37	180' doubling	All	50.6	41.7	6.8	0.9	23.7	2.3	11.1	8.0	Tuomanen and Cozens (1987)
W7	Min	37	360' doubling	All	49.9	43.0	6.2	0.9	21.8	4.2	9.9	8.1	Tuomanen and Cozens (1987)
W7	Min	37	828' doubling	All	47.8	45.9	5.2	1.1	19.5	7.6	6.8	11.4	Tuomanen and Cozens (1987)
W7	PB	30	Exp	All	54.07	41.4	4.36	0.15	21.54	2.19	30.3	4.97	Glauner (1988)
KN 126	PB	30	Exp	All	55.30	40.5	4.06	0.13	21.40	1.66	33.3	5.47	Glauner (1988)
KN 126	PB	37	Exp	All	54.47	41.2	4.16	0.13	21.38	1.54	36.86	6.80	Glauner *et al.* (1988)
KN 126	PB	42	Exp	All	55.37	40.8	3.68	0.14	20.62	2.34	37.90	5.76	Glauner *et al.* (1988)
KN 126	LB	30	Exp	All	51.52	43.2	5.02	0.19	22.36	2.76	25.80	4.98	Glauner *et al.* (1988)
KN 126	LB	30	Sta	All	39.40	50.5	9.85	0.22	26.80	5.20	17.80	8.55	Glauner *et al.* (1988)
MC4100	Min	37	50' doubling	At 5'	62.2	34.2	3.2	0.4	19.2	0.3	42.6	1.1	de Jonge *et al.* (1989)
MC4100	Min	37	50' doubling	At 35'	58.2	39.2	2.5	0.1	21.0	0.4	70.5	1.3	de Jonge *et al.* (1989)
MC4100	Min	37	50' doubling	At 60'	61.6	35.5	2.8	0.1	19.4	0.3	54.2	1.7	de Jonge *et al.* (1989)
MC4100	Min	37	50' doubling	At 80'	59.2	38.2	2.6	0.1	20.6	0.3	68.0	1.9	de Jonge *et al.* (1989)
MC4100 Δ*lysA* Δ*ftsA*	Min	28	70' doubling	At 5'	63.7	33.9	2.3	0.1	18.5	0.1	45.4	0.4	de Jonge *et al.* (1989)
MC4100 Δ*lysA* Δ*ftsA*	Min	28	70' doubling	At 55'	57.4	40.4	2.1	0.1	21.6	0.1	68.0	0.4	de Jonge *et al.* (1989)
MC4100 Δ*lysA* Δ*ftsA*	Min	37[f]	50' doubling	At 5'	66.7	31.7	1.5	0.1	16.8	0.1	53.0	0.3	de Jonge *et al.* (1989)
MC4100 Δ*lysA* Δ*ftsA*	Min	37[f]	50' doubling	At 35'	67.1	31.2	1.6	0.1	16.5	0.2	76.0	0.2	de Jonge *et al.* (1989)
MC4100 Δ*lysA*	Min	28	Exp	All	52.5	42.3	4.9	0.3	23.6	1.1	31.3	4.1	de Jonge (1990)
MC4100 Δ*lysA*	Min	37	Exp	All	53.0	41.1	5.6	0.3	23.6	0.9	31.1	3.7	de Jonge (1990)

MC4100 Δ*lysA* Δ*ftsA*	Min	28	Exp	All	54.5	41.1	4.2	0.2	22.4	1.1	30.2	3.7	de Jonge (1990)
MC4100 Δ*lysA* Δ*ftsA*	Min	37[f]	Exp	All	54.2	40.5	5.0	0.3	22.5	1.3	30.3	3.0	de Jonge (1990)
[g]MC4100 Δ*lysA pbpB*^{r1}	Min	28	Exp	All	52.9	41.8	5.0	0.3	23.4	1.0	30.0	4.1	de Jonge (1990)
JM83	Min	37	Exp	All	52.4	42.5	4.8	0.3	21.6	3.0	27.5	8.5	Korat *et al.* (1991)
JM83 (+*dacB*)	Min	37	Exp	All	63.0	34.7	2.2	0.1	11.9	7.0	29.9	9.0	Korat *et al.* (1991)
JM83	LB	37	Exp	All	55.3	40.0	4.5	0.2	20.5	2.6	35.0	7.9	Korat *et al.* (1991)
JM83 (+*dacB*)	LB	37	Exp	All	67.2	30.6	2.2	0.0	12.2	4.6	36.4	8.7	Korat *et al.* (1991)
HB101	LB	42	Exp	All	51.1	43.5	5.2	0.2	21.9	3.5	22.3	9.3	Korat *et al.* (1991)
HB101 (+*dacB*)	LB	42	Exp	All	62.1	34.5	3.3	0.1	14.4	5.2	23.1	10.2	Korat *et al.* (1991)
W7	Min	37	40' doubling	All	49.80	44.80	5.00	0.40	22.51	3.54	17.20	4.10	Vincent *et al.* (1991)
W7 (128×Ofl)	Min	37	40' doubling	All	49.30	45.96	4.37	0.37	19.48	6.67	28.33	3.35	Vincent *et al.* (1991)
W7 (128×Pef)	Min	37	40' doubling	All	48.25	48.73	2.70	0.32	19.01	7.39	24.27	2.96	Vincent *et al.* (1991)
W7 (3×Ofl)	Min	37	40' doubling	All	50.21	45.52	4.27	0.00	20.31	5.29	9.70	5.75	Vincent *et al.* (1991)
W7 (3×Pef)	Min	37	40' doubling	All	46.31	46.50	6.79	0.40	21.72	6.36	12.74	6.50	Vincent *et al.* (1991)
MC4100	LB	37	Exp	All	47.63	45.04	7.08	0.25	23.10	2.29	29.91	8.94	Obermann and Holtje (1994)
P678-54 minicells	LB	37	Exp	All	45.19	47.08	7.49	0.25	22.09	3.87	24.36	10.58	Obermann and Holtje (1994)
				AVERAGES	55.1	40.5	4.2	0.2	20.4	2.7	35.7	4.9	

[a] Incubation media, temperature (in °C), and growth phase - exponential (Exp) or stationary (Sta) phases - or the doubling time are indicated. Refer to the relevant references for recipes for minimal media (Min), Panassay broth (PB) and LB. [b] Either the total ("All") peptidoglycan composition was determined, or newly synthesised peptidoglycan following at the indicated time post induction. [c] Values were calculated (or recalculated) from raw data if provided as per Glauner (1988) and are expressed as the percentage of total muropeptides. Significant figures are as per the cited article. Mono-, di-, tri-, and tetra-mers are the four known peptide oligomerisation states. [d] Crosslinkages refer to the bond itself, rather than the muropeptides involved in crosslinkages, which is the sum of di-, tri- and tetra-mers. [e] Average muropeptide chain length was calculated (or recalculated) from raw data if provided as per Glauner (1988). [f] Filamentous growth. [g] Cells have pointed cell poles. [h] Strain growth media were supplemented with ofloxacin (Ofl) or pefloxacin (Pef).

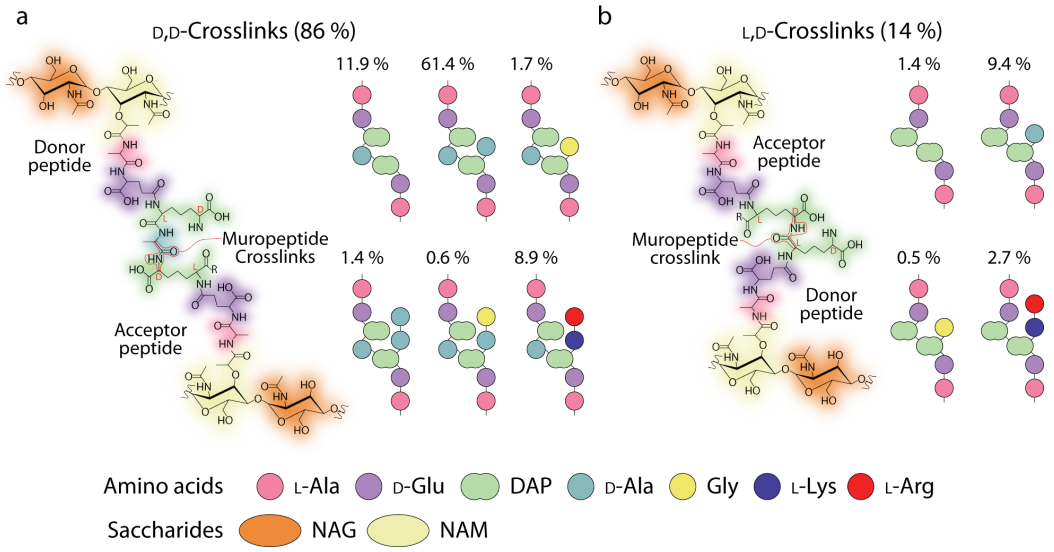


Figure 1.4.4 Peptide crosslinks between muropeptide subunits.

The two types of peptide crosslinkages that form between muropeptide subunits are shown. For clarity: the "acceptor" and "donor" peptides are labelled; the L and D centres of DAP are depicted in red, and; the residues and bonds that constitute the planar peptide crosslink are circled in red. The chemical structures of the **a,** D,D-crosslink between tetrapeptide-containing muropeptides and **b,** L,D-crosslink between tetra- and tri-peptide-containing muropeptides are shown. To the right of each structure is a cartoon depiction of the known dimeric muropeptides (glycan subunits are omitted for clarity) with the average abundance indicated as per Table 1.4.3.

Table 1.4.3 The distribution of dimeric peptides depends on the strain and growth conditions.

The distribution of dimeric peptides are indicated. Only references where the author(s) determined the abundance of all known dimeric peptides were included and used to determine dimeric peptide averages. Dimeric muropeptides contain crosslinkages beween the peptide component, where each crosslinkage may comprise: tripeptides with (3-Lpp) or without (3) covalently bound Lpp, tetrapeptides with (4g) or without (4) a glycine at position 4 and pentapeptides with (5g) or without (5) a glycine at position 5.

				Distribution of dimeric peptides[b]										
	Growth conditions[a]			**D,D-crosslinkages**						**L,D-crosslinkages**				
***E. coli* strain**	Media	°C	phase	4-3	4-4	4-4g	4-5	4-5g	4-3-Lpp	3-3	3-4	3-4g	3-3-Lpp	**References**
KN 126	PB	37	Exp	8.83	71.21	3.65	0.42	0.80	7.98	0.70	5.17	0.17	1.07	Glauner *et al.* (1988)
MC4100	LB	37	Exp	14.35	56.44	5.68	0.89	1.26	11.28	1.57	6.33	0.48	1.72	Obermann and Holtje (1994)
P678-54 minicells	LB	37	Exp	9.60	53.54	4.44	1.77	1.06	12.22	2.36	10.59	1.39	3.02	Obermann and Holtje (1994)
GC2700	LB	37	Exp	13.3	67.0	0.0	1.3	0.3	7.4	0.3	8.4	0.3	1.6	Signoretto *et al.* (1996)
GC2700[c]	LB + Mc	37	Exp	11.6	64.5	0.0	1.8	0.3	6.8	1.3	10.0	0.3	3.4	Signoretto *et al.* (1996)
GC2702	LB	37	Exp	12.5	67.2	0.0	1.3	0.0	7.2	0.6	9.1	0.3	1.9	Signoretto *et al.* (1996)
GC2702[c]	LB + Mc	37	Exp	12.6	55.6	0.0	2.0	0.9	8.3	2.3	12.9	0.6	4.9	Signoretto *et al.* (1996)
GC5391[c]	LB	37	Exp	12.8	55.8	0.0	1.7	0.6	9.7	1.9	12.8	0.6	4.2	Signoretto *et al.* (1996)
			AVERAGES	11.9	61.4	1.7	1.4	0.6	8.9	1.4	9.4	0.5	2.7	

[a] Incubation media, temperature in °C and growth phase - exponential (Exp) or stationary (Sta) - are indicated. Refer to the relevant references for the recipes to Panassay broth (PB) and LB, and for the concentration of mecillinam (Mc). [b] Values were calculated (or recalculated) from raw data if provided as per Glauner (1988) and are expressed as a percentage of total dimeric muropeptides. Significant figures are as per the cited article. Because this table averages a set of 8 data points and Table 1.4.2 averages a set of 36, the average values for total percentage D,D- and L,D-crosslinks vary between the two tables. [c]. Cells are coccal-shaped.

differences derive from variations in the number of amino acids, type and number crosslinks and type of disaccharide pair and are influenced by the media type, incubation temperature, growth phase and bacterial strain (Tables 1.4.1-1.4.3). In *E. coli* for example, between early logarithmic and mid-stationary phases, the number of monomeric peptides and the average chain length decreases, whereas the number of dimeric peptides and L,D-crosslinks increases (Table 1.4.2). Despite this, over 50 % of muropeptide species contain the L-Ala-D-Glu-DAP-D-Ala tetrapeptide from monomeric to tetrameric muropeptide crosslinked species (Tables 1.4.1-1.4.3).

Synthesis of the muropeptide subunit begins in the cytoplasm, where activated uridine diphosphate (UDP) precursors of NAG and NAM-pentapeptide are formed (Figure 1.4.5). Following their incorporation onto the universal glycan lipid carrier - undecaprenyl diphosphate - at the cytoplasmic leaflet, the newly synthesised muropeptide subunit is flopped to the periplasmic leaflet and incorporated into a nascent muropeptide chain (Typas *et al.*, 2012; Vollmer & Bertsche, 2008). A series of transpeptidases and transglycosylases then ensure that the nascent muropeptide chain is correctly incorporated into the murein sacculus (Typas *et al.*, 2012; Vollmer & Bertsche, 2008).

Peptidoglycan may be extensively remodelled or degraded, especially during cell division, but also during the movement of protein across the peptidoglycan barrier. While the average diameter of peptidoglycan pores is generally accepted to be about 2.06 nm (Demchick & Koch, 1996), pores of up to 10 nm in diameter have been reported in peptidoglycan (Turner *et al.*, 2013) potentially explaining

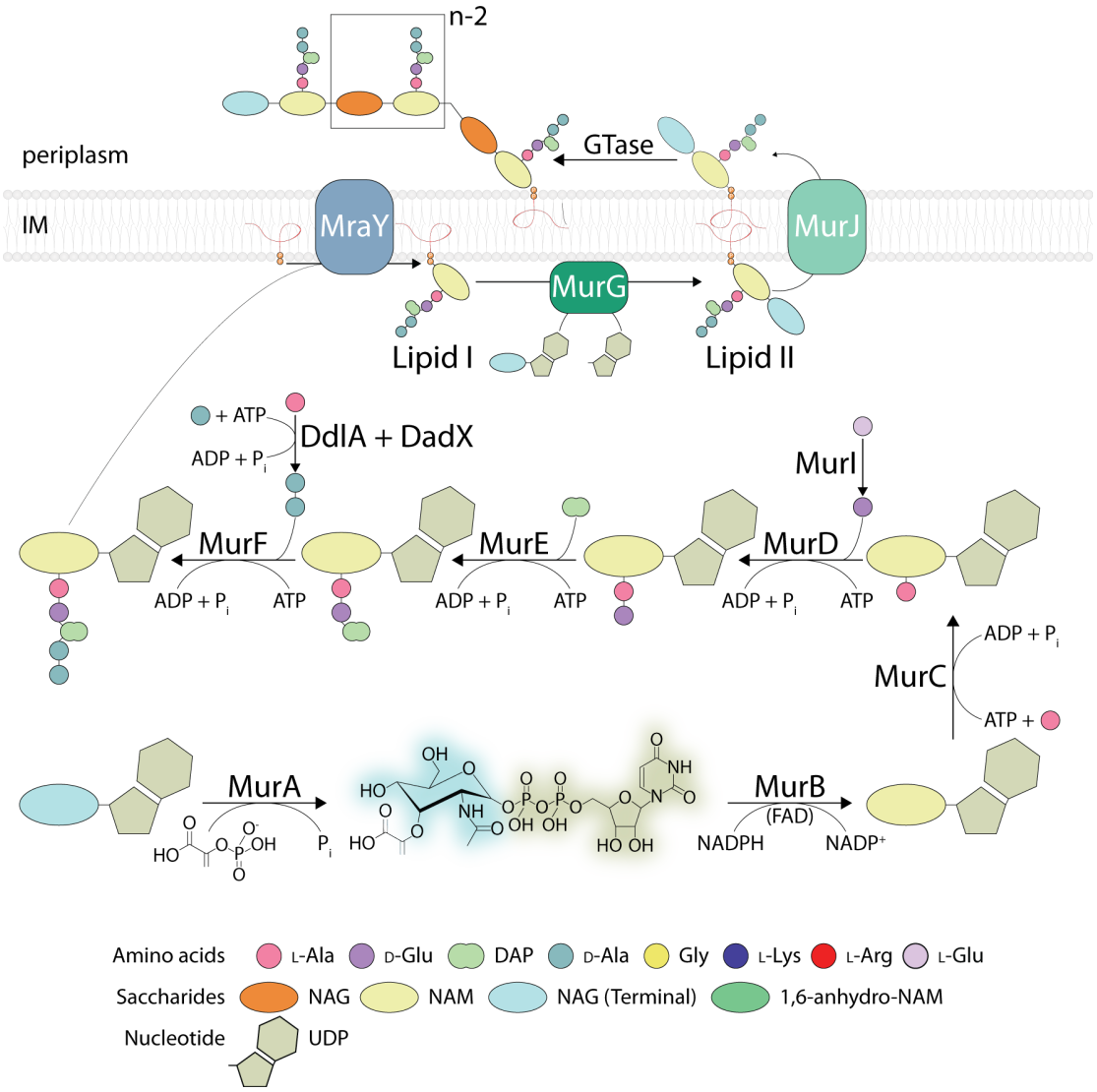


Figure 1.4.5 Muropeptide synthesis and chain elongation.

A cartoon depiction, including one chemical structure, of muropeptide synthesis in the cytoplasm is shown. After UDP-NAG is converted into UDP-NAM-pentapeptide by the a series of Mur enzymes, it is transferred onto undecaprenyl phosphate to form lipid I. UDP-NAG is then used as a substrate to convert lipid I into the muropeptide-containing lipid II, which is subsequently flopped by MurJ to the periplasm where one of four glycosyltransferases (GTase) attaches it to a growing muropeptide chain of "n" subunits (Typas *et al.*, 2012).

how folded proteins of up to 100 kDa might pass through the peptidoglycan layer (Vázquez-Laslop *et al.*, 2001). Although the larger pores were not as abundant as the ~2 nm pores (Turner *et al.*, 2013), it is thought that the numerous enzymes normally involved in degrading and remodelling the murein sacculus may also facilitate passage of larger substrates by selectively generating these larger pores (Typas *et al.*, 2012; Vollmer & Bertsche, 2008). Indeed, flagella, type 4 pili and the type 3, 4 and 6 secretion systems (T3SS, T4SS and T6SS, respectively) all encode dedicated peptidoglycan degradation proteins to ensure a suitable pore size is available for biogenesis of their periplasmic-spanning apparatus (Burkinshaw *et al.*, 2015; Chou *et al.*, 2012; Herlihey *et al.*, 2014; Scheurwater & Burrows, 2011; Zahrl *et al.*, 2005).

The peptidoglycan layer also provides an important structural stability to the OM (Figure 1.4.6). A common interaction is provided by Braun's lipoprotein (Lpp), which is anchored to the OM via its N-terminal lipid residues, and covalently attached to the L-centre of DAP through its C-terminal lysine residue (Magnet *et al.*, 2007) (Figure 1.4.6a). Roughly one-third of Lpp are bound to peptidoglycan, corresponding to 5.3% of Lpp-bound muropeptides (Table 1.4.2). Although the function of the remaining unbound pool of Lpp is unknown, they are surface-exposed (Cowles *et al.*, 2011) and thought to adopt a trimeric conformation (Choi *et al.*, 1986; Shu *et al.*, 2000).

The peptidoglycan-associated lipoprotein, Pal, and an integral OMP, OmpA, are also important for anchoring the OM to the peptidoglycan layer through a strong hydrogen-bonded network with DAP (Park *et al.*, 2012) (Figure 1.4.6b). Like Lpp,

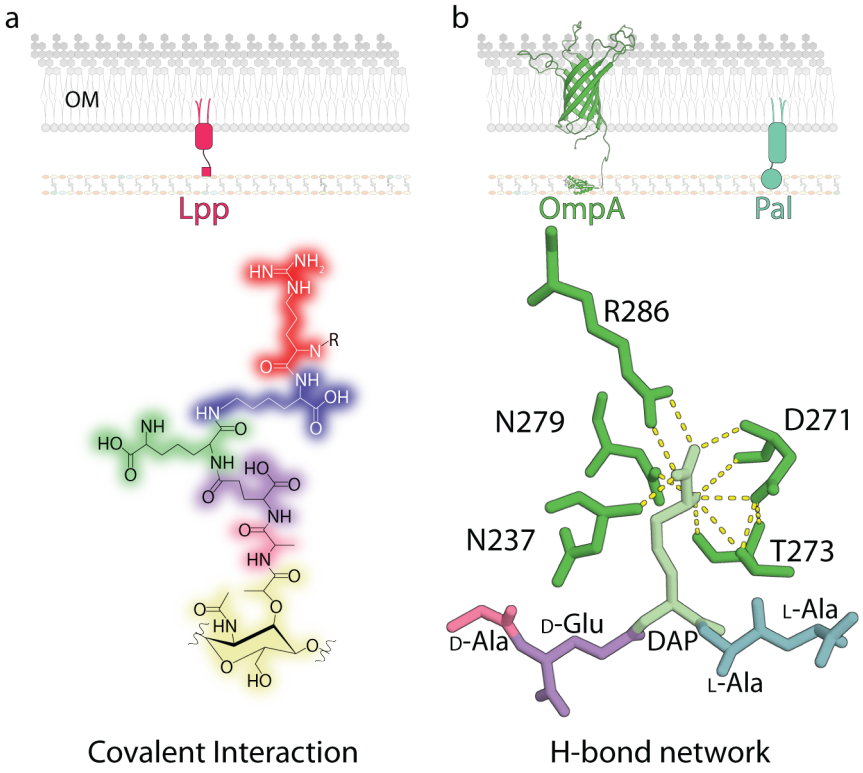


Figure 1.4.6 Major interactions between the cell wall and outer membrane.

The outer membrane is tethered to the cell wall by Braun's lipoprotein (Lpp), OmpA and Pal. **a,** Cartoon depiction of the covalent interaction between Lpp and a muropeptide. Below, the chemical structure of the terminal Lpp residues (Lys-Arg-R, where "R" represents the remaining residues comprising Lpp) covalently attached to a muropeptide (with NAG omitted) are shown. **b,** Ribbon diagram of OmpA (PDB: 2GE4) and a cartoon depiction of Pal. Both proteins are tethered to the cell wall via their similar OmpA-type peptidoglycan-binding domain, depicted as a ribbon diagram (PDB: 4G4Z) for OmpA and as a circle for Pal. Below, hydrogen bond network formed between DAP and the relevant OmpA residues is shown, as per Park *et al.* (2012).

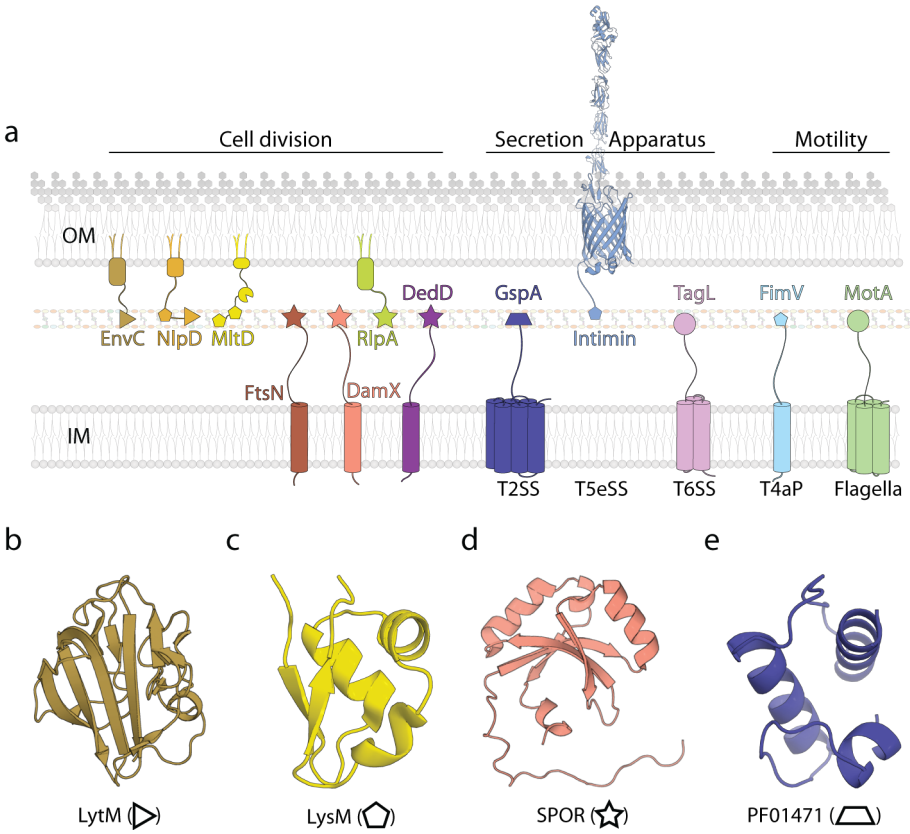


Figure 1.4.7 Minor interactions between the cell wall and outer membrane.

a, A number of proteins are "anchored" to the cell wall to provide structural support during cell division or to maintain the structural integrity of secretion and motility apparatus. In addition to the major peptidoglycan-OM interactions depicted in Figure 1.4.6, four other binding domains are known. The shape used to depict each binding domain corresponds to the type of fold depicted in **b-e**, where circles represent the OmpA-type peptidoglycan-binding domain. The combined crystal structures of intimin (PDB: 4E1S, 1F02 and Phyre2 homology model of residues 455-653 modelled after the same domain from invasin PDB: 1CWV) are also shown. **b,** The LytM-like peptidoglycan-binding domain of EnvC (PDB: 4BH5) is shown. **c,** The C-terminal (i.e. the second) LysM-like peptidoglycan-binding domain of MltD (PDB: 1E0G) is shown. **d,** The SPOR peptidoglycan-binding domain of DamX (PDB: 2LFV) is shown. **e,** A Phyre2 homology model of the PF01471 peptidoglycan-binding domain modelled after the same domain from EpsAB (PDB: 4G54).

Pal may also be surface-exposed, although Michel *et al.* (2015) found that rather than an individual cell harbouring a dual Pal orientation, Pal was surface-exposed in an "all or nothing" manner. A number of other proteins also anchor to the peptidoglycan layer (Scheurwater & Burrows, 2011) (Figure 1.4.7). These proteins may facilitate interactions between peptidoglycan and cell division or peptidoglycan degradation apparatus, or play a structural role for secretion or mobility systems (Figure 1.4.7a). At least four other peptidoglycan-binding domains are known (Figure 1.4.7b-e), and several others are suspected. For example, Pucciarelli and García-del Portillo (2003) suggest proteins such as InvH, PrgH and PrgK harbour as yet uncharacterised peptidogylcan-binding domains.

1.5 Inner membrane protein insertion and translocation pathways

Following synthesis, proteins with an extracytoplasmic destination may follow one of several pathways to ensure correct localisation. The vast majority of envelope proteins are transported in an unfolded state via the Sec pathway for IM insertion or translocation (Denks *et al.*, 2014), whereas about 28 proteins require cytoplasmic assembly before they are exported through the twin-arginine translocation (Tat) pathway (Palmer *et al.*, 2005; it should be noted therein that the authors misreported their own tabulated data, which listed 28 proteins, as "27" Tat substrates). Additionally, bacteria have evolved several secretion systems designed to export virulence factors to the cell surface, extracellular milieu or host cells (Costa *et al.*, 2015). Although some of these virulence factors may utilise the Sec and/or Tat pathways to facilitate translocation of effector proteins, the type 1 secretion system (T1SS), T3SS and T6SS form an intermembrane-spanning platform capable of directly excreting effectors from

the cytoplasm (Costa *et al.*, 2015).

Inner membrane translocation or insertion via the Sec pathway is mediated by the essential SecYEG complex. Translocation across SecYEG occurs through a pore of only 2.2-2.4 nm in diameter (Bonardi *et al.*, 2011) and IM insertion occurs through a lateral gate in SecY that opens into the lipid bilayer (du Plessis *et al.*, 2009). The relatively small pore size means that SecYEG can only translocate unfolded polypeptides or accommodate a single transmembrane α-helix, so for polytopic inner membrane proteins, each α-helix must be transferred through the lateral gate before the next α-helix can be assembled (Luirink *et al.*, 2012) (Figure 1.5.1).

Sec-dependent polypeptides are transferred to SecYEG in a post- or co-translational manner, which normally dictates whether the nascent polypeptide chain will be translocated across or inserted into the membrane (Luirink *et al.*, 2012). As a rule of thumb, the post-translation pathway translocates periplasmic proteins and integral OMPs and the co-translation pathway inserts integral inner membrane proteins and exports lipoprotein (Denks *et al.*, 2014). The most notable exception includes a subset of integral OMPs, the serine protease autotransporters of *Enterobacteriaceae* (SPATEs), which have been described to translocate across the IM in a co-translational manner (Jong & Luirink, 2008; Jong *et al.*, 2010; Sijbrandi *et al.*, 2003; Szabady *et al.*, 2005).

The post-translation pathway requires the cytoplasmic tetrameric chaperone, SecB, and the ATPase, SecA. After translation, preprotein is kept in an unfolded state by SecB and transferred to SecA at the cytoplasmic interface of SecYEG. SecB

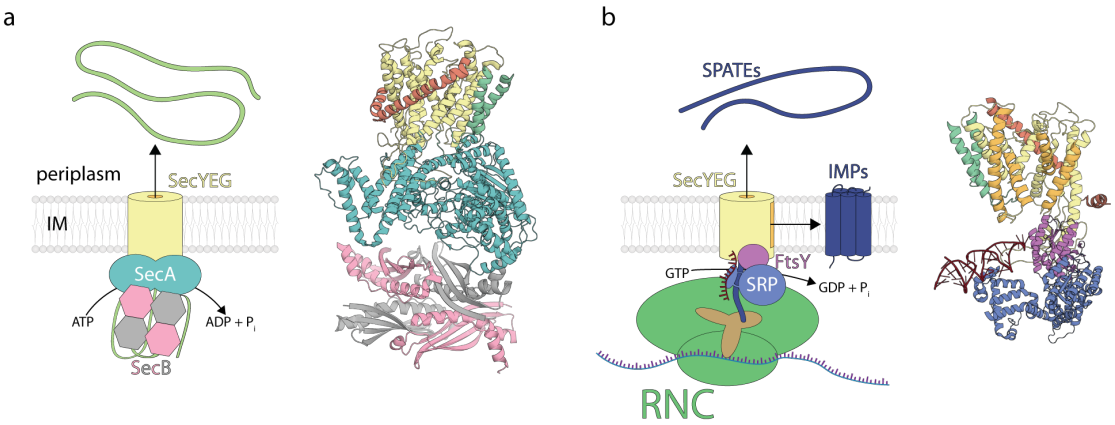


Figure 1.5.1 Sec-dependent translocation pathways.

The majority of proteins translocated across the IM follow the post-translational (SecA-dependent pathway). However, a subset of OMPs - collectively referred to as the SPATEs - translocate across the IM via the co-translational (SRP-dependent) pathway that is usually responsible for inner membrane protein insertion. **a,** A cartoon representation of the post-translational pathway is shown and the similarly coloured ribbon diagrams of the pathway's relevant components are also shown: SecA-SecYEG (PDB: 3DL8) and the SecB tetramer (PDB: 1QYN). For clarity, the SecB monomers are alternately coloured grey and pink, and in the ribbon diagram only, SecE and SecG are coloured maroon and green, respectively. **b,** A cartoon representation of the co-translational pathway is shown, where both the translocation pore and lateral gate of SecYEG are coloured orange. The RNC is depicted here as: green ribosome, brown (oversized) tRNA, blue and purple mRNA and dark blue nascent polypeptide chain. Ribbon diagrams of SecYEG (PDB: 3DL8) and the SRP-FtsY complex (PDB: 5GAD) are also shown. SecYEG is coloured as before, except transmembrane α-helices 2b, 3, 7 and 8 of SecY - which constitute the lateral gate - are coloured orange. For clarity, Ffh and the 4.5S RNA that constitute the SRP are coloured blue and dark brown, respectively.

dissociates from the complex as SecA energises the ATP-dependent translocation of the unfolded substrate polypeptide into the periplasm (Luirink *et al.*, 2012) (Figure 1.5.1a). In contrast, during the co-translation pathway - as the name suggests - the entire ribosome-nascent chain (RNC) complex is transferred to SecYEG for secretion of the unfolded polypeptide chain as it is being translated (Figure 1.5.1b). The signal-recognition particle (SRP), which is a ribonucleoprotein complex comprised of a 4.5S RNA and Ffh, is responsible for recognising the translating polypeptide within the RNC (Denks *et al.*, 2014). The SRP-RNC complex then binds to FtsY at the cytoplasmic interface of SecYEG, where FtsY in conjunction with SRP mediates the GTP-dependent secretion of polypeptide (Denks *et al.*, 2014).

The Sec pathway may recruit a number of auxiliary proteins to facilitate protein translocation or insertion, but the conditions for their recruitment are not well characterised. YidC, which may directly insert small IM proteins in a Sec-independent manner (Dalbey *et al.*, 2014), binds at the SecYEG lateral gate where it may facilitate or direct inner membrane protein insertion (Sachelaru *et al.*, 2013), but Jong *et al.* (2010) also suggest its role extends to the post-translocation chaperoning of SPATEs. The SecDF-YajC complex is thought to energise protein translocation or insertion through proton motive force, but it is considered to be a very low-abundancy partner protein to SecYEG (Pogliano & Beckwith, 1994). The two other known auxiliary proteins, PpiD and YfgM, are thought to work together by facilitating Sec translocation (Götzke *et al.*, 2015; Matern *et al.*, 2010). PpiD competes with YidC for binding at the SecYEG lateral gate (Sachelaru *et al.*, 2014) and YfgM was found to form a complex with PpiD and SecYEG-PpiD but not

YidC (Götzke *et al.*, 2014; Maddalo *et al.*, 2011).

The Tat machinery consists of an inner membrane protein complex, TatBC, that on contact with its substrate, recruits a variable number of TatA monomers to form an IM translocation pore (Tarry *et al.*, 2009) (Figure 1.5.2). Recently however, it has been suggested that TatA recognises Tat substrates independently of the TatBC complex (Taubert *et al.*, 2015). In either case, and in contrast to the Sec-dependent translocation, the Tat pathway translocates folded proteins through a variably-sized pore, ranging from 3 nm to 7 nm in diameter (Gohlke *et al.*, 2005). The differing pore sizes correlates well with the size range of the various Tat substrates, indicating that they enter an appropriately sized translocation pore that does not leak ions (Gohlke *et al.*, 2005).

Although for some proteins it remains unclear why they require pre-folding in the cytoplasm when assembly in the periplasmic assembly following Sec translocation could suffice (Natale *et al.*, 2008). Palmer and Berks (2012) reported that the majority of Tat-dependent proteins fall into at least one of three categories. Firstly, they may require complex cofactors (such as nucleotide-based cofactors or metal-sulphur clusters) that are not present in the periplasm and/or require cytoplasmic chaperones that are normally responsible for cofactor-protein complex assembly. Secondly, they may require the cytoplasmic environment to bind low-affinity metal cations that would otherwise be outcompeted by high-affinity metal ions if folding were to occur in the periplasm. Thirdly, to translocate proteins without a Tat-targeting signal, that form part of a special hetero-oligomeric complex. In this complex, at least one protein will have a Tat-targeting signal, so any protein in the

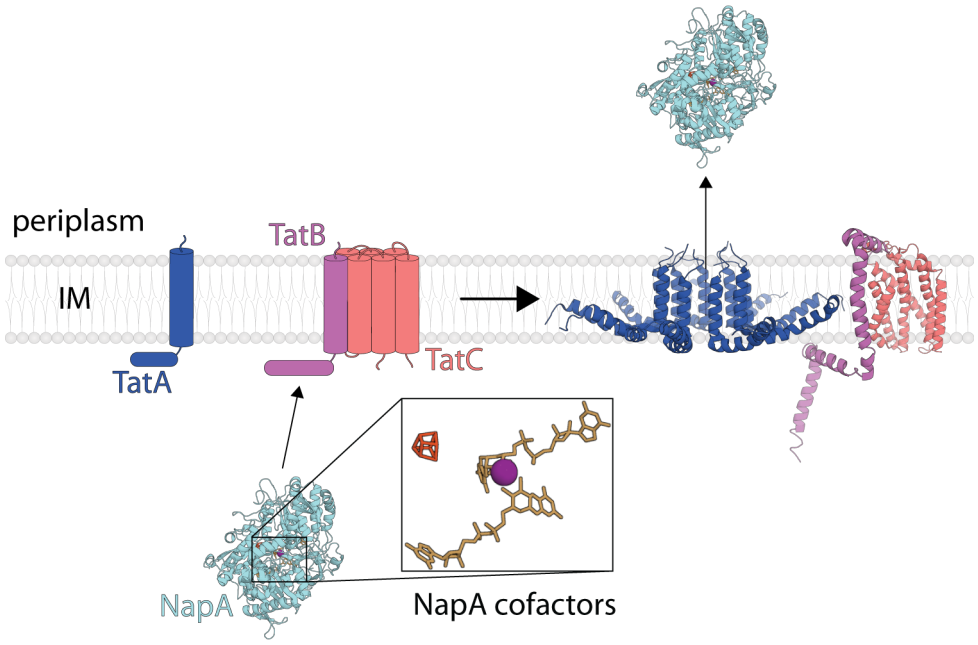


Figure 1.5.2 Tat-dependent translocation.

There are about 28 known or predicted Tat-dependent proteins in *E. coli (Palmer et al., 2005)* that assemble in the cytoplasm before translocation. The periplasmic nitrate reductase, NapA (PDB: 2NYA), is one such example and requires the cytoplasmic incorporation of an iron-sulphur cluster (red-orange) and a molybdenum ion (purple) coordinated to *bis*-molybdopterin guanine dinucleotide (brown) (Jepson *et al.*, 2007). A cartoon depiction is shown containing the TatBC pre-pore complex and monomeric TatA subunits. After NapA binds the pre-pore complex, TatBC recruits TatA to initiate its oligomerisation into a translocation pore through which NapA may enter the periplasm (Palmer & Berks, 2012). The final TatABC pore complex is depicted by ribbon diagrams of the TatA oligomer (PDB: 2LZS) interacting with TatB (PDB: 2MI2) and TatC (PDB: 4HTS).

complex without the signal will be translocated via the Tat-pathway in a 'piggy-back' manner.

1.6 Sequence motifs and signal peptides

The *E. coli* targeting sequences, commonly known as signal sequences or signal peptides, are located at the N-terminus of preproteins and prolipoproteins. Signal peptides are typically comprised of an N-terminal, positively-charged "n-region"; a hydrophobic "h-region" that typically spans the IM, and; a C-terminal, hydrophilic "c-region" (Pugsley, 1993) (Figure 1.6.1). An unusual exception are the SPATEs, which contain an extra n- and h-region, presumably to ensure targeting to SRP rather than SecB (Desvaux *et al.*, 2007; Szabady *et al.*, 2005). Excluding the majority of inner membrane proteins, which retain their signal sequence, signal peptides are cleaved by one of three dedicated signal peptidases (SPases) (Paetzel, 2014; Pugsley, 1993) (Figure 1.6.2a-c) in either their c-region (type 1 and 2 SPases) or n-region (type 3 SPases). The remaining membrane-spanning signal peptide by-product produced following type 1 or 2 SPase cleavage is then degraded into smaller peptides within its h-region by SppA (Kim *et al.*, 2008; Wang *et al.*, 2008) (Figure 1.6.2d) and following ejection into the cytoplasm, they are further degraded by cytoplasmic proteases and ultimately recycled (Novak & Dev, 1988; Novak *et al.*, 1986).

Signal sequences contain consensus motifs that determine SPase specificity and Sec- or Tat-dependence (Figure 1.6.1). Indeed, a number of algorithms have been developed to predict the presence and type of signal peptide with high confidence (Bendtsen *et al.*, 2005a; Bendtsen *et al.*, 2005b; Juncker *et al.*, 2003; Petersen *et al.*,

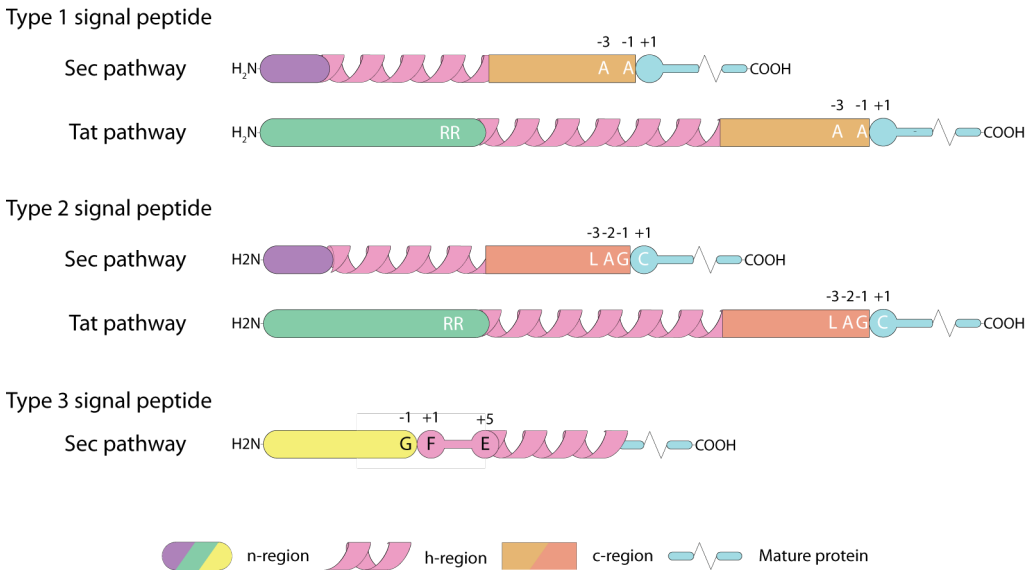


Figure 1.6.1 Sequence motifs of various signal peptides.

Several examples of signal peptides are shown. The single letter amino acids are used to indicate the most conserved residues that are required for cleavage by the indicated SPase, and the amino acid position is also indicated relative to the cleavage site. The "RR" indicated in the Tat signal peptide n-region represents the conserved twin-arginine sequence motif. Tat signal peptides are about 50 % longer than Sec pathway signal peptides as indicated, due to larger n- and/or h-regions (Paetzel, 2014; Pugsley, 1993). Although pre-pilin may contain amino acids that correspond to a c-region, which sometimes harbours a type 1 SPase motif, they are considered to not have a c-region because the mature protein starts before any apparent c-region (Pugsley, 1993).

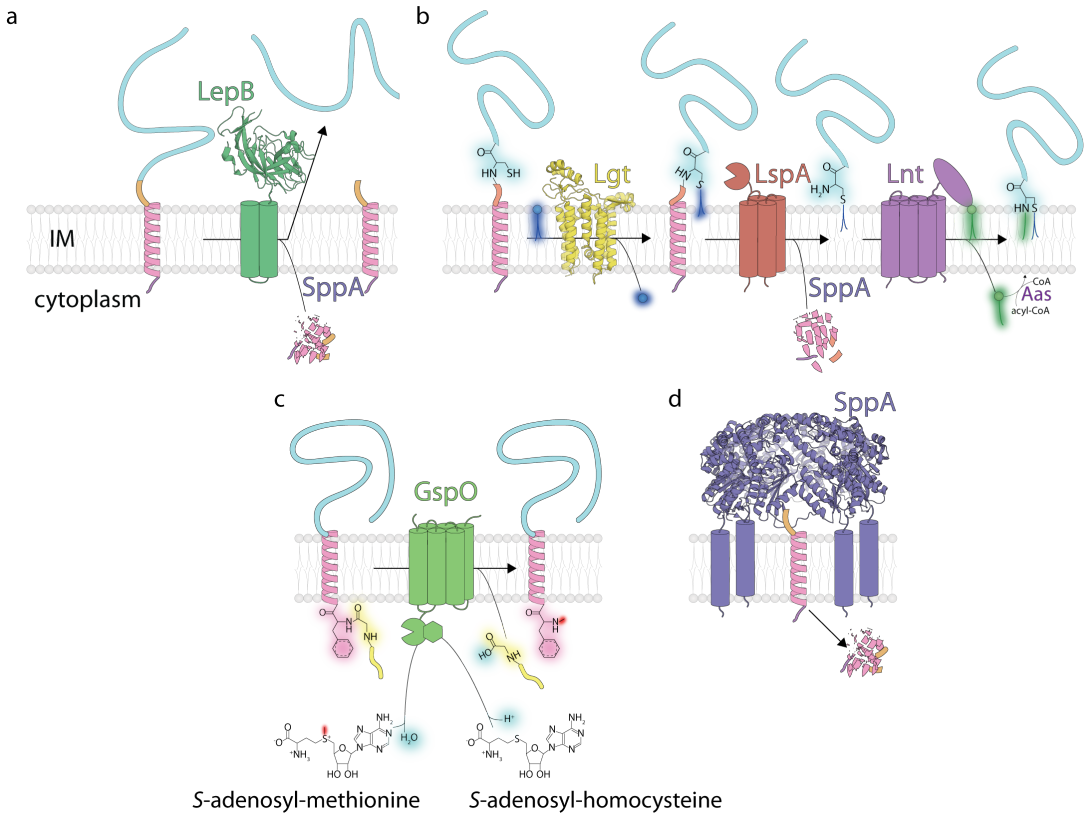


Figure 1.6.2 Mechanisms of signal peptide processing.

Shown are the three methods by which *E. coli* removes types 1-3 signal peptides (**a-b**) and how residual signal peptides are further processed by SppA (**d**). **a,** LepB (PDB: 3S04, periplasmic domain only) removes type 1 signal peptides, where the signal peptide by-product is further processed by SppA. **b,** The three-step prolipoprotein lipidation mechanism is indicated, where a ribbon diagram of Lgt (PDB: 5AZC) is shown. **c,** GspO cleaves type 3 signal peptide and *N*-methylates the +1 phenylalanine residue using *S*-adenosyl-methionine (Strom *et al.*, 1993), the methyl donor normally involved in DNA methylation (Sánchez-Romero *et al.*, 2015). For clarity, the relevant residues in the c- and h-region are indicated and atoms that are ultimately transferred to a new molecule are highlighted red (methyl group) or blue (water). **d,** Following type 1 or 2 signal peptide cleavage, the transmembrane signal peptide is further degraded by the tetrameric signal peptide peptidase, SppA (PDB: 3BEZ, periplasmic domain only).

2011). Tat signal peptides contain the eponymous "twin-arginine" sequence motif within their n-region: Ser/Thr-Arg-Arg-x_{pol}-Phe-Leu-Lys, where the two arginines are invariable and x_{pol} represents a polar residue (Berks, 1996). However, it is not the twin-arginine sequence motif that prevents targeting of Tat-dependent polypeptides to the Sec translocon, but the lower hydrophobicity of the h-region and the presence of basic residues in the c-region of Tat signal peptides (Cristóbal *et al.*, 1999).

Type 1 SPases cleave a conserved sequence motif within the c-region of preprotein (Figure 1.6.2a), which follows the "-3, -1" rule that requires small and neutral residues at the -3 and -1 positions relative to the cleavage site (Tuteja, 2005). In *E. coli*, alanine often occupies these positions, so the motif is sometimes known as the "Ala-x-Ala" motif: $Ala_{(-3)}$-$x_{any(-2)}$-$Ala_{(-1)}$, where x_{any} may be any residue (Tuteja, 2005). Prolipoprotein instead contains a "lipobox" sequence motif with an invariable +1 cysteine residue: $Leu_{(-3)}$-Ala/$Ser_{(-2)}$-Gly/$Ala_{(-1)}$-$Cys_{(+1)}$ (Zückert, 2014).

Following Sec or Tat translocation, prolipoproteins are converted into mature lipoprotein by a three-step process that includes the type 2 SPase and triacylation of the new N-terminal cysteine residue (Buddelmeijer, 2015; Kovacs-Simon *et al.*, 2011; Zückert, 2014) (Figure 1.6.2b). The first step involves the removal of a diacylglycerol moiety from PG by Lgt. While the glycerol-1-phosphate by-product is typically recycled, the diacylglycerol moiety is transferred to the thiol group of the +1 cysteine residue to generate diacylglyceryl-prolipoprotein. Cleavage of the type 2 signal peptide by LspA then occurs within the "lipobox" motif to generate

(i) apolipoprotein and (ii) a signal peptide by-product that is processed by SppA. The final step involves the transfer by Lnt of a 1-acyl group moiety from a phospholipid substrate to the free amine group of the diacylated cysteine, thereby generating the triacylated lipoprotein. Although Gupta *et al.* (1991) suggested that Lnt is capable of using PG or CL as substrates, it typically uses PE, where the resulting lysophospholipid by-product is either (i) degraded by PldB into smaller components for recycling or (ii) re-acyclated by Aas (Zhang & Rock, 2008).

Type 3 SPases, originally referred to as type 4 SPases for their ability to cleave type 4 prepilin signal peptides (Pugsley, 1993), recognise a $Gly_{(-1)}$-$Phe/Met_{(+1)}$ motif within the n-region of prepilin (or prepseudopilin) components in type 4 pili and type 2 secretion system (T2SS) machinery (Berry & Pelicic, 2015). This means that mature pilin retain their h-region to remain anchored in the IM following SPase processing; however, the reason for this IM-anchoring remains unknown, especially considering the extraction of these subunits from the IM to generate the type 4 pilus or T2SS pseudopilus would be energetically expensive (Craig & Li, 2008). Many type 3 SPases also methylate the new N-terminal residue following cleavage (Figure 1.6.2c), for which a conserved $Glu_{(+5)}$ is essential (Pasloske & Paranchych, 1988); however, the reason for the *N*-methylation of pilin subunits remains enigmatic, especially since it has not been observed in some bacterial and all archael homologues (Berry & Pelicic, 2015). Several additional sequence motifs important for inner membrane protein biogenesis are also known, but will not be discussed here: reverse signal anchors, helical-hairpin motifs and start- and stop-transfer sequences (Xie & Dalbey, 2008; Xie *et al.*, 2007).

1.7 Outer membrane lipoprotein biogenesis

Lipoproteins contain a sorting signal (Zückert, 2014) that determines whether it will remain anchored to the IM (~10%) or transferred to the OM inner leaflet (~90%) (Tokuda & Matsuyama, 2004). Yamaguchi *et al.* (1988) found that modifying the +2 residue of OM lipoproteins to aspartic acid abolished its OM localisation. Although the residues at the +3 and +4 positions also contribute to OM localisation in the presence of $Asp_{(+2)}$ (Gennity & Inouye, 1991) or IM retention in the absence of $Asp_{(+2)}$ (Seydel *et al.*, 1999), the IM retention signal is commonly referred to as the "+2 rule" (Zückert, 2014).

The Lol (lipoprotein OM localisation) pathway (Figure 1.7.1a) is the main pathway for lipoprotein transfer to the OM. Mature lipoprotein without an IM retention signal are recognised by the LolE component of the LolCDE complex. Once LolC recruits an empty LolA periplasmic chaperone, LolD energises the ATP-dependent transfer of mature lipoprotein from LolE to LolA. LolA is then ejected into the periplasm and transfers the mature lipoprotein to LolB (itself an OM lipoprotein), which inserts the lipoprotein into the OM periplasmic leaflet.

The T2SS, famous for its role in excreting cholera toxin (from *V. cholerae*) and heat-labile enterotoxin (from pathogenic *E. coli*), also plays a role in OM lipoprotein biogenesis (Rondelet & Condemine, 2013). However, rather than insertion into the periplasmic leaflet, the T2SS is thought to insert its lipoprotein substrates into the OM outer leaflet (Zückert, 2014) (Figure 1.7.1b). Lipoprotein substrates of the T2SS, including PulA from *Klebsiella* species and SslE from *E. coli*, must contain an IM retention (Lol avoidance) signal for T2SS-dependent

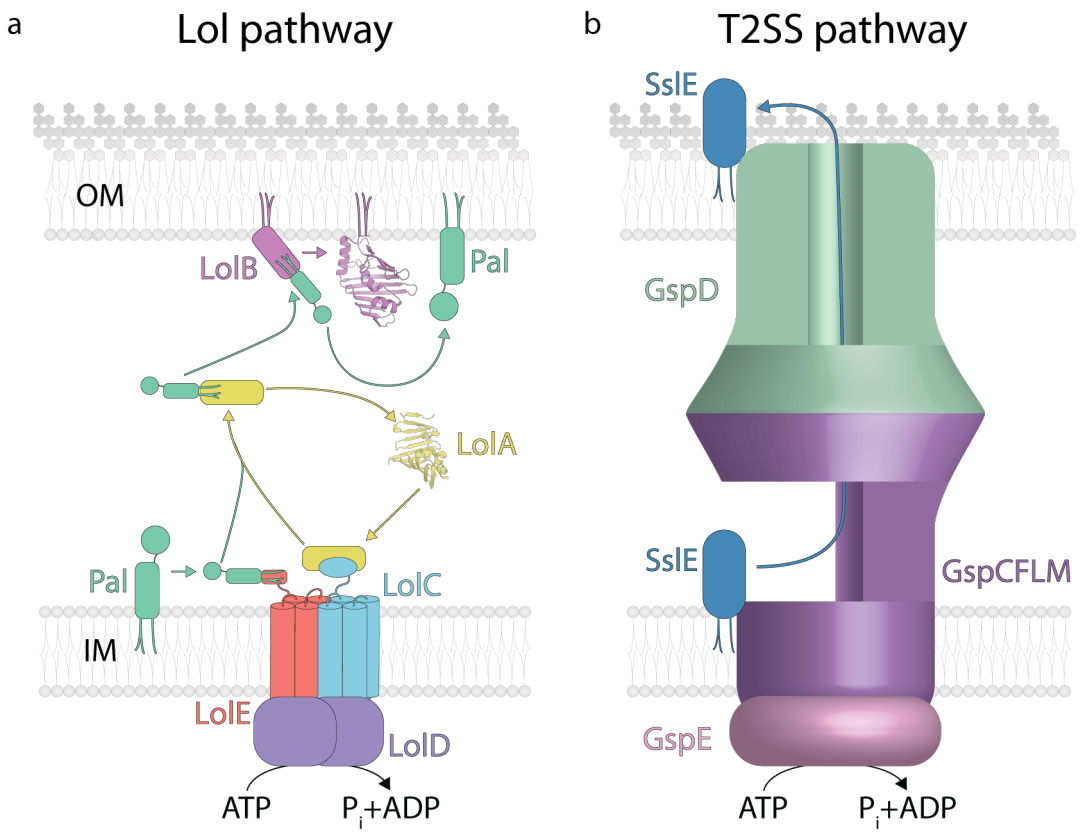


Figure 1.7.1 Outer membrane lipoprotein transfer pathways.

The majority of lipoproteins are transferred to the OM via the Lol pathway, but some utilise the T2SS machinery. Lipoproteins secreted by the T2SS are known to be surface-exposed, but without a lateral gate, it is unknown how the T2SS deposits lipoproteins into the OM (Zückert, 2014). **a,** The inner membrane complex recruits nascent lipoprotein and the periplasmic chaperone LolA (PDB: 1IML). Pal is then transferred to LolA, and LolB (PDB: 1IWM) inserts Pal into the OM after receiving the lipoprotein from LolA. Coloured arrows represent the movement of the similarly coloured subunit or substrate. **b,** SslE contains a Lol avoidance signal at its N-terminus: $Cys_{(+1)}$-$Asp_{(+2)}$-$Gly_{(+3)}$-$Gly_{(+4)}$, and utilises the T2SS for secretion into the LPS layer (shown) or excretion into the extracellular milieu (not shown) (Baldi *et al.*, 2012; DeCanio *et al.*, 2013). Although the exact mechanism for T2SS is unknown, a series of pseudopilin subunits are thought to form a platform comprising the major subunit GspG and the minor subunits GspIJKH that pushes a substrate to the surface (not shown) (Costa *et al.*, 2015).

secretion, but how the T2SS extracts lipoprotein from the IM (which is energetically costly) or shields its triacyl moiety during periplasmic transit remains unknown (Rondelet & Condemine, 2013). Interestingly, by swapping the signal sequence of PulA for the signal peptide of periplasmic MalE, PulA was still properly secreted by the T2SS machinery (Poquet *et al.*, 1993), making it difficult to define a role for its lipidation.

In addition to the monotopic lipoproteins already discussed, there are several interesting cases of polytopic lipoproteins (Figure 1.7.2). These lipoproteins are transported to the OM via the Lol pathway and are inserted into the OM in a BAM-dependent manner (NalP, RcsF or BamC) or via an unknown mechanism (e.g. bitopic species of Lpp or Pal) that may be BAM-independent (Konovalova & Silhavy, 2015; Wilson & Bernstein, 2016). Indeed, Dunstan *et al.* (2015) recently demonstrated that transmembrane lipoproteins with either α-helical barrels (Wza) or β-barrels (CsgG) do not require the BAM complex or the TAM for assembly.

The BAM-dependent polytopic lipoproteins are tethered either to (i) the OM periplasmic leaflet (NalP and BamC) or (ii) the outer leaflet (RcsF) by their lipid moieties, and are thought to be threaded through the pore of a β-barrel (Konovalova *et al.*, 2014; Oomen *et al.*, 2004; Roussel-Jazédé *et al.*, 2013; Webb *et al.*, 2012). The NalP lipoprotein is an autotransporter that assembles via the type 5 secretion system (T5SS) pathway, whereby it secretes its N-terminal passenger domain through its C-terminal β-barrel pore. While neither RcsF nor BamC contain their own transmembrane domain, RcsF spans the OM by threading its

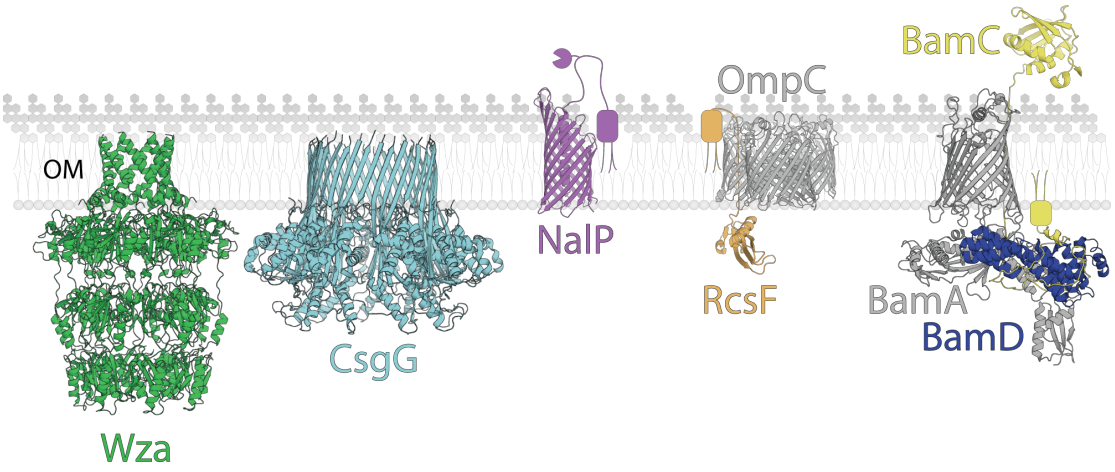


Figure 1.7.2 Polytopic lipoproteins.

Ribbon diagrams of various bitopic lipoproteins are shown. Wza (PDB: 2W8I) assembles into an 8-stranded octameric α-helical barrel (Dong *et al.*, 2006). CsgG is also membrane-spanning but it folds into a 36-stranded nonameric β-barrel. For clarity, the secondary structural data that was absent from the CsgG crystal structure file (PDB: 4UV3) was added as per Goyal *et al.* (2014). NalP is a lipoprotein autotransporter that assembles via T5SS to secrete its N-terminal extracellular domain through its C-terminal 12-stranded β-barrel transmembrane domain (PDB: 1UYN) (Leyton *et al.*, 2012). Although the orientation of the lipidated portion of NalP has not yet been determined, it is depicted in the LPS leaflet because NalP twice cleaves itself presumably extracellularly (Roussel-Jazédé *et al.*, 2013). RcsF is inserted into the LPS leaflet and spans the OM by threading its C-terminal periplasmic domain (PDB: 2Y1B) through the lumen of OmpC (PDB: 2J1N) (Konovalova *et al.*, 2014). Conversely, Webb *et al.* (2012) found that BamC has a C-terminal extracellular domain but its N-terminus is inserted into the periplasmic leaflet of the OM. BamC interacts with BamD in the periplasm before spanning the OM by presumably threading through BamA to display its C-terminal helix-grip domain (PDB: 3SNS) extracellularly. BamA, BamD and the N-terminal portion of BamC are also shown (PDB: 5D0O).

periplasmic domain through the lumen of OmpC or OmpF (Konovalova *et al.*, 2014) and, although not determined experimentally, BamC may similarly thread its extracellular helix-grip domain(s) through the lumen of BamA (Webb *et al.*, 2012).

1.8 Periplasmic quality control

Proteins that reach the periplasm are generally subject to various quality control mechanisms to ensure correct protein maturation. Periplasmic quality control proteins can broadly be categorised into proteases, folding catalysts and chaperones (Goemans *et al.*, 2014). The importance of these factors is underscored by the cell's stress response being activated by the misfolding and aggregation of periplasmic proteins (Kim, 2015; Raivio, 2014). Proteins that do misfold and aggregate in the periplasm are targeted for degradation by the dedicated protease, DegP (Ge *et al.*, 2014b). However, in contrast to the protease activity of DegP, periplasmic chaperones may protect folded proteins from denaturation by facilitating refolding (Thoma *et al.*, 2015) or sequester unfolded proteins via holdase activity either to promote their correct assembly (Hennecke *et al.*, 2005) or to prevent their misfolding (Burmann *et al.*, 2013).

Periplasmic chaperones may have dedicated substrates, such as the pilotins of T2SS, T3SS and T4SS (Burkinshaw & Strynadka, 2014; Dunstan *et al.*, 2013; Ilangovan *et al.*, 2015) or the fimbrial chaperones from the chaperone-usher family (Geibel & Waksman, 2014), while others have a much broader substrate range. In addition to LptA and LolA - which were respectively discussed in Sections 1.3 (LPS biogenesis) and 1.5 (lipoprotein biogenesis) - several general

(i.e. promiscuous) periplasmic chaperones involved in OMP biogenesis have been discovered: DegP, SurA, Skp, and FkpA (Goemans *et al.*, 2014) (Figure 1.8.1).

DegP has dual protease and chaperone activity at or below 37 °C (Figure 1.8.1a), but during heat-shock, DegP predominantly uses its protease activity to degrade protein aggregates (Rizzitello *et al.*, 2001; Skorko-Glonek *et al.*, 2007; Spiess *et al.*, 1999) (Figure 1.8.1b), although whether DegP indeed has chaperone-like activity has recently been questioned (Ge *et al.*, 2014b). OMP biogenesis is thought to rely primarily on SurA-mediated transport from the IM to the OM under normal conditions, whereas Skp and DegP - both of which may also bind periplasmic proteins - are thought to work together in an auxiliary OMP biogenesis pathway (Behrens *et al.*, 2001; Sklar *et al.*, 2007) (Figure 1.8.1a). The remaining periplasmic chaperones - FkpA, Spy, HdeA and HdeB - primarily function during the cell's stress response (Goemans *et al.*, 2014).

During heat-shock, FkpA is thought to replace SurA and Skp as the main OMP chaperone (Ge *et al.*, 2014a) (Figure 1.8.1b). Whereas Spy primarily binds periplasmic proteins during general OM stress (Stull *et al.*, 2016), it has also been implicated in LptD assembly when overexpressed in *E. coli* Δ*fkpA* Δ*skp* double knockout strains (Schwalm *et al.*, 2013). Under acid-induced stress, HdeA (at pH 2-3) and HdeB (at pH 4-5) protect chaperones, proteins and lipoproteins from denaturing (Dahl *et al.*, 2015; Hong *et al.*, 2005; Malki *et al.*, 2008; Zhang *et al.*, 2011). SurA and DegP are still involved in OMP biogenesis at low pH, but require the protection of HdeA to ensure they themselves do not misfold (Zhang *et al.*, 2011).

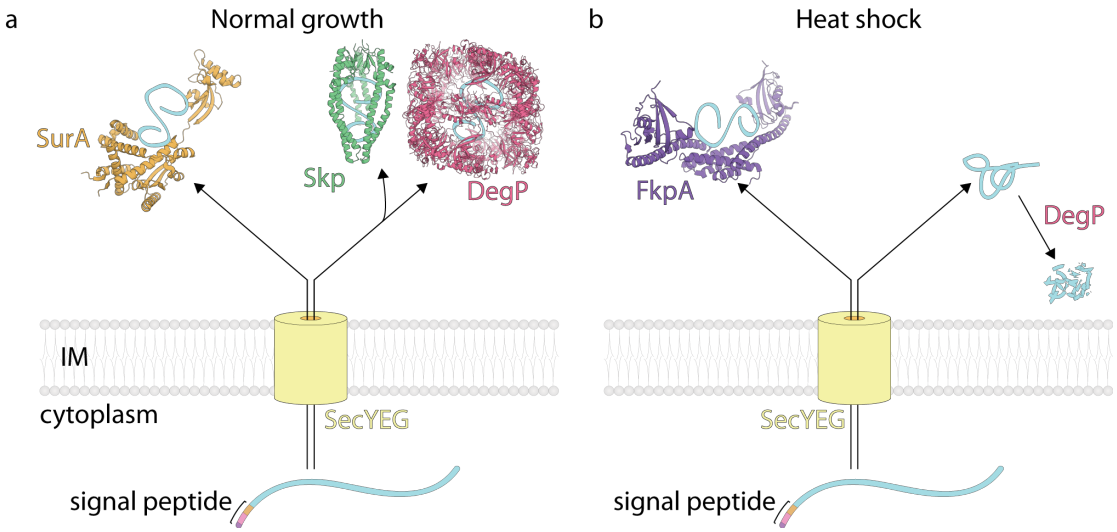


Figure 1.8.1 Periplasmic chaperones for outer membrane protein biogenesis.

Following Sec translocation, unfolded β-barrel proteins must traverse the aqueous and oxidising periplasm to the OM without misfolding and aggregating. As such, a number of periplasmic chaperones are recruited to ensure these unfolded polypeptides remain in a folding-competent state. **a,** During normal growth (37 °C), nascent β-barrel proteins (blue) are transferred to the OM usually via the SurA (PDB: 1M5Y) pathway, but otherwise by the Skp/DegP pathway. Although it is not clear how these two pathways interact, or how Skp (PDB: 1SG2) interacts with DegP (PDB: 2ZLE), these pathways are mostly redundant, where OMP biogenesis is generally still possible in the absence of either pathway (Sklar *et al.*, 2007). **b,** During heat-shock (44 °C), FkpA is the preferred periplasmic chaperone, whereas SurA and Skp only play minor roles during OMP biogenesis (Ge *et al.*, 2014a). During heat shock, nascent β-barrel proteins are more prone to misfolding and aggregating, so DegP's contribution to OMP biogenesis is mainly by proteolytically degrading these aggregates (Spiess *et al.*, 1999). Dimeric FkpA was generated from the monomeric structure (PDB: 1Q6U) using the "symmetry mates" function of Pymol.

The third class of periplasmic quality control proteins is the periplasmic folding catalysts, which may be further divided into: the Dsb (disulphide bond forming) proteins and peptidyl-prolyl *cis-trans* isomerases (PPIases) (Goemans *et al.*, 2014). DsbA catalyses the formation of cystine (i.e. cystine is the oxidised, disulfide-bonded form of two cysteine molecules), but problems may arise if DsbA substrates contain more than two cysteine residues (Figure 1.8.2a). DsbA has a tendency to oxidise consecutive cysteine residues and can potentially produce non-native cystine species (Kadokura *et al.*, 2004). To overcome this problem, the disulphide bond isomerase, DsbC, rearranges the disulphide bonds to ensure native cystine is formed, although how it determines what is native cystine is unknown (Berkmen, 2012). Recently, a novel folding catalyst was identified, BepA, that was found to selectively act as a disulphide bond isomerase for LptD (Narita *et al.*, 2013). Interestingly, BepA was also shown to harbour independent protease activity, where it was capable of proteolytically degrading LptD or BamA (Narita *et al.*, 2013).

In the highly oxidative periplasmic environment, cysteines with a free thiol group, are susceptible to oxidation and may be converted to highly unstable sulphenic acid species (Dutton *et al.*, 2008). Sulphenic acids readily react with free thiols to form non-native cystine or are irreversibly oxidised to sulphinic or sulphonic acid species (Reddie & Carroll, 2008). Although these reactive oxygen species may be targeted for degradation by DegP, DsbG is mostly responsible for reducing sulphenic acids into the native thiol form (Depuydt *et al.*, 2009), and DsbC is involved in the reduction of subsequent non-native cystines (Denoncin *et al.*, 2014) (Figure 1.8.2b).

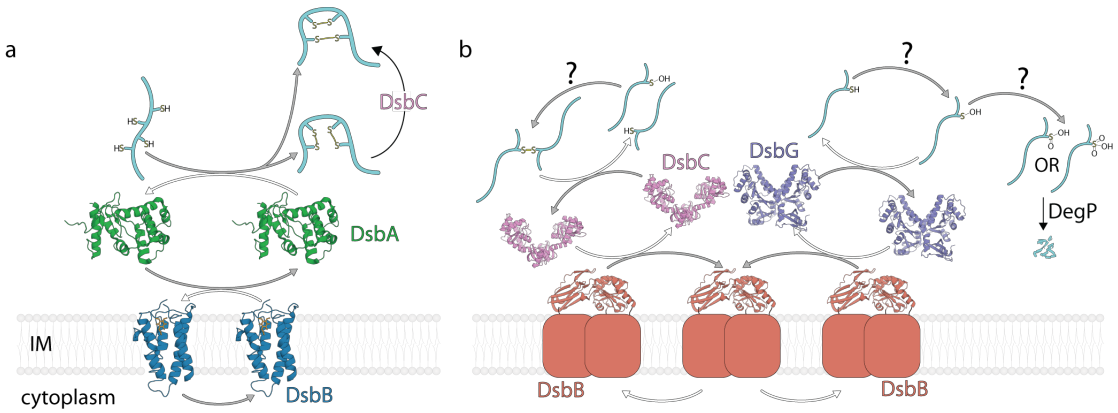


Figure 1.8.2 Redox pathways of cysteinyl thiols.

Arrows used to indicate oxidative pathways are coloured grey and those involving reduction are coloured white. The names of each Dsb protein are indicated next to the "active" (oxidised or reduced) form of that protein. DsbA (PDB: 2ZUP) is maintained in an "active" (oxidised) form by DsbB (also PDB: 2ZUP), which in turn is kept oxidised by ubiquinone (brown) or menaquinone (not shown) generated during aerobic or anaerobic metabolism (not shown), respectively (Bader *et al.*, 2000). DsbC (PDB: 1EEJ) and DsbG (PDB: 1V57) are likewise kept in an "active" reduced conformation by DsbD (PDB: 1Z5Y), which in turn is similarly maintained in a reduced form by cytoplasmic thioredoxin (not shown) (Berkmen, 2012). **a,** Following Sec translocation, DsbA oxidises nascent polypeptide; however, if more than two cysteines are present, incorrect cysteine pairs may be oxidised. DsbC then ensures correct cystine formation, but it is unknown whether DsbC recruits DsbA to re-oxidise the polypeptide (Berkmen, 2012). **b,** The oxidative periplasm (indicated by a question mark) may cause proteins with an odd number of cysteines to oxidise into sulphenic acids, which readily undergo further oxidation into cystine, sulphinic acids and sulphonic acids. DsbC and DsbG are recruited to reduce these incorrect cystines and sulphenic acids, respectively, or are otherwise degraded by DegP.

The second group of folding catalysts are the PPIases, which are important for ensuring prolyl peptide bonds adopt a *trans* conformation (Reimer *et al.*, 1998; Schmidpeter & Schmid, 2015) (Figure 1.8.3). While the equilibrium between *cis* and *trans* isomers for the majority of standard peptide bonds (380/400) greatly favours the *trans* isomer - resulting in 99.97% of these peptide bonds adopting the *trans* conformation in proteins (Jabs *et al.*, 1999) - the equilibrium involving the isomerisation of the peptide bond between a prolyl nitrogen and any other residue only slightly favours the *trans* isomer (Reimer *et al.*, 1998). PPIases are therefore required to push the equilibrium toward the *trans* isomer to ensure about 94.8 % of prolyl peptide bonds adopt the *trans* conformation (Reimer *et al.*, 1998; Schmidpeter & Schmid, 2015).

There are four *E. coli* proteins with periplasmic PPIase domains: SurA, FkpA, PpiA and PpiD, but a quadruple Δ*surA* Δ*fkpA* Δ*ppiA* Δ*ppiD* mutant remains viable, leading Justice *et al.* (2005) to conclude that PPIase activity is dispensable for viability. Indeed, because cytoplasmic PPIases are still expressed, such as trigger factor or SlyD (Schmidpeter & Schmid, 2015), it is possible that periplasmic PPIase activity is functionally redundant. Perhaps the periplasmic PPIases form a "backup" isomerisation pathway responsible for isomerising any peptidyl-prolyl *cis* isomers that were missed in the cytoplasm, or have since reverted to a *cis* isomer. Despite this, the functional consequences for an extracytoplasmic protein containing one or more peptidyl-prolyl *cis* isomers have not been thoroughly investigated, indicating that a quadruple periplasmic PPIase mutant may remain viable because PPIase substrates nevertheless remain functional.

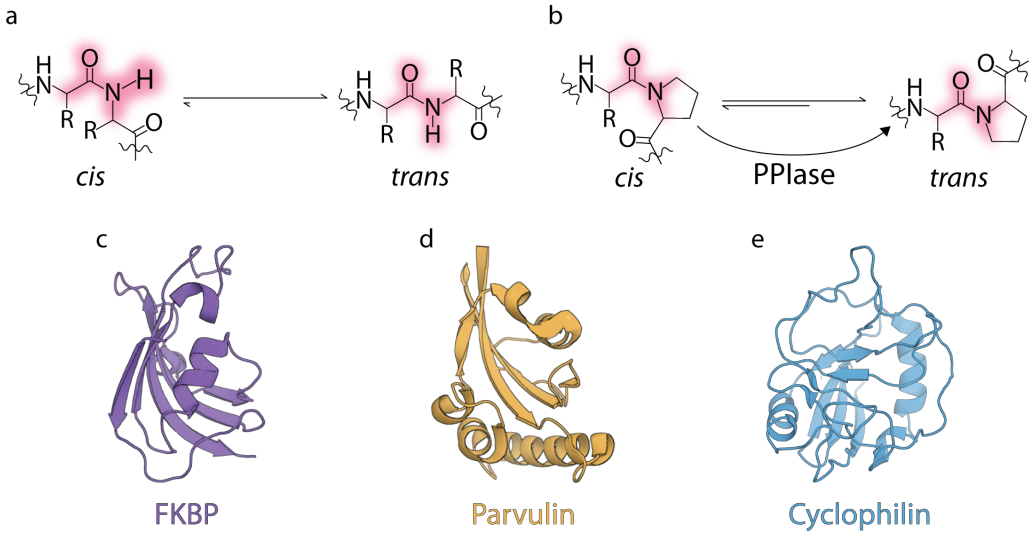


Figure 1.8.3 Peptide bond isomerisation.

The planar peptide bonds (and atoms that exist in the plane) are highlighted in pink and may adopt a *cis*- or *trans*-conformation. **a,** The majority of peptide bonds heavily favour the *trans*-isomer, except **b,** those involving proline, where the equilibrium between *trans*- and *cis*-isomers is almost equal. PPIases are recruited to ensure proline peptide bonds adopt the *trans*-isomer. *E. coli* contains four periplasmic PPIases, representing the three major PPIase families. **c,** Shown here is a ribbon diagram of the FKBP PPIase domain from FkpA (PDB: 1Q6U), corresponding to residues 123-225 of the mature protein. **d,** PpiD and SurA both contain the parvulin PPIase domain; the second parvulin domain from SurA (PDB: 1M5Y), corresponding to residues 262-365 of the mature protein, is depicted here. **e,** Shown is a ribbon diagram of PpiA (PDB: 1CLH), which is entirely comprised of the cyclophilin PPIase domain.

1.9 β-Barrel protein biogenesis

Omp85 proteins are a structurally conserved superfamily present in all domains of life. They adopt a β-barrel conformation and usually have at least one polypeptide transport associated (POTRA) domain, which are important for binding substrates and/or auxiliary proteins (Heinz & Lithgow, 2014; Simmerman *et al.*, 2014). Functionally, the Omp85 superfamily may be split into (i) translocators and (ii) β-barrel insertases: the bacterial archetypes of which are FhaC and BamA, respectively (Heinz & Lithgow, 2014). FhaC contains two POTRA domains and forms the translocating component of a two-partner secretion system where it is responsible for assembling and secreting FHA via the T5SS (Leo *et al.*, 2012) (Figure 1.9.1).

BamA is essential for cell viability and is the main component of the BAM complex. In addition to BamA, the BAM complex is comprised of four lipoproteins: BamB, BamC, BamD and BamE, of which, BamD is essential, and BamC is bitopic (Noinaj *et al.*, 2015; Webb *et al.*, 2012) (Figure 1.9.2). Although the precise interaction network between the BAM complex's components, periplasmic chaperones and substrate proteins is still under investigation, it is generally accepted that the BAM complex accepts substrate β-barrel proteins from periplasmic chaperones - such as Skp and SurA - and catalyses their insertion into the OM. To date, the only known exception - but likely not the only exception - is CsgG (Dunstan *et al.*, 2015): a 36-stranded β-barrel comprised of nine lipoprotein monomers (Cao *et al.*, 2014; Goyal *et al.*, 2014), which has an as-yet-undetermined mode of assembly.

Omp85-family translocators and insertases each contain similar but distinct

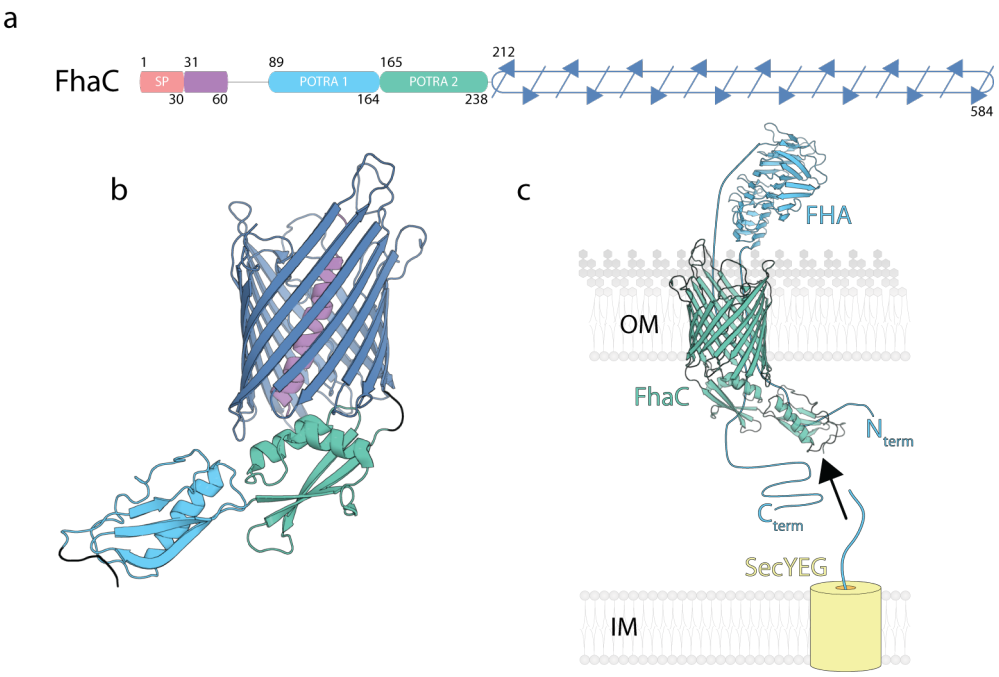


Figure 1.9.1 FhaC structure and mechanism.

a, Structural map of FhaC, based on its crystal structure (PDB: 2QDZ), where the POTRA domain borders are defined as per Clantin *et al.* (2007). **b,** Ribbon diagram of FhaC coloured as in "**a**". **c,** The putative Omp85 translocation mechanism whereby FhaC (now coloured green for clarity) secretes its partner protein, FHA (PDB: 1RWR). The N-terminal domain of FHA (indicated as N_{term}) interacts with POTRA domain 1 while FHA threads through the barrel lumen and assembles into a β-helical structure in the extracellular milieu (Clantin *et al.*, 2007). Once the C-terminus of FHA (indicated as C_{term}) threads through the lumen, FHA dissociates from POTRA domain 1 (Clantin *et al.*, 2007).

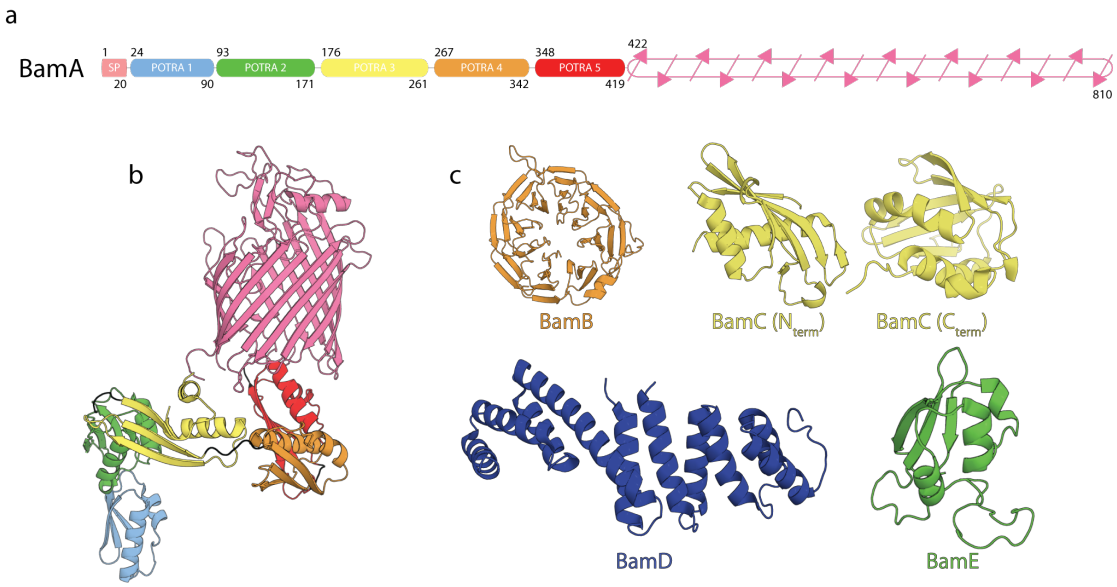


Figure 1.9.2 The BAM complex.

a, Structural map of BamA based on its crystal structure (PDB: 5D0O), where the POTRA domain borders are defined as per Selkrig *et al.* (2015). **b,** Ribbon diagram of BamA, coloured as in "**a**". **c,** Ribbon diagrams of the BAM lipoproteins, where the similar colours between the POTRA domains and lipoproteins are coincidental: BamB (PDB: 3Q7O), BamC (PDB: 2YH6 and 3SNS for its N- and C-terminal globular domains, respectively), BamD (PDB: 3Q5M) and BamE (PDB: 2KXX).

transmembrane domains (Heinz & Lithgow, 2014), including three subtle structural differences that enable BamA to act as an insertase. Firstly, unlike Omp85 translocators, which contain an "open" exit pore, the exit pore of Omp85 insertases is capped by their extracellular loop 6, suggesting a lateral transfer mechanism (Noinaj *et al.*, 2015). Secondly, there is a significant difference in the size of the aromatic girdle encircling BamA (9-20 Å) compared to FhaC (16-19 Å) (Noinaj *et al.*, 2013). Based on the observation that the aromatic girdle is positioned between the "head" and "tail" groups of LPS or phospholipid species (Schulz, 2002), Noinaj *et al.* (2013) determined that the membrane near the short side of BamA's aromatic girdle, when compared to the long side, was 16 Å thinner and contained a threefold increase in lipid disorder. This structural difference is thought to facilitate OMP insertion by destabilising the protein-lipid interface to lower the activation energy required for protein insertion.

Thirdly, the strong hydrogen-bonding interactions between the first and last β-strands in FhaC is greatly diminished in BamA, enabling the opening and closing of what has now been termed BamA's "lateral gate" (Noinaj *et al.*, 2014; Noinaj *et al.*, 2013). Following lateral gate opening, Omp85 insertases are thought to use their terminal β-strand to act as a template for substrate OMP β-strand assembly, a process termed "β-augmentation" (Estrada Mallarino *et al.*, 2015; Noinaj *et al.*, 2014; Noinaj *et al.*, 2015). The current working hypothesis is that substrate OMPs partially enter the lumen of BamA, where extracellular domains may translocate across the OM through the hydrophilic pore, whereas the transmembrane portion directly enters the disordered regions of the lipid bilayer and assembles via β-augmentation (Noinaj *et al.*, 2014) (Figure 1.9.3a).

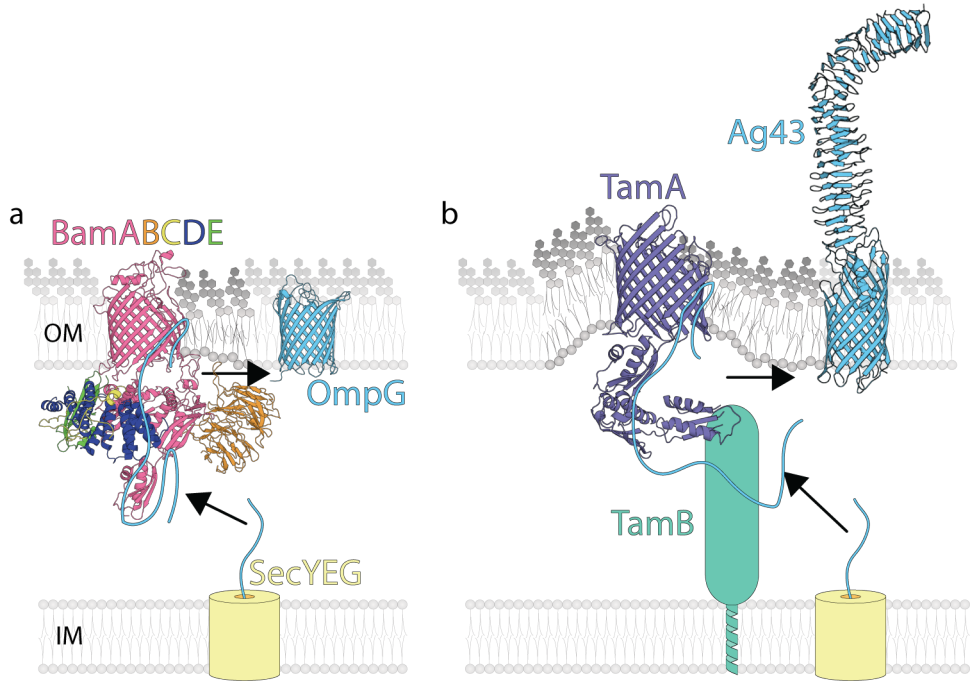


Figure 1.9.3 Insertion mechanisms of the BAM complex and the TAM.

Shown here are the putative mechanisms by which **a,** the BAM complex (PDB: 5D0O) and **b,** the TAM (PDB: 4C00 for TamA) assemble their substrates. The lopsided aromatic girdle of BamA or TamA destabilises the lipid-protein interface (darker regions of the bilayer) to enhance OMP insertion. In both examples, nascent OMP translocates across the IM via the Sec pathway and is transferred to the β-barrel insertase via periplasmic chaperones (not shown). **a,** In this example, assembly of OmpG (PDB: 2X9K) is shown. **b,** The TAM acts to further destabilise the lipid bilayer where the rigid POTRA domains of TamA are thought to push the OM away from the IM (Selkrig *et al.*, 2015; Shen *et al.*, 2014). In this example, assembly of Ag43 is shown. The structure of Ag43 is comprised of its N-terminal passenger domain (PDB: 4KH3) and a Phyre2 homology model of residues 705-1039 modelled after the same domain from AIDA (PDB: 4MEE).

Interestingly, recent molecular dynamics simulations revealed that for BamA in isolation, the "lateral gate" does not open and despite numerous interactions between the POTRA domain and the OM, the authors did not report significant disruptions to the lipid interface (Fleming *et al.*, 2016). Considering that a functional and efficient BAM complex does not consist of BamA alone, the BAM lipoproteins are likely required for lateral gate opening/closing and to facilitate OM disruption, where at least one BAM lipoprotein - BamE - has been shown to specifically bind PG, presumably to enhance substrate insertion (Endo *et al.*, 2011; Knowles *et al.*, 2011).

A second β-barrel assembling machine related to the BAM complex, is the TAM. The TAM is comprised of (i) TamB (formerly known as YtfN), which contains a large periplasmic β-helical subunit anchored to the IM, and; (ii) an Omp85 component harbouring three POTRA domains, TamA (formerly known as YtfM) (Figure 1.9.4). Although each of the BAM components are separately encoded within *E. coli* (Keseler *et al.*, 2013), a 4-basepair overlap exists between *tamA* and *tamB* (Selkrig *et al.*, 2012). This overlap suggests a potentially important transcriptional coupling process whereby TamA and TamB are recruited by the same (or neighbouring) SecYEG complexes to ensure both subunits are always in relatively close proximity for TAM assembly.

While the TAM is dispensable for viability *in vitro*, both subunits are required for proper TAM function *in vivo* (Selkrig *et al.*, 2012). A structural comparison of BamA with TamA revealed the presence of a lateral gate and a lopsided aromatic girdle, indicating they may share a similar insertion mechanism (Gruss *et al.*, 2013;

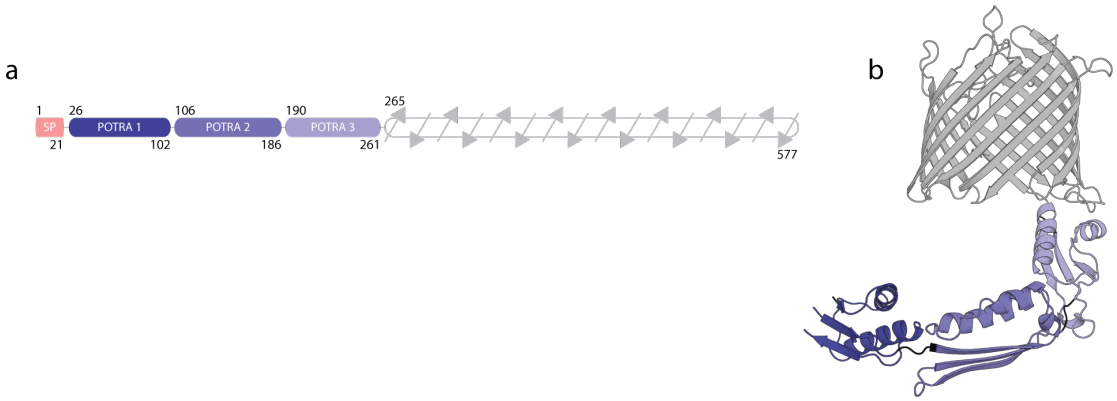


Figure 1.9.4 POTRA domains of TamA.

a, Structural map of TamA based on its crystal structure (PDB: 4C00), where the POTRA domain borders are defined as per Selkrig *et al.* (2015). **b,** Ribbon diagram of TamA, coloured as in "**a**".

Selkrig *et al.*, 2015). Intriguingly however, the flexibility observed in BamA's POTRA domains is not reflected in TamA, where TamA's POTRA domains instead form a rigid body that pushes onto TamB to further destabilise the OM, thereby promoting OMP insertion (Selkrig *et al.*, 2015; Shen *et al.*, 2014) (Figure 1.9.3b). Additionally, the exit pore "capping" performed by loop 6 in TamA is further stabilised in BamA by loops 4 and 7 (Noinaj *et al.*, 2015), perhaps indicating that exit pore "opening" during lateral transfer is relatively simpler than in BamA.

The BAM complex has been suggested to recognise a "β-signal motif" found within the last β-strand of its substrate, because its homologue in mitochondria, Sam50, does (Höhr *et al.*, 2015). This motif is reported to include a terminal phenylalanine or tryptophan and usually a tyrosine at the third-last residue (Paramasivam *et al.*, 2012; Robert *et al.*, 2006; Struyve *et al.*, 1991). Several additional studies have also supported this hypothesis by showing there is a reduction in OMP assembly if the β-signal motif is removed (de Cock *et al.*, 1997; Gessmann *et al.*, 2014; Lehr *et al.*, 2010; Walther *et al.*, 2009): PhoE (GMTYQF), YadA (SFNIEW), OmpA (GVSYRF) and OmpC (GLVYQF), where the last six residues reported to comprise the terminal β-strand are indicated in brackets.

However, it is hard to reconcile with the β-signal hypothesis following crystal structure inspection of several OMPs, which have no obvious β-signal within their terminal β-strand (Gruss *et al.*, 2013; Koronakis *et al.*, 2000; Phan *et al.*, 2011; Remaut *et al.*, 2008; Snijder *et al.*, 1999): OmpLA (VGVGVM), FimD (SGGVLA), PapC (ASGGAT), TamA (QFYIGL) and TolC (SLPIYQ). If the β-signal does not need to be within the last β-strand, then FimD's sixteenth β-strand may contain that β-

signal (GLNTAF). But the potential redundancy in this "signal" is high, and this "match" could be coincidental: considering that a similar motif may be found in the penultimate β-strand of CsgG (AGVFRF) (Cao *et al.*, 2014; Goyal *et al.*, 2014), and yet CsgG has been demonstrated not to require the BAM complex (Dunstan *et al.*, 2015). Indeed, transmembrane β-strands are already characterised by the presence of alternating hydrophobic and polar residues that are capped by aromatic residues constituting the aromatic girdle (Jackups & Liang, 2005; Schulz, 2002), so the "β-signal" may instead be a general structural feature.

The logic used to claim a protein is BAM-dependent based on the presence of an apparent β-signal could be considered a form of non sequitur, especially considering that TolC and TamA, which are both BAM-dependent, do not harbour the motif (Bennion *et al.*, 2010; Dunstan *et al.*, 2015; Malinverni *et al.*, 2006; Werner & Misra, 2005). Furthermore, if an unbiased approach was taken, after inspecting the structures of several OMPs that reportedly contain the β-signal motif - including PhoE, YadA, OmpC, and most OmpA structures - it appears that the terminal phenylalanine is not part of the terminal β-strand (Arora *et al.*, 2001; Basle *et al.*, 2006; Cowan *et al.*, 1992; Johansson *et al.*, 2007; Lou *et al.*, 2011; Pautsch & Schulz, 2000; Shahid *et al.*, 2012). While it is entirely possible that the BAM complex does recognise a signal, rather than an ability to recognise the general structural characteristic of transmembrane β-strands, more experimental evidence is required.

Unlike the BAM complex, the TAM has remained enigmatic in that only a single family of proteins have been shown to assembled by the TAM: the AIDA-like

autotransporters (Selkrig *et al.*, 2012) (Figure 1.9.3b). As such, it is difficult to identify any sequence or structural motifs within TAM substrates that may cause OMP biogenesis using the TAM rather than the BAM complex. A recent study by Heinz *et al.* (2015) investigating the evolutionary history of the TAM revealed that while TamA is found almost exclusively in *Proteobacteria* (see also Heinz & Lithgow, 2014), TamB is instead distributed among most Gram-negative lineages and is potentially under positive selection (Xu *et al.*, 2011). Indeed, Heinz *et al.* (2015) reported that the original TAM and BAM complex may not have been separate. Rather, the original β-barrel assembling machine may have consisted of an inner membrane TamB-like component that lead to the TAM, and an outer membrane BamA-like component that lead to the BAM complex. Even today, the *Spirochaetes* appear to harbour a similar *bamA-tamB* operon (Selkrig *et al.*, 2014).

The reason why two β-barrel assembling machines have been simultaneously selected for in *Proteobacteria* is baffling. Surely one machine would be sufficient to assemble all β-barrel proteins? Given the prevalence of the TAM, there must be something that the TAM does that the BAM complex cannot do. Because the AIDA-like autotransporters are the only known TAM substrates (Selkrig *et al.*, 2012), and about 40% of *Proteobacteria* do not contain autotransporters (Celik *et al.*, 2012), it is difficult to define a role for the TAM in the context of that simple evolutionary purpose. By searching for a more substantive and diverse list of TAM substrates, it was thought that questions pertaining to the co-evolution of the BAM complex and the TAM, and whether they are two competing machines or act in synergy may be answered.

Chapter 2 – Materials and Methods

2.1 Bacterial strains and growth conditions

The *E. coli* strains used in this study are listed in Table 2.1.1. In several cases, *E. coli* DH5α and BL21 Star™ (DE3) (and its derivatives) were transformed to harbour one or two plasmids as listed in Table 2.1.2. *E. coli* were incubated in 2× YT media (CSH Protocols, 2014) or SOC media (CSH Protocols, 2012), but usually and unless otherwise stated, *E. coli* were incubated in lysogeny broth (LB) media (Lennox formulation) (CSH Protocols, 2006a). Incubation media supplemented with carbohydrates was filter sterilised using Stericups® with 0.22 μm Express™ Plus membranes (Merck Millipore) and media with low volumes (less than 50 mL) was filter sterilised using an Acrodisc® syringe filter with a 0.2 μm Supor® membrane (PALL Life Sciences). Otherwise, media was sterilised by autoclaving at 121 °C for 20 minutes at 15 psi. When required, growth media was supplemented with antibiotics at the following concentrations: 100 $\mu g.mL^{-1}$ ampicillin, 30 $\mu g.mL^{-1}$ kanamycin, 34 $\mu g.mL^{-1}$ chloramphenicol. Where solid media was required, 15 $g.L^{-1}$ agar (Merck Millipore, 12177) was added to LB growth media prior to autoclaving.

Unless otherwise indicated, liquid cultures of *E. coli* were incubated at 37 °C on an orbital platform rotating 200 strokes per minute (25 mm orbit). Except where stated otherwise, *E. coli* were routinely incubated on solid LB media for 16-24 hours at 37 °C. Two-to-three colonies were then selected and used to inoculate fresh LB (5-10 mL liquid) media where cells were subsequently incubated for 12-16 hours. This saturated culture was typically used to prepare glycerol stock solutions or to inoculate fresh media. To estimate the growth phase of *E. coli*, the

Table 2.1.1 *E. coli* strains used in this study.

***E. coli* strains**	**Characteristics**	**Purpose**	**Source or Reference**
DH5α	F⁻ *endA1 glnV44 thi-1 recA1 relA1 gyrA96 deoR nupG purB20* φ80d*lacZ*ΔM15 Δ(*lacZYA-argF*)U169, hsdR17(*rK⁻ mK⁺*), λ⁻	PCR template and vector storage	Invitrogen
K-12 strain MG1655	F⁻ λ⁻ *ilvG⁻ rfb-50 rph-1*	PCR template	Bachmann (1996)
CFT073	Uropathogenic *E. coli* isolated from the blood and urine of a woman with acute pyelonephritis	PCR template	Mobley *et al.* (1990)
E2348/69	Enteropathogenic *E. coli* isolated in Taunton, United Kingdom, in 1969 during an outbreak of infantile diarrhoea	PCR template	Levine *et al.* (1978)
BL21	F⁻ *ompT gal dcm lon hsdSB*(*rB⁻mB⁻*) [*malB⁺*]$_{K-12}$(λ^{S})	gene expression	Invitrogen
BL21 (DE3)	BL21 λ(DE3)	gene expression	Invitrogen
BL21 Star™ (DE3)	BL21 (DE3) *rne131*	gene expression	Invitrogen
BL21 Star™ (DE3) Δ*tamA*	BL21 Star™ (DE3) Δ*tamA*::*nptII* (Kanamycin resistant)	gene expression	Made by M. J. Belousoff, Heinz *et al.* (2016), and Stubenrauch *et al.* (2016)
BL21 Star™ (DE3) Δ*tamB*	BL21 Star™ (DE3) Δ*tamB*::*nptII* (Kanamycin resistant)	gene expression	Made by M. J. Belousoff, Stubenrauch *et al.* (2016)
BL21 Star™ (DE3) Δ*bamB*	BL21 Star™ (DE3) Δ*bamB*::*nptII* (Kanamycin resistant)	gene expression	Made by M. J. Belousoff, Stubenrauch *et al.* (2016)
BL21 Star™ (DE3) Δ*bamC*	BL21 Star™ (DE3) Δ*bamC*::*nptII* (Kanamycin resistant)	gene expression	Made by M. J. Belousoff, Stubenrauch *et al.* (2016)
BL21 Star™ (DE3) Δ*bamE*	BL21 Star™ (DE3) Δ*bamE*::*nptII* (Kanamycin resistant)	gene expression	Made by M. J. Belousoff, Stubenrauch *et al.* (2016)

Table 2.1.2 *E. coli* plasmids used in this study.

Name	**Characteristics**	**Purpose**	**Source or Reference**
pET-15b	Base *E. coli* expression vector, pBR322 origin of replication, confers resistance to ampicillin	Base vector	Novagen
pET-22b(+)	Base *E. coli* expression vector, pBR322 origin of replication, confers resistance to ampicillin	Base vector	Novagen
pETDuet-1	Base *E. coli* expression vector, pBR322 origin of replication, confers resistance to ampicillin	Base vector	Novagen

pACYCDuet-1	Base *E. coli* expression vector, p15A origin of replication, confers resistance to chloramphenicol	Base vector	Novagen
pKS02	pET-15b (*Nco*I/*Nde*I) Ω MB45/MB46 PCR product from *E. coli* K-12 MG1655 gDNA template (*Nco*I/*Nde*I; 2639 bp), confers ampicillin resistance	*fimD* expression vector	Made by K. S. Tan, Stubenrauch *et al.* (2016)
pKS06	pET-15b (*Nco*I/*Nde*I) Ω MB53/MB54 PCR product from *E. coli* K-12 MG1655 gDNA template (*Nco*I/*Nde*I; 1736 bp), confers ampicillin resistance	PCR template	Made by K. S. Tan, unpublished
pKS07	pET-15b (*Nco*I/*Xho*I) Ω MB55/MB56 PCR product from *E. coli* K-12 MG1655 gDNA template (*Nco*I/*Xho*I; 1057 bp), confers ampicillin resistance	*phoE* expression vector	Made by K. S. Tan, unpublished
pMB11	pET-15b (*Nco*I/*Nde*I) Ω MB35/MB36 PCR product from *E. coli* K-12 MG1655 gDNA template (*Nco*I/*Nde*I; 1484 bp), confers ampicillin resistance	*tolC* expression vector	Made by M. J. Belousoff, Stubenrauch *et al.* (2016)
pWAK05	TEV-*fimD*-C_{strep} expression vector, RSF1010 origin, confers kanamycin resistance. FimD harbours an internal TEV site (within loop 7 of the transmembrane domain) and a C-terminal *strep*-II tag	PCR template	Made by J. Lillington, Stubenrauch *et al.* (2016)
pCJS29	pET-15b (*Nco*I/*Nde*I) Ω MB45/CSP37 PCR product from pWAK05 DNA template (*Nco*I/*Nde*I; 2669 bp), confers ampicillin resistance	*TEV-fimD-*C_{strep} expression vector	This study, Stubenrauch *et al.* (2016)
pCJS30	pET-15b(*Nco*I/*Nde*I) Ω CSP32/CSP36 PCR product using *E. coli* E2348/69 gDNA template (*Nco*I/*Nde*I; 2819 bp), confers resistance to ampicillin	*eaeA* expression vector	This study, Heinz *et al.* (2016)
pCJS31	pET-15b (*Nco*I/*Bam*HI) Ω CSP132/CSP47 PCR product from *E. coli* CFT073 gDNA template (*Nco*I/*Bam*HI; 2521 bp), confers ampicillin resistance	*yqiG* expression vector	This study
pCJS32	pET-22b(+) (*Nde*I/*Bam*HI) Ω CSP133/CSP51 PCR product from *E. coli* CFT073 gDNA template (*Nde*I/*Bam*HI; 2520 bp), confers ampicillin resistance	*papC* expression vector	This study, Stubenrauch *et al.* (2016)
pCJS33	pET-15b (*Nco*I/*Bam*HI) Ω CSP134/CSP55 PCR product from *E. coli* CFT073 gDNA template (*Nco*I/*Bam*HI; 2653 bp), confers ampicillin resistance	*yfcU* expression vector	This study, Stubenrauch *et al.* (2016)
pCJS37	pETDuet-1 (*Aat*II/*Xho*I) Ω CSP142/CSP143 PCR product from *E. coli* E2348/69 gDNA template (*Aat*II/*Xho*I; 2528 bp), confers ampicillin resistance	*ecpC* expression vector	This study
pCJS40	pET-15b (*Nco*I/*Bam*HI) Ω CSP146/CSP91 PCR product from *E. coli* K-12 MG1655 gDNA template (*Nco*I/*Bam*HI; 2446 bp), confers ampicillin resistance	*ybgQ* expression vector	This study, Stubenrauch *et al.* (2016)

pCJS42	pETDuet-1 (*Nco*I/*Bam*HI) Ω MB47/CSP119 PCR product from *E. coli* K-12 MG1655 gDNA template (*Nco*I/*Bam*HI; 2356 bp), confers ampicillin resistance	*lptD* expression vector	This study
pCJS43	pETDuet-1 (*Nde*I/*Xho*I) Ω CSP121/CSP122 PCR product from *E. coli* K-12 MG1655 gDNA template (*Nde*I/*Xho*I; 582 bp), confers ampicillin resistance	*lptE* expression vector	This study
pCJS44	pCJS43 (*Nco*I/*Bam*HI) Ω pCJS42 (*Nco*I/*Bam*HI; 2356 bp), confers ampicillin resistance	Dual *lptD* (MCS1) and *lptE* (MCS2) expression vector	This study
pCJS49	pET-15b (*Nco*I/*Bam*HI) Ω CSP141/CSP75 PCR product from *E. coli* CFT073 gDNA template (*Nco*I/*Bam*HI; 2587 bp), confers ampicillin resistance	*htrE* expression vector	This study, Stubenrauch *et al.* (2016)
pCJS50	pET-15b (*Nco*I/*Bam*HI) Ω CSP152/CSP107 PCR product from *E. coli* CFT073 gDNA template (*Nco*I/*Bam*HI; 4249 bp), confers ampicillin resistance	*fdeC* expression vector	This study, Heinz *et al.* (2016)
pCJS51	pET-22b(+) (*Nco*I/*Eco*RI) Ω CSP155/CSP156 PCR product from pKS02 DNA template (*Nco*I/*Eco*RI; 2533 bp), confers ampicillin resistance. FimD contains the pET-22b(+)-based *pelB* signal sequence followed by an N-terminal *strep*-II tag instead of its own 45-residue signal sequence.	N_{strep}-*fimD* expression vector	This study, Stubenrauch *et al.* (2016)
pCJS52	pETDuet-1 (*Nco*I/*Eco*RI) Ω MB45/CSP156 PCR product from pKS02 DNA template (*Nco*I/*Eco*RI; 2638 bp), confers ampicillin resistance	*fimD* expression vector	This study
pCJS58	pCJS52 (*Nde*I/*Xho*I) Ω CSP158/CSP159 PCR product from pKS06 DNA template (*Nde*I/*Xho*I; 1734 bp), confers ampicillin resistance	Dual *fimD* (MCS1) and *tamA* (MCS2) expression vector	This study
pCJS60	pCJS52 (*Nde*I/*Xho*I) Ω CSP162/CSP163 PCR product from *E. coli* E2348/69 DNA template (*Nde*I/*Xho*I; 1179 bp), confers ampicillin resistance	Dual *fimD* (MCS1) and *bamB* (MCS2) expression vector	This study
pCJS64	pET-15b (*Nco*I/*Nde*I) Ω MB45/CSP172 PCR product from pKS02 DNA template (*Nco*I/*Bsp*EI; 559 bp) Ω CSP173/MB46 PCR product from pKS02 DNA template (*Bsp*EI/*Nde*I; 2100 bp), confers ampicillin resistance. FimD harbours a hexahistidine-tag between P180 and G181 (P135/G136 of mature FimD)	$His_{135/136}$-*fimD* expression vector	This study

pCJS65	pET-15b (*Nco*I/*Nde*I) Ω MB45/CSP168 PCR product from pKS02 DNA template (*Nco*I/*Bsp*EI; 787 bp) Ω CSP169/MB46 PCR product from pKS02 DNA template (*Bsp*EI/*Nde*I; 1872 bp), confers ampicillin resistance. FimD harbours a hexahistidine-tag between P256 and G257 (P211/L212 of mature FimD)	*His*$_{211/212}$-*fimD* expression vector	This study, Stubenrauch *et al.* (2016)
pCJS66	pET-15b (*Nco*I/*Nde*I) Ω MB45/CSP170 PCR product from pKS02 DNA template (*Nco*I/*Bsp*EI; 1843 bp) Ω CSP171/MB46 PCR product from pKS02 DNA template (*Bsp*EI/*Nde*I; 816 bp), confers ampicillin resistance. FimD harbours a hexahistidine-tag between S608 and D609 (S563/D564 of mature FimD)	*His*$_{563/564}$-*fimD* expression vector	This study, Stubenrauch *et al.* (2016)
pCJS69	pACYCDuet-1 (*Nco*I/*Nde*I) Ω pKS06 (*Nco*I/*Nde*I; 1736 bp), confers chloramphenicol resistance	*tamA* complementation vector	This study, Heinz *et al.* (2016), and Stubenrauch *et al.* (2016)
pCJS72	pACYCDuet-1 (*Nde*I/*Xho*I) Ω CSP160/CSP161 PCR product from *E. coli* DH5α gDNA template (*Nde*I/*Xho*I; 3780 bp), confers chloramphenicol resistance	*tamB* complementation vector	This study, Stubenrauch *et al.* (2016)
pCJS73	pACYCDuet-1 (*Nde*I/*Xho*I) Ω pCJS60 (*Nde*I/*Xho*I); 1179 bp), confers chloramphenicol resistance	*bamB* complementation vector	This study, Stubenrauch *et al.* (2016)
pCJS74	pACYCDuet-1 (*Nde*I/*Xho*I) Ω CSP164/CSP165 PCR product from *E. coli* K-12 MG1655 gDNA template (*Nde*I/*Xho*I; 342 bp), confers chloramphenicol resistance	*bamE* complementation vector	This study, Stubenrauch *et al.* (2016)

optical density at 600 nm (OD_{600}) was measured with an Ultrospec 10 cell density meter (GE Healthcare Life Sciences).

2.2 Storage conditions

Following routine incubation, an aliquot of the saturated culture was then transferred to a cryogenic vial and mixed with an equal volume of 40 % v/v glycerol (20 % v/v final concentration), before snap freezing in liquid nitrogen and transferring to a long-term -80 °C storage freezer. Although plasmids were stored in *E. coli* BL21 Star™ (DE3) (and its derivatives) for subsequent pulse chase analysis, stock plasmids were stored in *E. coli* DH5α (which is a Δ*recA* mutant). When required, and without thawing the glycerol stock, a small portion of the frozen the glycerol culture was used to inoculate solid LB media for routine incubation.

2.3 Transformations

2.3.1 $CaCl_2$ Chemically competent cells and transformations

To prepare chemically competent cells with $CaCl_2$, saturated overnight cultures were diluted 1:100 into fresh 200 mL LB and incubated until an OD_{600} of 0.5. The cells were chilled for 30 minutes on ice, subjected to centrifugation (4415 **rcf**, 15 minutes, 4 °C) and the cell pellet was resuspended in 28 mL ice-cold 0.1 M $CaCl_2$. After chilling and following centrifugation, the cell pellet was resuspended in 1 mL ice-cold 0.1 M $CaCl_2$ and incubated on ice for 2 hours. Ice-cold LB media containing 30% v/v glycerol (500 µL) was then added to the suspension and 50 µL aliquots were snap frozen in liquid nitrogen. The cell aliquots were stored at -80 °C.

To transform $CaCl_2$ chemically competent *E. coli* strains, cells were first thawed on ice. Purified plasmid (1 µL) was added and, after gentle mixing, the sample was incubated on ice for 20 minutes. The sample was then heat-shocked for 45 seconds at 42 °C, incubated on ice for 2 minutes, and transferred to 950 µL of non-selective SOC or 2× YT media for a 45-minute incubation. Selective solid media was then inoculated with 10-100 µL of transformation mixture. Following a 24-hour incubation at 37 °C, a well-isolated single colony of transformants was used to inoculate fresh LB (liquid) media for overnight incubation, which was then used to prepare a glycerol storage stock.

2.3.2 Preparing Electrocompetent cells and transformations

Electrocompetent cells were prepared if the *E. coli* strain was to be transformed with a second plasmid. To prepare electrocompetent cells, saturated overnight culture was diluted 1:50 with 30 mL of fresh LB and incubated until the OD_{600} reading was between 0.4 and 0.6. Cells were centrifuged four times (3488 **rcf**, 10 minutes, 4 °C) where the cell pellet was resuspended using increasingly smaller volumes of ice-cold 10 % v/v glycerol: 12 mL, then 6 mL, then 3 mL, then 300 µL. Aliquots (50 µL) were either used for transformation immediately, or snap frozen in liquid nitrogen and stored at -80 °C. When required, frozen aliquots were thawed on ice before incubation with plasmid.

To transform electrocompetent *E. coli*, cells were incubated on ice for 3 minutes with 1 µL of purified plasmid. After transferring to pre-chilled ice-cold electroporation cuvettes (1 mm gap), cells were electroporated at 1800 V, 200 Ω and 25 µF and then immediately transferred to 250 µL of non-selective 2× YT or

SOC media. Following a 1-hour incubation, selective solid media was inoculated with 10-100 µL of transformation mixture and incubated overnight. Liquid media of the same composition was then inoculated with a well-isolated single colony of transformants and after overnight incubation, the saturated culture was used to prepare a glycerol storage stock.

2.4 Plasmid isolation

Plasmids were isolated from saturated overnight cultures using the Wizard® Plus SV Miniprep DNA purification system (Promega) kit as per manufacturer's instructions.

2.5 Plasmid construction

The plasmids constructed in this study are given in Table 2.1.2. Genes of interest were cloned in to pET-15b, pET-22b(+), pETDuet-1 or pACYCDuet-1 and were subsequently used in pulse chase and/or complementation experiments. Following each purification step listed below, samples were sometimes stored at -20 °C, and thawed on ice for use at a later time.

2.5.1 PCR

Primers were designed using Serial Cloner version 2.6.1 (SerialBasics). The oligonucleotides were synthesised by Integrated DNA Technologies (USA), and their sequences are listed in (Table 2.5.1.1). PCR, used to amplify a gene of interest for cloning purposes, was performed using Expand™ High Fidelity PCR System (Sigma Aldrich) as per manufacturer's instructions and using the appropriate primers as indicated in Table 2.1.2. If however, the indicated template DNA was

Table 2.5.1.1 PCR primers used in this study.

PCR primers were used to generate and/or sequence the indicated inserts. Primers may have a restriction site incorporated within their 5' end as indicated, otherwise not applicable (N/A) is shown.

Name	Sequence	Purpose	Site	Source or reference
MB35	5'-tttgacgtccatgggcaagaaatt gctccccattcttatc-3'	*tolC* cloning	*Nco*I	Designed by M. J. Belousoff, Stubenrauch *et al.* (2016)
MB36	5'-gttgacgtcatatgtcagttacgg aaagggttatgac-3'	*tolC* cloning	*Nde*I	Designed by M. J. Belousoff, Stubenrauch *et al.* (2016)
MB45	5'-tttgacgtccatgggctcatatctg aatttaagactttaccagc-3'	*fimD* cloning	*Nco*I	Designed by M. J. Belousoff, Stubenrauch *et al.* (2016)
MB46	5'-gttgacgtcatatgttaacgacatt cagctgatagctg-3'	*fimD* cloning	*Nde*I	Designed by M. J. Belousoff, Stubenrauch *et al.* (2016)
MB47	5'-tttgacgtccatgggcaaaaaacg tatccccactctcc-3'	*lptD* cloning	*Nco*I	Designed by M. J. Belousoff, unpublished
MB53	5'-tttgacgtccatgggccgctatatc cgacagttatgctg-3'	*tamA* cloning	*Nco*I	Designed by M. J. Belousoff, unpublished
MB54	5'-gttgacgtcatatgtcataattctg gccccagacc-3'	*tamA* cloning	*Nde*I	Designed by M. J. Belousoff, unpublished
MB55	5'-tttgacgtccatgggcaaaaagag cactctggcattagtg-3'	*phoE* cloning	*Nco*I	Designed by M. J. Belousoff, Stubenrauch *et al.* (2016)
MB56	5'-gttgacgtctcgagttaaaactgat acgtcatgccaac-3'	*phoE* cloning	*Xho*I	Designed by M. J. Belousoff, Stubenrauch *et al.* (2016)
KS FimD 899 F	5'-cggtgatccacggtattgctc-3'	*fimD* sequencing	N/A	Designed by K. S. Tan, unpublished
KS LptD 852 F	5'-cgctggcttgatggaactgg-3'	*lptD* sequencing	N/A	Designed by K. S. Tan, unpublished
CSP32	5'-ccggcccatggttactcatggtttt tatgcccg-3'	*eaeA* cloning	*Nco*I	This study, Heinz *et al.* (2016)
CSP36	5'-ccggccatatgttattttacacaag tggcataagc-3'	*eaeA* cloning	*Nde*I	This study, Heinz *et al.* (2016)
CSP37	5'-ccggccatatgttatttttcgaact gcgggtggctcca-3'	*strep*-II tag cloning	*Nde*I	This study, Stubenrauch *et al.* (2016)
CSP39	5'-tgacagtcagcatgacggacaat-3'	*fimD* sequencing	N/A	This study

CSP42	5'-tgacgctgggtgatggttatactca-3'	*fimD* sequencing	N/A	This study
CSP43	5'-tcgtgtgacatgctgtagctggcact-3'	*fimD* sequencing	N/A	This study
CSP44	5'-tgacggtagttcactggacttctta-3'	*eaeA* sequencing	N/A	This study
CSP45	5'-gctaacacgggcactaacgagtgat-3'	*eaeA* sequencing	N/A	This study
CSP47	5'-gccggggatccttatttgactgttttacaaggtaaaatt-3'	*yqiG* cloning	*Bam*HI	This study
CSP48	5'-ggacgctcggtaaaccgcgaca-3'	*yqiG* sequencing	N/A	This study
CSP49	5'-catctgcgaagtattactgcttcccg-3'	*yqiG* sequencing	N/A	This study
CSP51	5'-ccggctctagaatgcgtggaatgaaagacagaat-3'	*papC* cloning	*Bam*HI	This study, Stubenrauch *et al.* (2016)
CSP52	5'-aaacgacaaacagaaatttcacatg-3'	*papC* sequencing	N/A	This study
CSP53	5'-ccatattcacataacggtcatta-3'	*papC* sequencing	N/A	This study
CSP55	5'-gccggggatccttattgagttgctgaaattgacgt-3'	*yfcU* cloning	*Bam*HI	This study, Stubenrauch *et al.* (2016)
CSP56	5'-tgacgacgatgacagcagtaa-3'	*yfcU* sequencing	N/A	This study
CSP57	5'-acgttaacctggtagtgcgtt-3'	*yfcU* sequencing	N/A	This study
CSP75	5'-gccggggatccttactgtatctgacaacgaagc-3'	*htrE* cloning	*Bam*HI	This study, Stubenrauch *et al.* (2016)
CSP76	5'-aatcgctatttacaacgcgatct-3'	*htrE* sequencing	N/A	This study
CSP77	5'-ctgtttaattgtgtatcaatactctg-3'	*htrE* sequencing	N/A	This study
CSP80	5'-cgattttgctggtcaccgattt-3'	*ecpC* sequencing	N/A	This study
CSP81	5'-tgacagattggcaccaacagt-3'	*ecpC* sequencing	N/A	This study
CSP91	5'-gccggggatccttattaatgacaaggtaaaatcaggcg-3'	*ybgQ* cloning	*Bam*HI	This study, Stubenrauch *et al.* (2016)
CSP92	5'-tcggcgaaaccgatttcagtt-3'	*ybgQ* sequencing	N/A	This study
CSP93	5'-tacgttcatccagcgtatcgtt-3'	*ybgQ* sequencing	N/A	This study
CSP107	5'-gccggggatccttatttctcctcagcgccttcagt-3'	*fdeC* cloning	*Bam*HI	This study, Heinz *et al.* (2016)
CSP108	5'-tggcgcaagcctgatgtatgaa-3'	*fdeC* sequencing	N/A	This study
CSP109	5'-gcaacaaacttcagcgtttgttg-3'	*fdeC* sequencing	N/A	This study
CSP110	5'-caacaaacgctgaagtttgttgc-3'	*fdeC* sequencing	N/A	This study
CSP111	5'-agtcacggaaaccactttcagtt-3'	*fdeC* sequencing	N/A	This study
CSP119	5'-gccggggatccttacaaagtgttttgatacggcagaa-3'	*lptD* cloning	*Bam*HI	This study
CSP120	5'-ccaaatatggatgccaccatca-3'	*lptD* sequencing	N/A	This study
CSP121	5'-ccggccatatgcgatatctggcaacattgttgt-3'	*lptE* cloning	*Nde*I	This study
CSP122	5'-gccggctcgagttagttacccagcgtggtggaga-3'	*lptE* cloning	*Xho*I	This study

CSP123	5'-caacgattactgttagcgttgatg-3'	*fdeC* sequencing	N/A	This study
CSP124	5'-catcaacgctaacagtaatcgttg-3'	*fdeC* sequencing	N/A	This study
CSP125	5'-ctgactgttattgccggagag-3'	*fdeC* sequencing	N/A	This study
CSP126	5'-ctctccggcaataacagtcag-3'	*fdeC* sequencing	N/A	This study
CSP132	5'-ccggcccatggacaataaaaacacgttttcccgg-3'	*yqiG* cloning	*Nco*I	This study
CSP133	5'-ccggccatatgcgtggaatgaaagacagaat-3'	*papC* cloning	*Nde*I	This study, Stubenrauch *et al.* (2016)
CSP134	5'-ccggcccatgggtatgcctaatcactcaaattttcg-3'	*yfcU* cloning	*Nco*I	This study, Stubenrauch *et al.* (2016)
CSP141	5'-ccggcccatggctattaaatctactaatcaccttacc-3'	*htrE* cloning	*Nco*I	This study, Stubenrauch *et al.* (2016)
CSP142	5'-ccggcgacgtcatgcctttacgacggttctc-3'	*ecpC* cloning	*Aat*II	This study
CSP143	5'-gccggctcgagttaactctcattcttctcctgt-3'	*ecpC* cloning	*Xho*I	This study
CSP146	5'-ccggcccatggatatttatcgactctcttttgtatcct-3'	*ybgQ* cloning	*Nco*I	This study, Stubenrauch *et al.* (2016)
CSP152	5'-ccggcccatggcacgttataaaacagacaataaacag-3'	*fdeC* cloning	*Nco*I	This study, Heinz *et al.* (2016)
CSP156	5'-ggccggaattcttaacgacattcagctgatagct-3'	*fimD* cloning	*Eco*RI	This study, Stubenrauch *et al.* (2016)
CSP158	5'-ggccgcatatgcgctatatccgacagttatgctg-3'	*tamA* cloning	*Nde*I	This study, Heinz *et al.* (2016), and Stubenrauch *et al.* (2016)
CSP159	5'-ggccgctcgagttataattctggcccagaccgat-3'	*tamA* cloning	*Xho*I	This study, Heinz *et al.* (2016), and Stubenrauch *et al.* (2016)
CSP160	5'-ggccgcatatgagtttatggaaaaaatcagcctc-3'	*tamB* cloning	*Nde*I	This study, Stubenrauch *et al.* (2016)
CSP161	5'-ggccgctcgagttaaaactcgaactgatagagcaaatc-3'	*tamB* cloning	*Xho*I	This study, Stubenrauch *et al.* (2016)
CSP162	5'-ggccgcatatgcaattgcgtaaattactgctgc-3'	*bamB* cloning	*Nde*I	This study, Stubenrauch *et al.* (2016)
CSP163	5'-ggccgctcgagttaacgtgtaatagagtacacggtt-3'	*bamB* cloning	*Xho*I	This study, Stubenrauch *et al.* (2016)
CSP164	5'-ggccgcatatgcgctgtaaaacgctgactg-3'	*bamE* cloning	*Nde*I	This study, Stubenrauch *et al.* (2016)
CSP165	5'-ggccgctcgagttagttaccactcagcgcaggtt-3'	*bamE* cloning	*Xho*I	This study, Stubenrauch *et al.* (2016)
CSP168	5'-ctatctccggagtggtgatggtgatgatgcggtattatgtctcgctcaagccag-3'	*$His_{211/212}$-fimD cloning*	*Bsp*EI	This study, Stubenrauch *et al.* (2016)

CSP169	5'-cgctgtccggattacgttcccggctgacgctg-3'	*His$_{211/212}$-fimD cloning*	*Bsp*EI	This study, Stubenrauch *et al.* (2016)
CSP170	5'-ctatctccggagtggtgatggtgatgatgagaacgcagccagtggctgaaag-3'	*His$_{563/564}$-fimD* cloning	*Bsp*EI	This study, Stubenrauch *et al.* (2016)
CSP171	5'-ctatctccggagacagtaaatctcagtggcgacatgc-3'	*His$_{563/564}$-fimD* cloning	*Bsp*EI	This study, Stubenrauch *et al.* (2016)
CSP172	5'-ctatctccggagtggtgatggtgatgatggggatcccataactcaggaggaatataacc-3'	*His$_{135/136}$-fimD* cloning	*Bsp*EI	This study
CSP173	5'-ctatctccggaggtattaatgccggattgctcaattataatttcag-3'	*His$_{135/136}$-fimD* cloning	*Bsp*EI	This study
CSP183	5'-gtaaatatcaacctcgtcaccctgaatg-3'	*tamB* sequencing	N/A	This study
CSP184	5'-cattcagggtgacgaggttgatatttac-3'	*tamB* sequencing	N/A	This study
CSP185	5'-ccagaccgtatttgctcttaacg-3'	*eaeA* sequencing	N/A	This study
T7 Promoter	5'-taatacgactcactataggg-3'	pET-15b or pET-22b(+) sequencing	N/A	Novagen
T7 Terminator	5'-gctagttattgctcagcgg-3'	pET-15b, pET-22b(+) or MCS2 sequencing	N/A	Novagen
pET Upstream	5'-atgcgtccggcgtaga-3'	MCS1 sequencing (pETDuet-1 only)	N/A	Novagen
DuetDOWN1	5'-gattatgcggccgtgtacaa-3'	MCS1 sequencing	N/A	Novagen
DuetUP2	5'-ccagaccgtatttgctcttaacg-3'	MCS2 sequencing	N/A	Novagen

genomic DNA (gDNA), part of a single colony from an overnight (solid media) incubation of the indicated strain was suspended in the PCR reaction mixture. Following PCR amplification, the sample was purified using the Wizard® SV Gel and PCR Clean-Up System (Promega) kit.

2.5.2 Restriction digestion and alkaline phosphatase treatment

All restriction enzymes were from New England Biolabs and where possible, high-fidelity versions of enzymes were used. Plasmids or purified PCR amplicons were "double digested" as indicated in Table 2.1.2 as per manufacturer's instructions for one hour. CutSmart® buffer was typically used during double digestion reactions. If double restriction digests included *Bsp*EI, NEBuffer 3.1 was used instead of CutSmart® buffer and normal-fidelity versions of the second restriction enzyme were used. The linearised plasmid was then usually treated with alkaline phosphatase (Promega, M182A) for 1 hour as per manufacturer's instructions. After alkaline phosphatase treatment and/or restriction enzyme digestion, DNA samples were purified by gel electrophoresis using 0.8-1.5 % w/v molecular biology grade agarose (Scientifix, 9010E) in TAE buffer as per Sambrook and Russell (2006). DNA was stained using SYBR® Safe (Thermo Fisher Scientific) as per manufacturer's instructions. Slices corresponding to the digested vector or insert were extracted and DNA was purified using the Wizard® SV Gel and PCR Clean-Up System (Promega) kit.

2.5.3 DNA Ligation and plasmid sequencing

The concentration of purified vector and insert DNA was estimated using the NanoDrop 1000 Spectrophotometer (Thermo Scientific) as per manufacturer's

instructions. Ligation using the T4 DNA Ligase (New England Biolabs) system was performed as per the manufacturer's instructions, using the overnight ligation protocol and usually a 1:3 vector:insert ratio. However, for construction of pCJS64, pCJS65 and pCJS66 that required ligation of two DNA inserts to a pET-15b vector, a vector:insert:insert ratio of 1:1:1 was instead used. Ligation mixture (10 µL) was used to transform chemically competent *E. coli* DH5α (instead of 1 µL purified plasmid) as described in Section 2.3.1, but the saturated overnight culture was also used for plasmid isolation (as per Section 2.4). The plasmid was then sequenced by Macrogen, Inc. (South Korea) using the relevant sequencing primers listed in Table 2.5.1.1 specific for the gene of interest and/or the base vector. Silent mutations likely generated during PCR amplification, as well as intentional mutations introduced from the specific oligopeptide sequence, such as fusion tags, are listed in Appendix 1.

2.6 Polyacrylamide Gel Electrophoresis

Proteins were analysed by polyacrylamide gel electrophoresis (PAGE) using Mini-PROTEAN® Tetra vertical electrophoresis cell (Bio-Rad) equipment as per manufacturer's instructions. For PAGE analysis, 10 µL of sample was loaded into polyacrylamide gels. When preparing gradient gels, the Hoefer™ SG 50 gradient maker (GE Healthcare Life Sciences) was used as per manufacturer's instructions.

2.6.1 SDS-PAGE

Sodium dodecyl sulphate (SDS)-PAGE analysis was performed as previously described (Simpson, 2006) with several exceptions. The 4× resolving gel buffer was instead comprised of 0.4 % w/v SDS, 2 mM EDTA and 1.5 M Tris, pH 8.8. The

4× stacking gel buffer was instead comprised of 0.08% w/v SDS, 2 mM EDTA and 1.0 M Tris, pH 6.8. The SDS sample buffer (also known as SDS loading buffer) was instead comprised of 10 % v/v glycerol, 1 % w/v SDS, 100 mM DTT, 0.01 % w/v bromophenol blue and 100 mM Tris-HCl, pH 6.8. The 30 % T, 2.6 % C acrylamide/bis solution was exchanged for 40 % T, 3.3 % C acrylamide/bis solution (Bio-Rad, #1610146) and recipes for preparing the polyacrylamide gel solutions were adjusted accordingly.

One of two molecular weight standards was used for SDS-PAGE analysis. Kaleidoscope™ or Dual color Precision Plus Protein™ Pre-stained standards (Bio-Rad) - which have the same molecular weight marker profile - were diluted 1:5 or 1:10 with SDS sample buffer, respectively and without boiling, 10 µL was loaded into polyacrylamide gels for SDS-PAGE analysis. Unless otherwise indicated, 12% polyacrylamide gels were used during SDS-PAGE analysis.

2.6.2 SN-PAGE

Semi-native (SN)-PAGE analysis was performed using a similar method to SDS-PAGE, but buffers were of a different composition. Instead of SDS running buffer, MES running buffer was used, containing 50 mM 2-(*N*-morpholino)ethanesulphonic acid (MES), 50 mM Tris, 2 mM EDTA and 0.2 % w/v SDS (the pH was not adjusted, but was about pH 7.3). Instead of SDS loading buffer (also known as SDS sample buffer), SN sample buffer was used, containing 40 mM Tris (pH 6.8), 10 % v/v glycerol, 0.2 % w/v SDS, 0.01 % w/v bromophenol blue.

Instead of 4× resolving and stacking gel buffers, 4× SN gel buffer was used

comprising 2.0 mM EDTA and 1.5 M Tris pH 8.8 (i.e. no stacking layer). SN-PAGE analysis was performed using 4-16 % gradient gels. The same molecular weight markers used for SDS-PAGE analysis were used, except that the markers were diluted in SN sample buffer for SN-PAGE analysis.

2.6.3 BN-PAGE

Proteins were analysed by blue native (BN)-PAGE using the BE-250 (Bio Craft) electrophoresis equipment as per manufacturer's instructions. BN-PAGE molecular weight standards were freshly prepared before use and were comprised of 1 $mg.mL^{-1}$ ferritin, 1 $mg.mL^{-1}$ catalase, 1 $mg.mL^{-1}$ BSA, 1 mM EDTA, 50 mM NaCl, 7.5 % w/v glycerol, 55 mM 6-aminohexanoic acid, 25 mM imidazole (pH 7.0) and 0.2 % w/v Coomassie Brilliant Blue G-250.

BN-PAGE analysis was performed as previously described by Wittig *et al.* (2006) and as modified by Shiota *et al.* (2015), using 4-14 % gradient gels, but with several exceptions. Cathode buffer B was replaced by deep blue cathode buffer, comprising 0.05 % w/v Coomassie Brilliant Blue G-250 in 50 mM tricine-imidazole, pH 7.0 (i.e. tricine adjusted with 1 M imidazole to pH 7.0). Cathode buffer B/10 was replaced by slightly blue cathode buffer, comprising 0.001% w/v Coomassie Brilliant Blue G-250 in 50 mM tricine-imidazole, pH 7.0. The 49.5 % T, 3.0 % C acrylamide/bis solution was exchanged for the 40 % T, 2.6 % C acrylamide/bis solution (Bio-Rad, #1610148) and recipes for preparing the polymerising solutions were adjusted accordingly. The polymerising solutions additionally contained 3.75 % w/v glycerol (for 4 % T acrylamide/bis polymerising solution) or 20.25 % w/v glycerol (for 14 % T acrylamide/bis

polymerising solution).

2.7 Staining, destaining and drying SDS-PAGE gels

Following SDS-PAGE analysis, gels may be stained with Coomassie staining solution, comprising 50 % v/v methanol, 10 % v/v acetic acid and 0.05 % w/v Coomassie Brilliant Blue R-250. Gels were incubated in Coomassie staining solution for at least 1 hour at 40-50 rpm (32 mm orbit), and then destained with destaining solution, comprising 5 % v/v methanol and 7 % v/v acetic acid. Gels were usually destained for up to 24 hours at 40-50 rpm (32 mm orbit).

After destaining was completed, gels were transferred to gel drying solution, comprising 20 % v/v ethanol, 10 % v/v glycerol and 5 % v/v acetic acid. Gels were incubated in gel drying solution for at least 1 hour at 40-50 rpm (32 mm orbit) and heat-dried for one hour using DryGel Jr SE540 (Hoefer Scientific instruments) or the Slab GEL Dryer model 1125 (Bio-Rad) as per manufacturers' instructions, and subsequently scanned using CanoScan 8600F with software version 5.0.1.2 (Canon).

2.8 Western transfer

Where necessary, following SDS-PAGE or SN-PAGE analysis, proteins were transferred from gels to 0.45 μm hydrophobic immobilon-P PVDF transfer membranes (Merck Millipore), or transferred to 0.45 μm Protran nitrocellulose hybridisation transfer membranes (PerkinElmer). Western transfer was performed using mini Trans-Blot® electrophoretic transfer cell (Bio-Rad) as per manufacturer's instructions, except that transfer buffer A also contained 10 % v/v

methanol.

Following BN-PAGE analysis, proteins were denatured by transferring the gel to denaturing buffer pre-heated to 65 °C and comprising 4 % w/v SDS, 1 % w/v DTT and 50 mM Tris-HCl, pH 6.8. Following incubation for 20 minutes at 37 °C and 40-50 rpm (25 mm orbit), the gel was rinsed twice in water and incubated for a further 10 minutes at 40-50 rpm (25 mm orbit) in CAPS buffer, comprising 10 % v/v methanol and 10 mM 3-(cyclohexylamino)-1-propanesulphonic acid (CAPS), pH 11. The denatured proteins were then transferred to 0.45 µm PVDF using Trans-Blot® electrophoretic transfer cell (Bio-Rad) as per the manufacturer's instructions, except that CAPS buffer was used as the transfer buffer.

2.9 Immunoblotting

The antibodies used for immunoblotting are listed in Table 2.9.1. Primary and secondary antibodies were diluted (using the dilution factor indicated in Table 2.9.1) in blocking solution, comprising 5 % w/v instant skim milk powder (Coles) in Western washing buffer. Western washing buffer was prepared as per CSH Protocols (2013) except the final concentration of Tween-20 was 0.2 % v/v. All incubations were performed at room temperature at 40-50 rpm (25 mm orbit) unless otherwise indicated. Primary and secondary antibodies are stored at -80 °C and -20 °C, respectively, and when required aliquots were thawed on ice and subsequently stored at 4 °C. Antibodies (either concentrated or diluted in blocking solution) were usually supplemented with 0.025 % w/v sodium azide for prolonged storage at 4 °C, except for antibodies conjugated to horse-radish peroxidase (HRP).

Table 2.9.1 List of antibodies used in this study for immunoblotting.

Antibody	Type	Dilution	Source or reference
Mouse αBamA	Primary	1:2,000	Webb *et al.* (2012)
Mouse αBamB	Primary	1:5,000	Webb *et al.* (2012)
Mouse αBamC	Primary	1:10,000	Webb *et al.* (2012)
Mouse αBamD	Primary	1:10,000	Webb *et al.* (2012)
Mouse αBamE	Primary	1:500	Webb *et al.* (2012)
Mouse α6×His	Primary	1:2,000	BD Pharmingen™, 51-9000012
Rabbit αBamA	Primary	1:20,000	T. Shiota, unpublished
Rabbit αDnaK	Primary	1:3,000	Clements *et al.* (2009)
Rabbit αTamA	Primary	1:5,000	Selkrig *et al.* (2012)
Rabbit αTamB	Primary	1:10,000	Selkrig *et al.* (2012)
Goat αMouse-HRP	Secondary	1:20,000	Sigma Aldrich, A4416
Goat αRabbit-HRP	Secondary	1:20,000	Sigma Aldrich, A6154

Following Western transfer, membranes were incubated for 1 hour in blocking solution at room temperature or incubated in 50 % v/v blocking solution in Western washing buffer for 16 hours at 4 °C with gentle rocking. After removing the blocking solution, the membrane was incubated for one hour in the diluted primary antibody solution. The primary antibody solution was removed and stored at 4 °C for reuse, and the membrane was washed twice for 5 minutes and twice for 10 minutes in Western washing buffer. The membrane was incubated for one hour in secondary antibody solution and washed as before.

Protein detection was performed using the enhanced chemiluminescence method with the ECL Prime Western blotting detection reagent (GE Healthcare Life Sciences) as per manufacturer's instructions. Chemiluminescent membranes were then exposed to Super RX-N film (Fujifilm) in an Amersham Hypercasesette™ (GE Healthcare Life Sciences) for up to 1 hour (usually 5 minutes). Alternatively, for detection of hexahistidine-tagged proteins, the more sensitive Amersham Hyperfilm ECL (GE Healthcare Life Sciences) was used. The

film was subsequently processed using the FPM-100A film processor (Fujifilm) or the SRX-101A medical film processor (Konica) as per manufacturers' instructions, and scanned using CanoScan 8600F with software version 5.0.1.2 (Canon).

For detection of *strep*-tag II, the immunoblotting method was followed with several exceptions. The blocking solution was comprised of 5 % w/v BSA (Sigma Aldrich, A7906) in washing buffer. The incubations with a secondary antibody and subsequent washing steps were omitted. Instead of the primary antibody solution, membranes were incubated for one hour with Precision Protein™ StrepTactin-HRP conjugate (Bio-Rad) diluted 1:5000 with blocking solution (no sodium azide). Because the *strep*-tag II is also fused to the pre-stained markers, the lane containing the standards was removed from the membrane prior to immunoblotting, and then reattached (for size referencing) before exposure to film. This was to ensure the extremely strong signal from the markers did not interfere with the much weaker signal from the protein of interest.

2.10 *E. coli* membrane preparation and analysis

Saturated overnight cultures were diluted 1:100 into fresh 50 mL LB and incubated until an OD_{600} between 1.2 and 2.0 was reached. Cells were isolated by centrifugation (4609 **rcf**, 10 minutes, 4 °C) and the cell pellet was resuspended in 10 mL ice-cold sonication buffer, comprising 300 mM NaCl, 2 mM EDTA and 10 mM Tris, pH 7.0. The sample was sonicated using a Branson Sonifier 450 as per manufacturer's instructions with a 45-50 % duty cycle and up to 30 ten-second pulses at output 5. Between each ten-second pulse, the sample was chilled for 2-3 minutes in an ice/water bath. The sonicated sample was subjected to

centrifugation (2668 **rcf**, 5 minutes, 4 °C) to remove unbroken cells and the supernatant was further centrifuged (16743 **rcf**, 10 minutes, 4 °C). This final pellet corresponded to a crude *E. coli* membrane preparation and after resuspension in 1 mL SEM buffer, comprising 250 mM sucrose, 1 mM EDTA and 10 mM MOPS-KOH (pH 7.2), it was snap frozen in liquid nitrogen (250 µL aliquots) and stored at -80 °C.

Aliquots were thawed on ice and 10 µL of membrane extract (1 % v/v final) was incubated with 0.6 % w/v SDS (in water) and boiled for 5 minutes. After isolating the insoluble material by centrifugation (13523 **rcf,** 5 minutes), the optical density of the supernatant at 280 nm (OD_{280}) was determined using the SmartSpec™ 3000 (Bio-Rad) as per manufacturer's instructions. The OD_{280} was used to estimate the crude protein concentration and determine the amount of membrane required for subsequent Western blot analysis following either SDS-PAGE or BN-PAGE. For SDS-PAGE analysis, a volume of membrane extract corresponding to 0.28 $\mu L.OD_{280}^{-1}$ (per lane) was subjected to centrifugation (13523 **rcf**, 5 minutes). The membrane pellet was resuspended in 15 µL (per lane) of SDS sample buffer, boiled for 3 minutes and 10 µL (per lane) was used for separation by SDS-PAGE. Following Western transfer, TamA and BAM complex proteins were detected by as per Section 2.9.

For BN-PAGE analysis, the volume of membrane extract required for analysis was: 2.8 $\mu L.OD_{280}^{-1}$ (per lane) for native BAM complex and 5.6 $\mu L.OD_{280}^{-1}$ (per lane) for native TAM. An appropriate volume of membrane extract was subjected to centrifugation (13523 **rcf**, 5 minutes, 4 °C) and the membrane pellet was

resuspended in 24 µL (per lane) of lysis buffer, comprising 10 mg.mL^{-1} DDM and 1 mM PMSF (added immediately before use) in BN-lysis buffer (Shiota *et al.*, 2015). Following a 30-minute incubation on ice, the sample was subjected to centrifugation (21130 **rcf**, 10 minutes, 4°C) and the supernatant was transferred to 6 µL (per lane) of 5×BN sample buffer, comprising 1 % w/v Coomassie Brilliant Blue G-250 and 25 mM 6-aminohexanoic acid in 75 % v/v lysis buffer (in water). For analysis, 20 µL (per lane) was analysed by BN-PAGE. Following Western transfer, BAM complex proteins and/or the TAM were detected as per Section 2.9.

2.11 Pulse chase analysis

An extensive discussion of the pulse chase analysis method is provided in Chapter 3, including most variations used during pulse chase optimisation. The optimised pulse chase analysis method will be described here, and the chase temperature, unless otherwise indicated, should be taken to be 30 °C. Also, unless otherwise indicated, samples were incubated at the indicated temperature statically and, if not indicated, the temperature should be assumed to be the chase temperature.

2.11.1 Cell preparation for pulse chase analysis

Following routine incubation, saturated overnight cultures were diluted 1:50 into fresh LB media, incubated until mid-log phase and normalised to an OD_{600} of 0.6. The sample was subjected to centrifugation twice (4613 **rcf**, 10 minutes, 4 °C), and the cell pellet washed (with M9-S media) then resuspended (after the second round of centrifugation) in an equal volume of M9-S media (equal to the normalised LB media volume), comprising 47.8 mM Na_2HPO_4, 22 mM KH_2PO_4, 8.56 mM NaCl, 11.1 mM glucose, 1.12 mM thiamine, 1 mM $MgCl_2$, 0.1 mM $CaCl_2$,

45.4 pg.mL^{-1} of standard non-sulphur containing amino acids, except 227 pg.mL^{-1} for leucine. The sample was diluted 1:2 with 40 % v/v glycerol and snap frozen (1.3 mL aliquots) in liquid nitrogen and stored at -80 °C.

2.11.2 General pulse chase analysis steps

Aliquots were thawed on ice, centrifuged (1693-3000 **rcf**, 5 minutes, 4 °C) and resuspended in 650 µL fresh M9-S media (i.e. volume before dilution with glycerol). Rifampicin (500 µg.mL^{-1}) was added to inhibit transcription by incubation at 37 °C for one hour with mixing at 400 rpm (3 mm orbit). Unless otherwise indicated, subsequent incubations were performed at the chase temperature without mixing. Following rifampicin treatment, 0.2 mM of pre-warmed IPTG was added and the sample was incubated for 5 minutes. Cells were 'pulsed' with 22 µCi.mL^{-1} of pre-warmed [^{35}S]-methionine and [^{35}S]-cysteine by addition of EXPRE^{35}S^{35}S [^{35}S]-Protein Labelling Mix (Perkin Elmer, NEG072). It should be noted that the volume of EXPRE^{35}S^{35}S [^{35}S]-Protein Labelling Mix added varied daily according to the residual sample radioactivity (based on its half-life of 87 days).

After a 40-second 'pulse', the samples were immediately placed on ice, centrifuged as before and resuspended in 650 µL of pre-warmed chase media, M9+S media. M9+S media has a similar composition to M9-S media, except that 1 mM $MgSO_4$ replaces $MgCl_2$ and 45.4 pg.mL^{-1} methionine and cysteine were also included. The 'chase' component was considered to have begun immediately on addition of M9+S media (rather than on resuspension of cell pellet) and may proceed up to 256 minutes. Routine incubations were for 32 or 60 minutes.

2.11.3 Pulse chase analysis to detect oligomerisation

For analysis of PhoE or TolC oligomerisation, at each chase time point duplicate 50 µL aliquots were transferred to 10 µL of 6× SN sample buffer and incubated for 10 minutes at 37 °C or 100 °C and then transferred to ice. Samples (10 µL) were separated by 4-16 % SN-PAGE.

2.11.4 Pulse chase analysis to detect surface localisation by protease shaving

For analysis by protease shaving, at each chase time point 50 µL aliquots were incubated on ice for 10 minutes with (+PK) or without (-PK) 50 $\mu g.mL^{-1}$ proteinase K (PK) (Promega, MC500). Alternatively, the 50 µL aliquots were incubated for 30 minutes at 30 °C with (+TEV) or without (-TEV) 100 $\mu g.mL^{-1}$ TEV protease (Selkrig *et al.*, 2015) and 1 mM DTT. Following proteolysis (+PK/+TEV) or 'mock' proteolysis (-PK/-TEV), 10 % v/v TCA was added and protein precipitates were collected by centrifugation (25000 **rcf**, 15 minutes, 4 °C). The precipitate was washed with 200 µL acetone, centrifuged as before, and after removing the supernatant, the pellet was allowed to air dry. The protein pellet was resuspended in 60 µL SDS sample buffer, boiled for 3 minutes and 10 µL was analysed by SDS-PAGE.

2.11.5 Detection of radioactivity

SDS-PAGE or SN-PAGE gels were sometimes dried (with or without the Coomassie staining step, but with destain to 'fix' proteins in the gel) as per Section 2.7. Alternatively, and unless otherwise indicated, gels were transferred to nitrocellulose or PVDF membrane as per Section 2.8. Following Western transfer, immunoblotting may be performed before or after detection of radioactivity, but

the membrane must be air-dried before radioactivity detection. The pre-stained molecular weight markers were marked using a red dye spiked with 0.01-0.1 % v/v EXPRE^{35}S^{35}S [^{35}S]-Protein Labelling Mix and the dye was allowed to air-dry or was covered with clear tape. The dried gel or membrane was transferred to a Molecular dynamics exposure cassette (GE Healthcare Life Sciences) and exposed to a Storage Phosphor Screen (GE Healthcare Life Sciences). Radiation was captured for 12-18 hours and analysed using Typhoon Trio (detection set to 320 nm). To prevent the dry gel from sticking to the storage phosphor screen, it was overlaid with cling wrap. Following analysis, residual information stored in the storage phosphor screens was removed as per manufacturer's instructions using a light box (Medialux).

2.12 Liquid chromatography-tandem mass spectrometry

Samples that were to be analysed by liquid chromatography-tandem mass spectrometry (LC-MS/MS) were prepared for pulse chase analysis, snap frozen in liquid nitrogen, thawed and resuspended in M9-S media as per Section 2.10. Duplicate samples were induced with 0.2 mM IPTG and incubated at 37 °C with mixing at 400 rpm (3 mm orbit). After a 30-minute induction, 0.5 mg.mL^{-1} of rifampicin was added for a subsequent 45-minute incubation. One sample was pulsed as per Section 2.10 and the second was "mock" pulsed by a further 40 second incubation without radioactivity. Cells were chased for 16 minutes and prepared as per Section 2.10, except that the entire 650 µL "mock" pulsed sample was TCA-precipitated, thrice acetone-washed and the pellet was resuspended in 100 µL SDS sample buffer with the aid of a micropestle (Sigma Aldrich) (i.e. concentrated 7.8 times more than the normal resuspension method). The

radioactive and non-radioactive (both 10 µL) preparations were analysed by SDS-PAGE using the same polyacrylamide gel.

The gel was cut in half, the radioactive sample was analysed as in Section 2.10 and the non-radioactive sample was subjected to Instant*Blue*™ Ultrafast protein stain (Sigma Aldrich) as per manufacturer's instructions, except that the gel was rinsed twice with water for 2-3 hours and then once overnight. Bands corresponding to the full length and fragment bands were excised from the gel, using the autoradiogram obtained from the duplicate radioactive sample as a guide. For FimD, in-gel tryptic digestion was performed as per Phu *et al.* (2011), and LC-MS/MS analysis was performed by the Monash Biomedical Proteomics Facility as per Stubenrauch *et al.* (2016). For FdeC, in-gel tryptic digestion and subsequent LC-MS/MS analysis was performed by N. P. Croft as per Heinz *et al.* (2016).

2.13 Computer-based methods

All figures were prepared using Adobe Illustrator release number 2015.2.1 and where required, images were cropped or re-sized using Adobe Photoshop release number 2015.1.2. Protein structures were predicted using Phyre2 (Kelley *et al.*, 2015) and subsequent rendering of the predicted structure or actual structures deposited into the RCSB protein data bank (PDB) (http://www.rcsb.org/) were performed using PyMOL™ Open-Source version 1.8.2.1 (Schrödinger).

ExPASy PeptideCutter (Gasteiger *et al.*, 2005) was used to predict the sensitivity of a protein substrate to a particular protease. To determine the primary sequence of mature protein following signal peptide cleavage, analysis of the preprotein by

SignalP version 4.1 was performed (Petersen *et al.*, 2011). To predict the number and orientation of IM transmembrane α-helices, TMHMM server version 2.0 was used (Krogh *et al.*, 2001). ExPASy ProtParam (Gasteiger *et al.*, 2005) was used to estimate molecular weight and calculate the distribution of amino acids or atoms within a protein. Multiple sequence alignments were performed using the MUSCLE algorithm (Edgar, 2004) via the webtool developed by Li *et al.* (2015a).

Following Densitometrical analysis of radiographs was performed using ImageQuant™ TL software version 7.0 (GE Healthcare Life Sciences) as per manufacturer's instructions with the 1-D gel analysis module. To account for loading and pipetting differences between timepoints, each lane was normalised to the bands it contained and the total density count was normalised to unity. Normalised density plots were prepared using GraphPad Prism® version 6.0h.

Chapter 3 – Developing a method to find novel TAM substrates

3.1 Introduction

Despite being ubiquitous throughout the *Proteobacterial* lineage, a characteristic of the TAM is that neither TamA nor TamB are essential for cell viability, which allows for the recovery and analysis of deletion mutants. With the advent of both whole genome sequencing (Loman & Pallen, 2015) and signature-tagged mutagenesis (Saenz & Dehio, 2005) techniques in the mid-late-1990s, novel virulence factors could be determined using high throughput screens of deletion mutant collections. Three such studies, and a fourth utilising the advanced INSeq technique which combines signature-tagged mutagenesis and high-throughput sequencing (Goodman *et al.*, 2011), reported virulence or colonisation defects following transposon-based inactivation of at least one TAM component.

Struve *et al.* (2003) employed signature-tagged mutagenesis to detect novel virulence-associated genes in *Klebsiella pneumoniae*. Although a Δ*tamB* transposon-insertional mutant showed no *in vitro* growth defects, the mutant strain was significantly outcompeted by the wildtype parent strain during co-challenge experiments in the mouse bladder, but not the gastrointestinal tract. Burall *et al.* (2004) performed signature-tagged mutagenesis to survey for novel virulence genes in *Proteus mirabilis*. They found that during an ascending urinary tract infection model, Δ*tamA* or Δ*tamB* transposon-insertional mutants were significantly outcompeted by the wildtype parent strain in the urine, bladder and kidney of co-challenged mice. The mutants showed no *in vitro* growth defects and were still capable of producing functional urease. The authors showed the mutants were still also capable of producing the major flagella (FlaA) and major

MRP (mannose-resistant *Proteus*-like) fimbrial (MrpA) subunits, but whether the strains could still produce functional flagella or MRP fimbriae was not assessed.

Kelly *et al.* (2006) employed signature-tagged mutagenesis to investigate novel virulence factors in *Citrobacter rodentium*. They found that *C. rodentium* Δ*tamB* transposon-insertional mutants were significantly attenuated for the ability to colonise the GI tract of mice during a co-challenge with the wildtype parent strain. Selkrig *et al.* (2012) confirmed and extended these results by showing that, in a similar animal model, both *C. rodentium* Δ*tamA* and Δ*tamB* mutants were attenuated for colonisation in the GI tract of mice, despite normal growth *in vitro*. Additionally, Selkrig *et al.* (2012) showed that both *Salmonella* Typhimurium Δ*tamA* and Δ*tamB* mutants were significantly more sensitive to complement-mediated killing in human serum than the isogenic wildtype parent strain, despite normal *in vitro* growth. Brooks *et al.* (2014) used INSeq to identify colonisation determinants of *Aliivibrio* (formerly *Vibrio*) *fischeri*, which are a fluorescent group of oceanic bacteria that engage in a symbiotic relationship with a host species to provide light. Although an *A. fischeri* Δ*tamB* transposon-insertional mutant had normal growth *in vitro*, it was attenuated for its ability to colonise its symbiotic host *Euprymna scolopes* (Hawaiian bobtail squid).

While these four signature-tagged mutagenesis studies have not formally satisfied molecular Koch's postulates, if these data are taken together it is clear that, with five independent mutants in four different bacteria - and with the data from Selkrig *et al.* (2012) also included (nine independent mutants in five different bacteria) - there is a general trend for the TAM to be needed for effective growth

in vivo, rather than *in vitro* growth. In contrast to this trend, *Aggregatibacter actinomycetemcomitans* Δ*tamB* (*tamB* was referred to as *morC*) mutants display severe *in vitro* morphological defects (Azari *et al.*, 2013; Gallant *et al.*, 2008) and an inability to secrete leukotoxin (LtxA) *in vitro*. The authors suggested that TamB (and therefore the TAM) may help to assemble the T1SS responsible for secreting LtxA, but a later report by the same group showed that T1SS levels in *A. actinomycetemcomitans* Δ*tamB* are normal (Smith *et al.*, 2015). By comparing the *A. actinomycetemcomitans* wildtype and Δ*tamB* proteomes, Smith *et al.* (2015) found that TamB has unusual affects on several quality control systems. They found that the levels of three cytoplasmic chaperones (HtpX, SlyD and GroES) were reduced and the level of a DegP homologue, DegQ, was increased, possibly explaining the severe morphological defects. The authors also found that two oxidative stress response proteins (SodA and MsrB) were down-regulated, possibly explaining the loss of LtxA secretion, because the leukotoxin is rapidly inactivated by reactive oxygen species (Balashova *et al.*, 2007). This association with quality control systems has not been reported for other *Proteobacteria*, and highlights the importance of the TAM in the *Pasteurellaceae*.

A screen of 324 stress conditions for 3979 *E. coli* strains (comprised of 3737 non-essential gene deletion mutants, 100 sRNA deletion mutants, 133 mutants with modified essential genes and their 9 associated control strains) revealed a comprehensive list of *E. coli* mutant phenotypes (Nichols *et al.*, 2011). Additional data analysis on the 324 tested conditions for *E. coli* Δ*tamA* and Δ*tamB* (Appendix 2) reveals that *E. coli* had wildtype-level sensitivity to vancomycin, erythromycin and SDS, as has been reported elsewhere (Stegmeier *et al.*, 2007). A total of eight

phenotypes were found for *E. coli* Δ*tamA* or Δ*tamB* mutants (Table 3.1.1), but the only differential phenotype common to both *E. coli* Δ*tamA* and Δ*tamB* was the increased resistance to bleomycin.

In mapping the function of the TAM, Selkrig *et al.* (2012) compared the extracellular secretome and outer membrane protein extracts of *C. rodentium* wildtype, Δ*tamA* and Δ*tamB* strains. Although there were no obvious differences across the secretomes, the authors observed that a 180 kDa AIDA-like autotransporter, p1121, was absent from outer membrane extracts of *C. rodentium* Δ*tamA*. The authors subsequently showed that in the absence of the TAM, p1121 would not insert into the outer membrane, and instead remained in the periplasm. To expand their results to other AIDA-like autotransporters, Selkrig *et al.* (2012) confirmed that both Ag43 and EhaA were substrates of the TAM that could not function in cell-cell aggregation in the absence of the TAM.

In order to survey for other protein substrates of the TAM, I sought to adapt a "pulse chase" assay (Leyton *et al.*, 2014; Stenberg *et al.*, 2007). Pulse chase analysis is traditionally accomplished by "pulsing" the cells with a radioactive label, and then "chasing" the cells with the non-radioactive isotopic form of the compound. The Stenberg *et al.* (2007) method modifies the traditional rifampicin-blocking and pulse chase techniques (Studier *et al.*, 1990; Tabor & Richardson, 1985) by inducing target gene expression after rifampicin treatment (Figure 3.1.1a). Normally, 200 μg/mL of rifampicin was added after a 30-minute induction to increase production of the target protein, but by reversing the order and reducing the induction time (and amount of protein), the subtleties of a protein's biogenesis

Table 3.1.1 Significant TAM mutant phenotypes indicated by colony size in the presence of various antimicrobial agents.

This table is a shorter version of Appendix 2 listing only the "stress" conditions where at least one concentration of each indicated antimicrobial revealed a significant positive or negative growth phenotype for *E. coli* TAM mutants. The data was originally reported by Nichols *et al.* (2011) as a normalised colony size distribution, where the authors defined significant mutant phenotypes as colony sizes located within the top fifth (more resistant) or bottom fifth percentile (more sensitive). Shown here are the calculated percentiles from the 3979 *E. coli* strains, showing 25 of 324 "stress" conditions. Significant phenotypes are highlighted in blue with white and bold typeface.

Condition	Target	Concentrations screened	Percentile	
			Δ*tamA*	Δ*tamB*
More resistant				
Actinomycin	RNA biosynthesis	2.5 µg/mL	79.8	78.5
		5.0 µg/mL	56.6	57.7
		10.0 µg/mL	94.4	67.6
		15.0 µg/mL	**95.3**	86.1
Bleomycin	DNA and RNA degradation	0.1 µg/mL	59.1	88.6
		0.5 µg/mL	**97.1**	90.2
		1.0 µg/mL	**99.0**	93.1
		2.0 µg/mL	**97.0**	**98.3**
Novobiocin	DNA gyrase	4 µg/mL	29.4	60.8
		6 µg/mL	37.0	73.0
		8 µg/mL	41.4	86.2
		10 µg/mL	66.4	76.8
		12 µg/mL	69.8	75.2
		30 µg/mL	**96.1**	53.6
More sensitive				
Cholate	Membrane	0.1 % (w/v)	68.9	66.5
		0.5 % (w/v)	**0.3**	44.7
		1.0 % (w/v)	14.7	47.4
		2.0 % (w/v)	35.4	55.0
Dibucaine	Proton motive force	0.4 mM	31.9	8.1
		0.8 mM	15.8	15.0
		1.2 mM	17.4	**2.4**
H_2O_2	Oxidative stress	0.1 mM	**1.3**	70.7
		0.5 mM	16.7	10.1
		1.0 mM	51.5	85.5
		2.0 mM	39.0	20.3

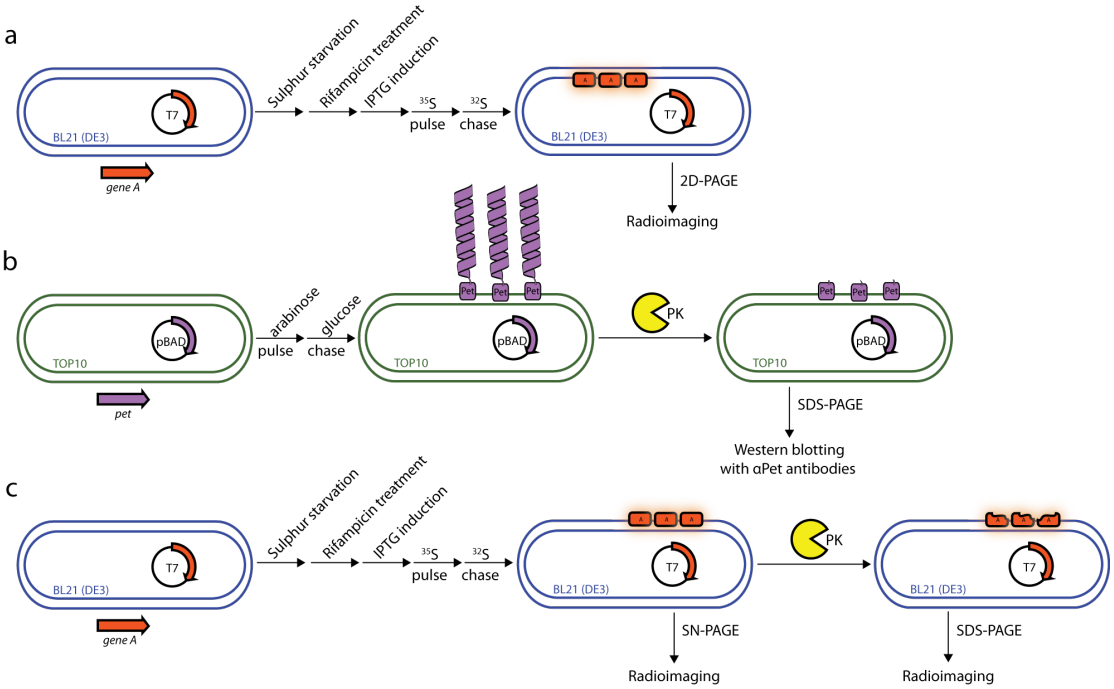


Figure 3.1.1 Schematic of pulse chase techniques.

a, Stenberg *et al.* (2007) developed a pulse chase assay to assess the assembly of the cytochrome bo3 complex at the IM. By radiolabelling the protein of interest with [^{35}S]-methionine, Stenberg *et al.* (2007) could forego the need to raise antibodies specific to their proteins of interest. **b,** The pulse chase method developed by Leyton *et al.* (2014) was used in conjunction with a protease shaving protocol using proteinase K (PK) to assess surface-exposure of Pet. By using the arabinose expression and glucose repression system, Leyton *et al.* (2014) had to use antibodies specific to Pet for detection. **c,** By combining both pulse chase techniques, the advantages of each method could be used to assess the TAM dependence of a range of OMPs.

pathway that may be lost during protein overexpression can be probed more easily.

The radiolabelling technique is beneficial over other detection techniques, because it foregoes the need for developing antibodies to the target protein, or the need to functionally characterise an immunolabelled target protein. Because Leyton *et al.* (2014) had prepared antibodies to their target protein, Pet, they could perform a non-radioactive pulse chase analysis. To ensure they could still monitor Pet's biogenesis pathway (by preventing *pet* overexpression), they used the pBAD system to "pulse" by inducing gene expression with arabinose and "chase" by repressing gene expression via glucose (Figure 3.1.1b). The main benefit of the Leyton *et al.* (2014) method is that they performed an extracellular protease shaving technique to distinguish between mature and intermediate forms of Pet. By combining the benefits of both methods (Figure 3.1.1c), the candidate TAM substrates may be assayed for TAM dependence.

3.2 Candidate proteins for pulse chase analysis

In considering the general properties of TAM substrates, it was clear that there is a general trend for the TAM to be associated with *in vivo* growth, as opposed to *in vitro* growth. Because the TAM is known to assemble AIDA-like autotransporters (Selkrig *et al.*, 2012), which are virulence factors involved in cell-cell adhesion and biofilm formation (Leyton *et al.,* 2012), the TAM may play a general role in assembling various virulence-related β-barrel proteins. By looking to the structures of the AIDA-like autotransporters (Figure 3.2.1a), it has been noted that the TAM might work in concert with the BAM complex to facilitate the insertion of

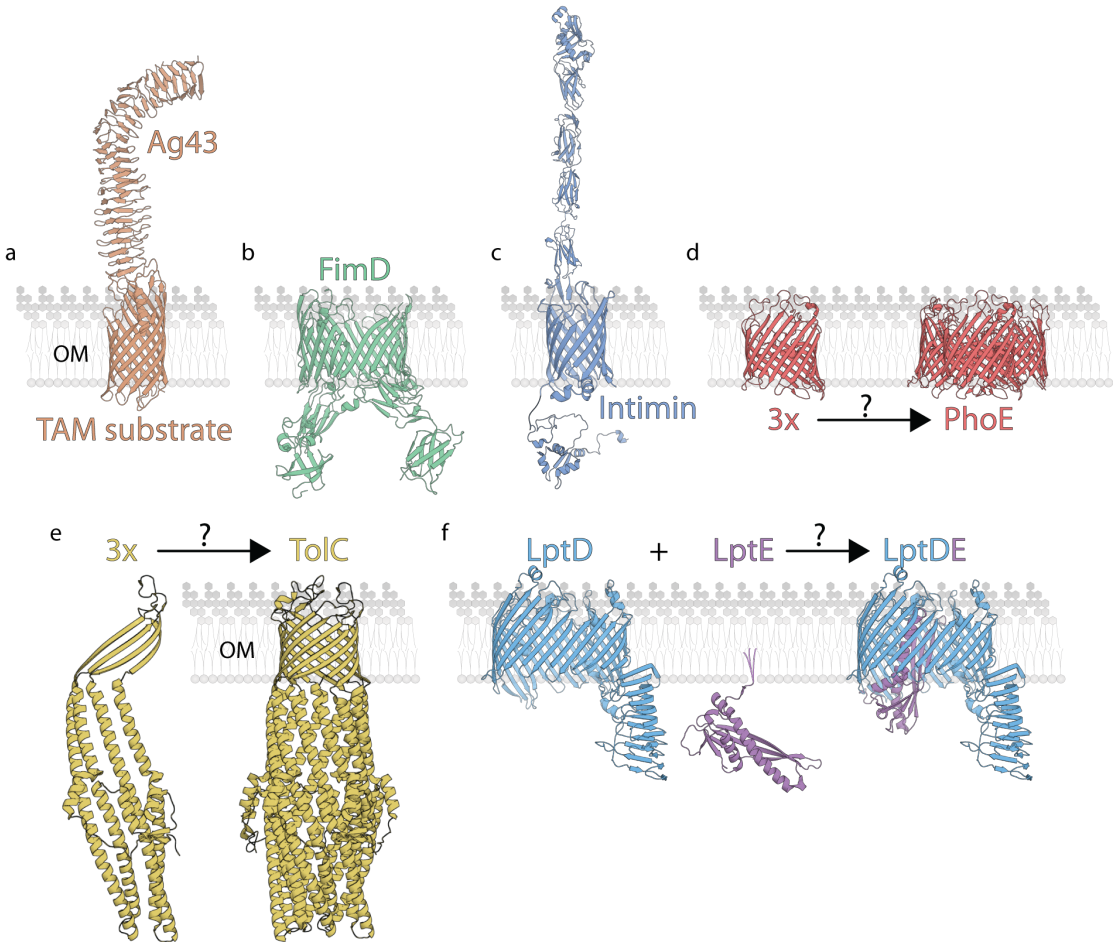


Figure 3.2.1 Known and putative TAM substrates.

Structures of the known (**a**) and six candidate TAM substrates (**b**-**f**), chosen for their complex tertiary (**b**, **c**, **e**, **f**) or quaternary structure (**d**-**f**). **a,** Ribbon diagram of Ag43, comprising an N-terminal extracellular β-helical passenger domain (PDB: 4KH3) secreted through its transmembrane domain (a Phyre2 homology model of residues 705-1039 modelled after the same domain from AIDA, PDB: 4MEE). **b,** Ribbon diagram of FimD (PDB: 4J3O). **c,** Ribbon diagram of intimin, comprising a periplasmic domain tethered to the cell wall (PDB: 2MPW), and a C-terminal extracellular passenger domain (PDB: 1F02 and a Phyre2 homology model of residues 455-653 modelled after the same domain from invasin, PDB: 1CWV) secreted through its transmembrane domain (PDB: 4E1S). **d,** Ribbon diagram of a PhoE monomer (PDB: 1PHO) that assembles into a PhoE trimer (trimers were generated from the monomeric structure using the "symmetry mates" function of Pymol). **e,** Ribbon diagram of a TolC monomer that adopts a trimeric quaternary structure (PDB: 1EK9). **f,** Ribbon diagrams of the lipoprotein - LptE - and a 26-stranded β-barrel - LptD - that "somehow" assemble into the LptDE complex (PDB: 4Q35).

particularly challenging β-barrel proteins (Selkrig *et al.*, 2014). Indeed, an interesting mechanistic feature of the TAM is that it further destabilises the OM following the initial destabilisation caused by the lopsided aromatic girdle of Omp85-family proteins (Noinaj *et al.*, 2013; Selkrig *et al.*, 2015; Shen *et al.*, 2014). This destabilisation presumably corresponds to a further decrease in the activation energy required for OMP biogenesis. Following from this hypothesis, a short list of candidate TAM substrates containing complex non-transmembrane domains or quaternary structure was prepared (Figure 3.2.1b-f) and tested for TAM dependence in chapters 4 and 5.

3.3 *E. coli* strains to be used for pulse chase analysis

The choice of *E. coli* strain is generally important for protein overexpression, especially if the protein is toxic to the cell or otherwise difficult to purify. However, because this assay required only very low amounts of protein, strain selection was less important. As such, there were only two *E. coli* strain requirements necessary for pulse chase analysis. The first was that the strains contain the λDE3 lysogen, which encodes T7 RNA polymerase under LacI repression. In order to utilise the rifampicin-blocking technique, the strains must also be rifampicin-sensitive, which therefore excluded the commonly-used rifampicin-resistant *E. coli* HMS174 (DE3) strain.

Plasmids encoding the λDE3 lysogen could potentially have been used, but because they would have restricted the types of additional plasmids and antibiotic resistance markers, strains with chromosomal λDE3 lysogens were used instead. Two commonly used strains that both encode a chromosomal copy of the λDE3

lysogen and are rifampicin-sensitive are BL21 (DE3) and its derivative: BL21 Star™ (DE3). Both strains are deleted for *lon* and *ompT*, which encoded proteases that could have potentially degraded the protein of interest. BL21 Star™ (DE3) additionally contains the *rne131* allele, which encodes a defective RNase E that cannot degrade mRNA, while retaining its 5S rRNA maturase activity. This causes increased mRNA stability, due to decreased mRNA degradation.

3.4 Plasmids for pulse chase analysis

In considering the plasmids that will be used for expression of the target gene, the only requirement was that the gene be transcribed from a T7 promoter. Additionally, if complementation analyses were required, it was important that two compatible T7 plasmids, or at least a T7 plasmid with two independent multiple cloning sites (MCSs), was used. Fortunately, commercial vectors from Novagen fulfilled these requirements: pET-15b, pET-22b(+), pETDuet-1 and pACYCDuet-1 (Figure 3.4.1). The three plasmids from the pBR322 incompatibility group - pET-15b, pET-22b(+) and pETDuet-1 - contain different MCSs. This meant that if a gene of interest contained an internal restriction site that would preclude its use with one pBR322-based vector, it was compatible with the alternate pBR322-based vectors. The fourth plasmid - pACYCDuet-1 - is from the p15a incompatibility group and was therefore used in conjunction with the pBR322-based plasmids during complementation analyses.

Each of these plasmids contains a variant T7*lac* promoter to reduce background (uninduced) gene expression. Novagen recommends the use of 1 mM IPTG for 2-3 hours for maximum expression from *T7lac* promoters, but considering the

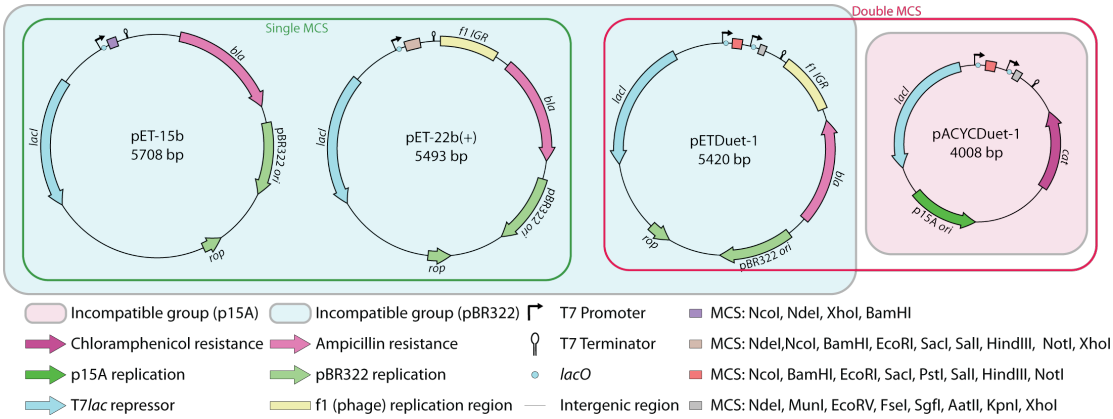


Figure 3.4.1 Base vectors used in this study.

Schematics of each base vector to scale. The major differences include the different resistance markers, MCSs and plasmid replication systems. Although the f1 intergenic region (IGR) responsible for ssDNA replication and packaging into phage particles is depicted, only the p15A and pBR322 dsDNA replication systems are relevant to this study.

purpose of the assay was to monitor protein biogenesis, induction was reduced to a 5-minute incubation with 0.2 mM IPTG.

3.5 Culture media to be used during pulse chase analysis

By using the EXPRE^{35}S^{35}S [^{35}S]-Protein Labelling Mix (Perkin Elmer, NEG072) comprising 73% [^{35}S]-methionine and 22% [^{35}S]-cysteine, it was possible to radiolabel the proteins of interest. To maximise incorporation of ^{35}S-methionine and -cysteine, in place of the naturally abundant ^{32}S-isotopic forms, it was important to starve the bacteria of sulphur. This was achieved by preparing a defined minimal media, such as an M9 minimal medium (Stenberg *et al.*, 2007) that excludes the sulphur-containing amino acids, but also free sulphates, which are important for cysteine and methionine biosynthesis (Sekowska *et al.*, 2000).

Depending on the application, M9 minimal media composition can greatly differ, but they all contain M9 salts (CSH Protocols, 2006b). M9 salts are comprised of an $HPO_4^{2-}/H_2PO_4^-$ buffer, sodium chloride for osmotic support, and ammonium chloride as an important source of nitrogen, especially for amino acid synthesis (CSH Protocols, 2006b). M9 minimal media is also commonly supplemented with a carbon source, such as glucose, and; calcium chloride and magnesium sulphate to enhance growth (Ca^+ and Mg^{2+}) and to provide an appropriate source of sulphur (SO_4^{2-}) (CSH Protocols, 2006b; 2010). Vitamins may also be added to support growth, especially if the organism is auxotrophic for a particular vitamin, or the intention is to reduce reagents required for vitamin biosynthesis.

In this study, two formulations of M9 media were required: one deficient in

sulphur to aid in sulphur starvation (M9-S media), and one containing sulphur to act as the "chase" medium (M9+S media). In designing M9 media, it was thought that by exchanging ammonium chloride for a defined amino acid mixture, such as the amino acid supplement powder without adenine, uracil, inositil and vitamin B_{10} (CSH Protocols, 2016), growth would be supported by reducing the need for amino acid synthesis.

Although only trace amounts of vitamins are required, one sulphur-containing vitamin, thiamine, requires cysteine for its biosynthesis (others typically use iron-sulphur clusters) (Leonardi & Roach, 2004; Sekowska *et al.*, 2000), so supplementation with thiamine should support growth in the absence of cysteine. M9+S media was therefore based on the standard M9 minimal medium formulation (CSH Protocols, 2010) and is comprised of: 47.8 mM Na_2HPO_4, 22 mM KH_2PO_4, 8.56 mM NaCl, 11.1 mM glucose, 1.12 mM thiamine, 1 mM $MgSO_4$, 0.1 mM $CaCl_2$, 45.4 pg/mL of standard amino acids, except 227 pg/mL for leucine. Whereas M9-S media was similar to M9+S media, except that methionine and cysteine were excluded, and 1 mM $MgCl_2$ was used instead of $MgSO_4$.

Considering M9-S media does not contain all the necessary nutrients to support *E. coli* growth, prolonged incubation times with M9-S media should be avoided. Instead, *E. coli* was incubated in LB media to the desired optical density before resuspension in M9-S media for sulphur starvation. The next series of steps, which include rifampicin treatment and introduction of radioactivity, are obviously detrimental to *E. coli* cells that have been transferred from rich media (LB) to minimal media (M9-S). Therefore, *E. coli* was incubated for a single doubling time

(30 minutes) before addition of rifampicin, to allow *E. coli* to adapt to M9-S.

With regards to the specific order of steps in the pulse chase assay, the method of Stenberg *et al.* (2007) was used as a guide, but the specific concentrations, temperatures and timing was then optimised for OMP synthesis and targeting. In considering the materials available and protocols for similar pulse chase analyses of OMP biogenesis (Crowlesmith *et al.*, 1981; Jansen *et al.*, 2000; Leyton *et al.*, 2014; Misra *et al.*, 1991; Pages & Bolla, 1988), the following steps were performed before optimisation was required.

After the initial 30-minute incubation in M9-S media, cells were incubated for a further 60 minutes in 200 µg/mL of rifampicin to block native *E. coli* transcription and to ensure sufficient sulphur starvation. A five-minute induction with 0.2 mM IPTG followed to ensure transcription of both T7 RNA polymerase and the plasmid-based target genes, before a 40-second "pulse" with EXPRE^{35}S^{35}S [^{35}S]-Protein Labelling Mix (Perkin Elmer, NEG072), corresponding to 22 µCi/mL of radioactivity.

Several pulse chase methods then require the “pulsed” cells to be "chased" with excess methionine and cysteine, but initial experiments showed that removing the IPTG and unincorporated radioactivity was important. This was achieved by recovering cells by centrifugation at low speeds, and removing the supernatant containing IPTG and unincorporated radioactivity. Cells were then resuspended in M9+S media in order to "chase" the biogenesis pathway of the target protein and aliquots were taken at a range of "chase" timepoints for subsequent analysis

to determine whether protein biogenesis was affected in the absence of the TAM.

3.6 Analysis of radiolabelled proteins

The list of target proteins includes monomeric β-barrel proteins with complex periplasmic or extracellular domains and β-barrel proteins that adopt quaternary structure (Figure 3.2.1b-f). Whether the TAM facilitates the folding of non-transmembrane domains or the assembly of quaternary structures, was assessed by utilising one of two techniques to probe β-barrel biogenesis. For multimeric β-barrel proteins, multimerisation was analysed by SN-PAGE to determine if there was any defect or change in the quaternary structure assembly rate. For monomeric β-barrel proteins, extracellular protease shaving using proteinase K as per Leyton *et al.* (2014) was performed as a means of measuring the surface-exposure of target proteins.

Regardless of the analysis technique used, central to the pulse chase method is the use of the [^{35}S]-methionine and -cysteine to radiolabel the proteins of interest. Storage phosphor screens are capable of retaining the energy produced by ionising radiation (including β^- rays), and upon laser-induced stimulation, light directly proportional to the amount of radioactivity in the sample is emitted. The image produced can therefore be used to calculate the relative abundance of a target protein, and subject to densitometrical analyses, was used to calculate protein folding rates.

3.7 Optimising conditions for analysis with *E. coli* BL21 Star™ (DE3)

Because BL21 Star™ (DE3) causes increased mRNA stability, it was important to

determine the "background" amount of radiolabelled protein due to translation of residual mRNA after rifampicin treatment. The amount of native radiolabelled protein was therefore determined (Figure 3.7.1a), and not surprisingly, addition of 200 μg/mL rifampicin significantly reduced the number of radiolabelled proteins in BL21, BL21 (DE3) and BL21 Star™ (DE3). While radiolabelled proteins in BL21 and BL21 (DE3) were comparable, BL21 Star™ (DE3) had a higher "background" of [^{35}S]-incorporated proteins. Because residual protein levels across the three strains were comparable, as adjudged by the relatively equal amounts of BamA and DnaK present, the higher "background" for the BL21 Star™ (DE3) strain is most likely due to increased mRNA stability as a direct result of the *rne131* allele.

To reduce "background" levels of radiolabelled protein observed in BL21 Star™ (DE3), rifampicin concentration was increased and/or chase temperature was decreased (Figure 3.7.1b). When rifampicin concentration was increased from 200 μg/mL to 500 μg/mL, there was a slight but discernible drop in the level of background labelling. However, a greater decrease was observed when the chase temperature was reduced from 37 °C to 30 °C, possibly due to the lower translation efficiency at suboptimal temperatures. Choosing the correct chase temperature may depend on the target protein or be useful for capturing intermediates, but as a starting point: subsequent pulse chase analyses with BL21 Star™ (DE3) were performed using 500 μg/mL rifampicin and a 30 °C chase temperature.

The background level of radiolabelled proteins in BL21 Star™ (DE3) harbouring

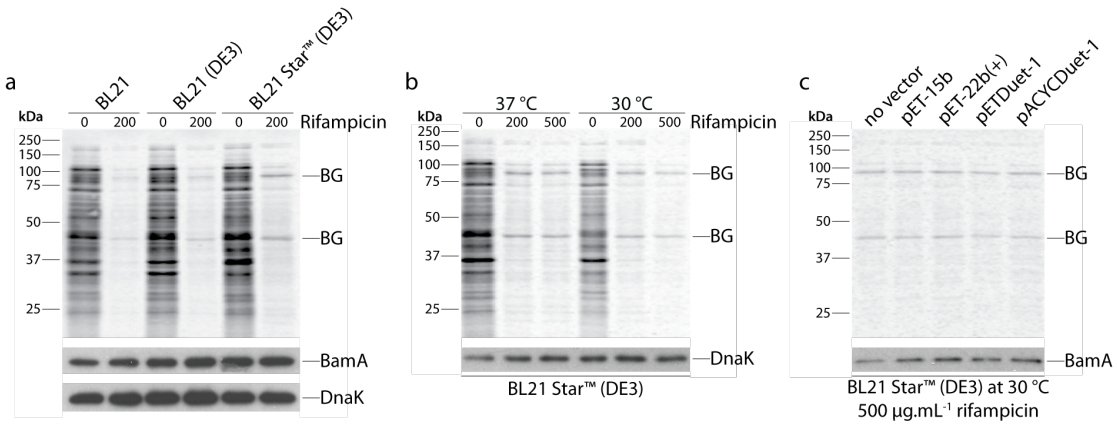


Figure 3.7.1 Optimising pulse chase conditions for *E. coli* strains and plasmids.

Following routine incubation in LB media to OD_{600} of 0.6, the indicated *E. coli* strains with (**c**) or without (**a**-**b**) the indicated plasmids were starved of methionine and cysteine by incubation in M9-S media (200 rpm [25 mm orbit], 30 min, 37 °C). Cells were treated with 0, 200 or 500 µg.mL^{-1} of rifampicin (200 rpm [25 mm orbit], 1 hour, 37 °C) as indicated, and then induced with 0.2 mM IPTG (static, 5 min) at the indicated temperature. Cells were "pulsed" with 22 µCi.mL^{-1} of a [^{35}S]-methionine/-cysteine mixture (NEG072, Perkin Elmer) (static, 40 sec) at the indicated temperature and then subjected to centrifugation (3000 **rcf**, 5 min, 4 °C). Proteins were "chased" by resuspending the cell pellet in M9+S media (static, 10 min) at the indicated temperature where aliquots were taken and boiled for 3 min in SDS sample buffer. Analysis was by SDS-PAGE, storage phosphorimaging and Western blotting using rabbit antibodies specific for BamA or DnaK as indicated. Sizes in kDa are indicated on the left. Prominent background (BG) bands are indicated on the right. The strains, rifampicin concentration and chase temperatures used are indicated above and/or below the autoradiograms.

one of the base vectors depicted in Figure 3.4.1 was determined (Figure 3.7.1c). There was no obvious difference between the amounts of radiolabelled protein in BL21 Star™ (DE3), regardless of whether it contains one of the four plasmids or not. While it is not clear what the identities of the background radiolabelled proteins were, they likely arise from translation of the prominent mRNA species present prior to rifampicin treatment. Based on the longest mRNA half-lives reported by Taniguchi *et al.* (2010), the 40-45 kDa band could be LpxB (43 kDa) and/or OmpC (40.4 kDa), and the 90-100 kDa band could be AcnB (94 kDa) and/or AceE (99.7 kDa).

3.8 Proof-of-principle with pulse chase analysis of FimD biogenesis

In the early stages of exploring whether the pulse chase technique discussed so far was feasible, *fimD* was selected from the list of candidate proteins in a proof-of-principle analysis. FimD is the usher component of the chaperone-usher secretion system that produces type 1 fimbriae, where each fimbrial subunit is secreted through the FimD pore (Phan *et al.*, 2011). FimD is one of the largest β-barrel proteins (24 β-strands), and contains several complex periplasmic domains (Figure 3.8.1), which are important for substrate binding and/or occluding the pore in the absence of substrate (Geibel & Waksman, 2014).

FimD biogenesis was assessed by protease shaving and it was noted that the fragmentation profile was dependent on the presence or absence of the TAM (Figure 3.8.2). When wildtype cells were exposed to extracellular PK, the portion of 90 kDa FimD assembled in the OM was cleaved into ~50 kDa and ~40 kDa fragments, respectively referred to as fragments A and C. Whereas, in the absence

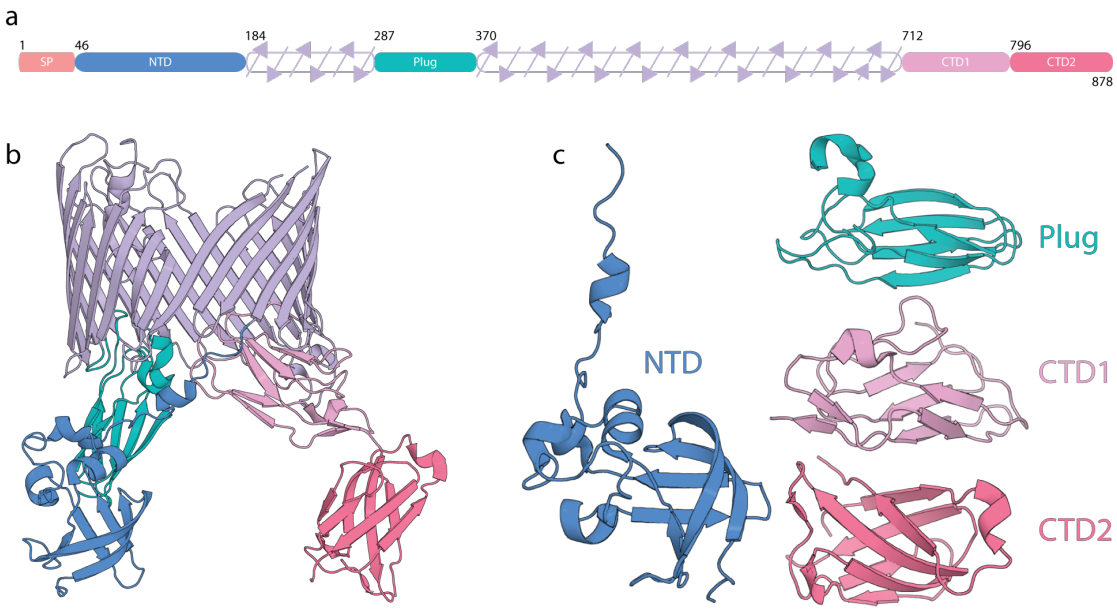


Figure 3.8.1 FimD structure.

a, Structural map of FimD, based on its crystal structure (PDB: 3RFZ). **b,** Ribbon diagram of FimD. **c,** Ribbon diagrams of the four periplasmic domains of FimD in isolation, comprising: the N-terminal domain (NTD), plug domain and two C-terminal domains (CTD1 and CTD2).

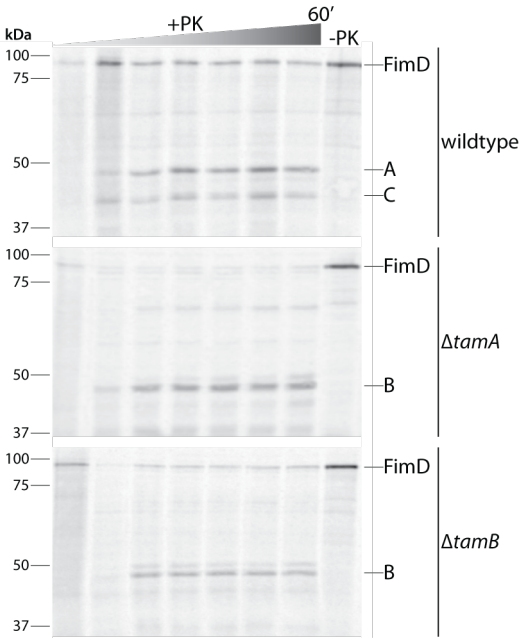


Figure 3.8.2 FimD biogenesis in the presence and absence of the TAM.

FimD biogenesis in the indicated *E. coli* BL21 Star™ (DE3) strains harbouring pKS02 (pET-15b with *fimD* encoded) was assessed by pulse chase analysis. Cells were incubated to OD_{600} of 0.6 in LB media before sulphur starvation in M9-S media (200 rpm [25 mm orbit], 30 min, 37 °C). Cells were then treated with 500 μg.mL^{-1} of rifampicin (200 rpm [25 mm orbit], 1 hour, 37 °C), induced with 0.2 mM IPTG (static, 5 min, 30 °C) and "pulsed" with 22 μCi.mL^{-1} of a [^{35}S]-methionine/-cysteine mixture (NEG072, Perkin Elmer) (static, 40 sec, 30 °C). Following centrifugation (3000 **rcf**, 5 min, 4 °C), cell pellets were resuspended in M9+S media and proteins were "chased" for 60 minutes. At 10 seconds, 2, 5, 10, 15, 30 and 60 minutes, aliquots were taken and treated with (+PK) or without (-PK, last timepoint only) 50 μg mL^{-1} proteinase K (static, 10 min, on ice). Proteins were then TCA-precipitated, washed with acetone and boiled for 3 minutes in SDS sample buffer. Analysis was by SDS-PAGE and storage phosphorimaging. Sizes in kDa are indicated on the left, and full-length FimD and its fragments A, B, and C are indicated on the right. The time increment is indicated above the autoradiograms as a graded triangle and the strains are indicated to the right of the autoradiograms.

of either TamA or TamB, FimD was apparently cleaved into at least one 45 kDa fragment, indicated as fragment B. While the TAM-dependence of FimD is explored and confirmed in Chapter 4, this initial experiment indicates that, in principle, the pulse chase method is a beneficial tool for identifying novel TAM substrates.

By analysing the biogenesis of FimD in wildtype *E. coli*, several aspects of the pulse chase assay were optimised. For storage phosphorimage analysis, radiation may be captured directly from the polyacrylamide gel or from radioactive membranes following Western transfer (Figure 3.8.3). However, it is clear that radiation from nitrocellulose membranes is more effectively captured than from radioactive polyacrylamide gels, regardless of whether the gels were stained with Coomassie staining solution or not (Figure 3.8.3a-b). Whereas radiation from nitrocellulose membranes or PVDF membranes was comparable when analysed by storage phosphorimaging (Figure 3.8.3c).

Stenberg *et al.* (2007) used LB medium during the "chase" phase of their pulse chase analysis of inner membrane proteins, so FimD biogenesis was assessed by pulse chase analysis using either LB or M9+S chase media (Figure 3.8.4a). Although a small proportion of full-length FimD was observed when LB chase medium was used (Figure 3.8.4a, top panel: "-PK lane"), few (if any) bands corresponding to FimD or its fragments were detected when the sample was subjected to protease shaving. Not surprisingly, when M9+S chase medium was used instead (Figure 3.8.4a, lower panel) biogenesis of FimD could be monitored.

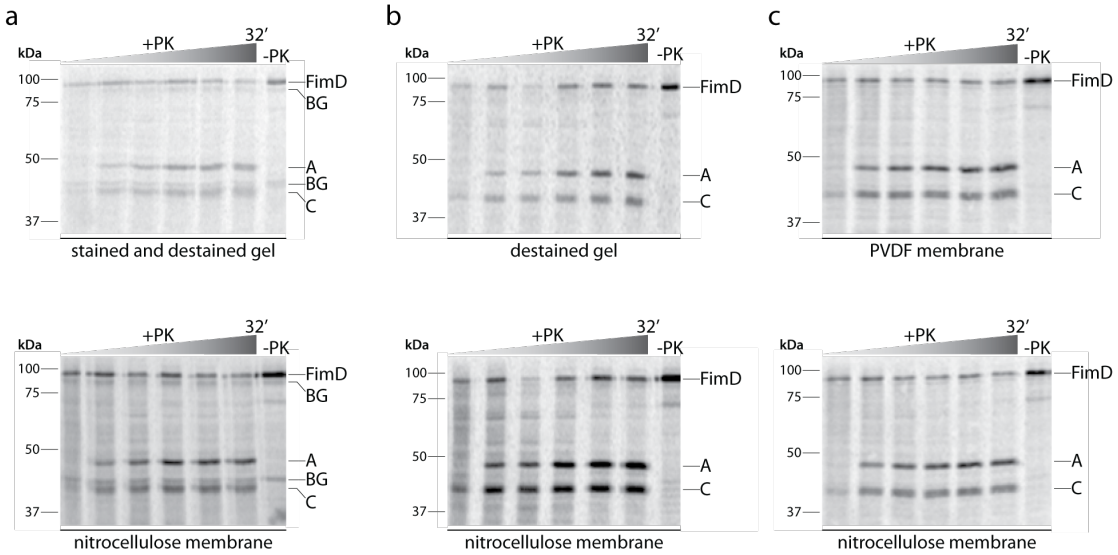


Figure 3.8.3 Radiation from membranes is captured more effectively than from radioactive polyacrylamide gels.

FimD biogenesis in *E. coli* BL21 Star™ (DE3) wildtype cells harbouring pKS02 (pET-15b with *fimD* encoded) was assessed by pulse chase analysis as per Figure 3.8.2. At 10 seconds, 2, 4, 8, 16 and 32 minutes, aliquots were taken and treated with (+PK) or without (-PK, last timepoint only) 50 μg mL^{-1} proteinase K (static, 10 min, on ice). Proteins were then TCA-precipitated, washed with acetone and boiled for 3 minutes in SDS sample buffer. Samples were applied in duplicate to a polyacrylamide gel and analysed by SDS-PAGE. The polyacrylamide gels were cut in half and the proteins in one gel-half were transferred onto nitrocellulose membranes (lower panels), whereas the other gel-half was either: incubated with (**a**) or without (**b**) Coomassie staining solution, destain solution (**a**-**b**) and gel drying solution (**a**-**b**) (40 rpm [32 mm orbit], 1 hour each), or; subjected to Western transfer onto PVDF membrane. Each dried gel/membrane or membrane/membrane pair were simultaneously analysed by storage phosphor imaging (i.e., analysis was performed for the same length of time in the same storage phosphorimaging cassette) and the same level adjustments were applied to the pair using Adobe Photoshop. Sizes in kDa are indicated on the left. Full-length FimD, its fragments A and C, and prominent background (BG) bands are indicated on the right. The time increment is indicated above the autoradiograms as a graded triangle and the radiation "source" is indicated below the autoradiograms.

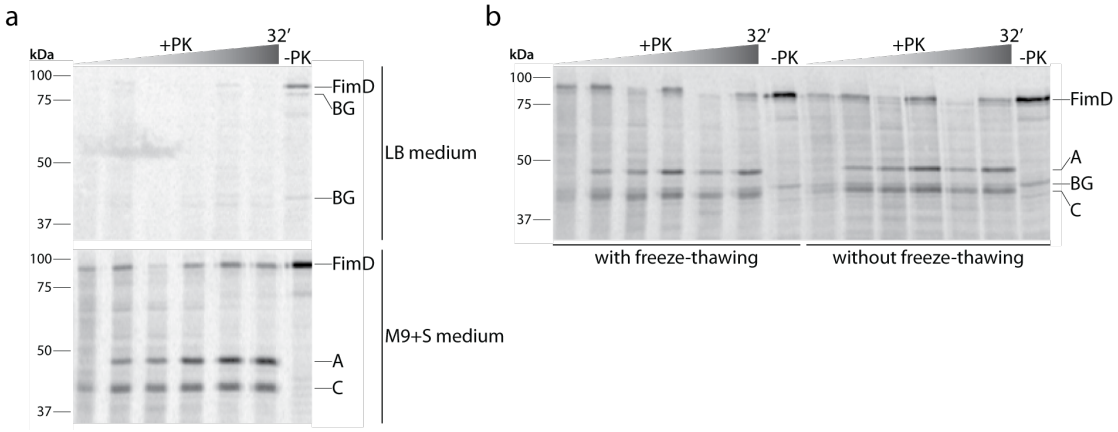


Figure 3.8.4 Optimisation of the pulse chase protocol.

FimD biogenesis in *E. coli* BL21 Star™ (DE3) wildtype cells harbouring pKS02 (pET-15b with *fimD* encoded) was assessed by pulse chase analysis. Aliquots were taken at 10 seconds, 2, 4, 8, 16 and 32 minutes, and treated with (+PK) or without (-PK, last timepoint only) 50 µg mL^{-1} proteinase K (static, 10 min, on ice). Proteins were then TCA-precipitated, washed with acetone and boiled for 3 minutes in SDS sample buffer. Samples were analysed by SDS-PAGE and storage phosphor-imaging. Sizes in kDa are indicated on the left. Full-length FimD, its fragments A and C, and prominent background (BG) bands are indicated on the right. The time increment is indicated above the autoradiogram as a graded triangle. **a,** Pulse chase analysis was performed simultaneously using either LB or M9+S "chase" media, as indicated to the right of the autoradiogram. The same level adjustments were applied to both autoradiograms using Adobe Photoshop and except for the level adjustments, the "M9+S" panel is identical to the "nitrocellulose" panel from Figure 3.8.3b. **b,** Pulse chase analysis was performed with or without freeze thawing as indicated to below the autoradiogram. The same batch of cells were either: (i) snap frozen with 20 % v/v glycerol in liquid nitrogen following sulphur starvation, (ii) or kept on ice. Snap-frozen cells were thawed on ice, subjected to centrifugation (3000 **rcf**, 5 min, 4 °C) and the cell pellet was resuspended in M9-S media. The subsequent steps were then performed simultaneously on both sample types.

Although the current pulse chase protocol allows analysis of up to eight samples concurrently (with a staggered time course), preparing "pulse-chase-ready" samples in bulk and snap-freezing them in liquid nitrogen for storage at -80 °C was thought to be more efficacious. This is because multiple technical replicate samples could be prepared, while minimising the number of incubation days required for sample preparation. The best time to snap-freeze the sample was immediately before rifampicin treatment and after the majority of cells have adapted from LB media to M9-S media.

Glycerol was also added before snap-freezing samples to maintain cell viability and prevent membrane damage and removed when cells were resuspended in M9+S chase medium. There were no obvious differences between samples prepared with or without snap-freezing, as adjudged by similar amounts and rates of FimD biogenesis (Figure 3.8.4b), so either method may be used. If the snap-freezing step was included, cells from the same batch were considered to be technical replicates; whereas, cells from different batches were considered to be biological replicates.

3.9 Conclusion

The TAM has been evolutionarily conserved together with the BAM complex as important machinery for assembling β-barrel proteins (Heinz & Lithgow, 2014; Heinz *et al.*, 2015). At the commencement of this project, only the AIDA-like autotransporters were known to be TAM substrates, but considering not all TAM-containing bacteria harbour autotransporters (Celik *et al.*, 2012; Heinz & Lithgow, 2014), it is difficult to ascribe a general role for the TAM in the context its

evolutionary purpose. In considering the previous techniques used to dissect a role for the TAM, a pulse chase assay was developed to determine novel TAM substrates. The usefulness of this assay was then subsequently confirmed in a proof-of-principle experiment to assess FimD biogenesis in the presence or absence of the TAM. Chapters 4 and 5 subsequently make use of the pulse chase assay to characterise the assembly of novel TAM substrates.

Chapter 4 – Fimbrial ushers are substrates of the TAM

4.1 Introduction

Fimbriae are proteinacious cell-surface appendages on bacteria that facilitate adherence to animal cells, abiotic surfaces and other bacteria. They were first observed in 1949 (Anderson, 1949) by use of electron microscopy and were described as “filaments”. In 1955, Duguid *et al.* named these filaments “fimbriae”, from the Latin word for fibres or threads and in 1959, Brinton suggested they be renamed to “pili”, from the Latin word for hair or fur. Both “fimbriae” and “pili” were used as umbrella terms to refer to all non-flagella appendages, and are now used interchangeably (Nuccio & Bäumler, 2007).

There are four major fimbrial types: conjugation pili, type 4 pili, curli fimbriae and chaperone-usher fimbriae. Conjugation pili (also known as sex pili) are assembled by the T4SS pathway (Cabezón *et al.*, 2015) and are used to transfer genetic material between a donor and recipient cell (Figure 4.1.1a). Type 4 pili may be further divided into type 4a or type 4b pili based on the composition of their subunits, but in both cases they are assembled by the T2SS pathway (Giltner *et al.*, 2012) where they are responsible for a bacterium's twitching motility, a flagella-independent form of motility (Figure 4.1.1b). Curli fimbriae (also known as thin, "aggregative fimbriae") are assembled by the type 8 secretion system pathway (Evans & Chapman, 2014) where they are important for surface attachment and cell-cell aggregation, ultimately leading to biofilm formation (Figure 4.1.1c). Chaperone-usher fimbriae (also known as ‘common pili’) mediate attachment to cells or other surfaces (Figure 4.1.1d) and assemble via the eponymous chaperone-usher secretion pathway (Lillington *et al.*, 2015).

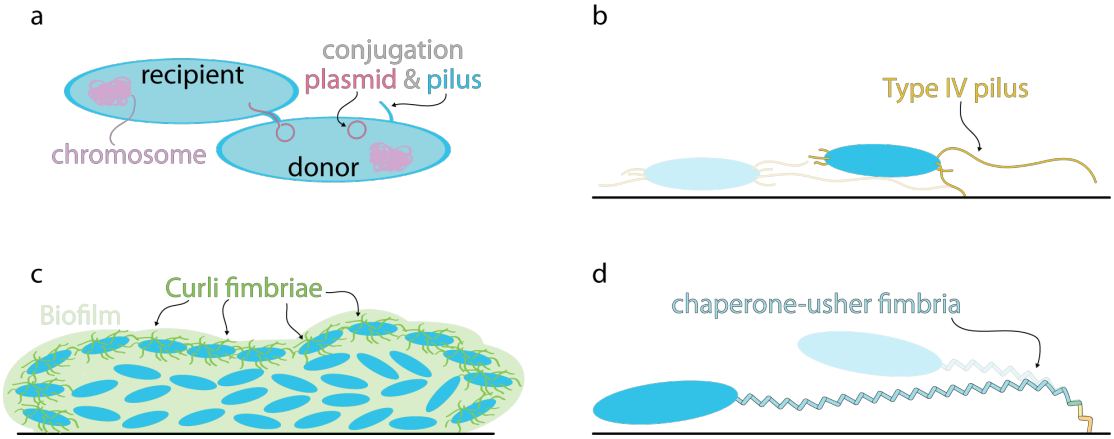


Figure 4.1.1 The four major fimbrial families.

A bacterium (blue) may produce one of the four major fimbriae as indicated, which are important for cell-cell attachment and/or surface (thick black line) adhesion. **a,** Conjugative pili are responsible for the dissemination of genetic material (i.e. the conjugative plasmid) between a donor and recipient cell, usually to provide a benefit, including antibiotic resistance or xenobiotic tolerance (Ilangovan *et al.*, 2015). **b,** The general twitching motility mechanism is illustrated. Type 4 pili extend toward a surface from the pole of the cell and following attachment, the bacterium is propelled to the attachment site by pilus retraction (Giltner *et al.*, 2012). **c,** Curli fimbriae are produced by a subset of cells within a bacterial population that will ultimately cause a biofilm layer to form that will protect the bacterial population. **d,** Chaperone-usher fimbriae are normally distributed peritrichously, but for clarity, a single type 1 fimbria is illustrated. Chaperon-usher fimbriae are important for colonisation of the urinary tract where a fimbria may uncoil and recoil in response to flushing mechanism of urine (Hospenthal *et al.*, 2016).

A given species of bacteria may encode a number of chaperone-usher fimbriae (Spurbeck *et al.*, 2011) that need to be tightly regulated. Differential expression of fimbriae may be important for different stages or types of infection, but the sheer number of fimbrial clusters encoded by bacteria suggests a functional redundancy. Indeed, type 1, P, S, F1C and Dr fimbriae all allow adhesion to human kidney epithelia, although they recognise different receptor molecules (Korea *et al.*, 2011). By comparing the sequences of various usher proteins, Nuccio and Bäumler (2007) developed a classification system that distributed fimbriae into six phylogenetic clades: α (alternate chaperone-usher fimbriae), β, γ (type 1 fimbriae-like), κ (flexible fimbriae), π (P fimbriae-like) and σ (archaic chaperone-usher fimbriae), where the γ clade were further divided into four sub-clades: γ1-γ4. One of the best-studied examples of chaperone-usher fimbriae are the type 1 fimbriae (γ1 clade) from *E. coli* encoded by the *fim* locus (Figure 4.1.2a).

During type 1 fimbrial biogenesis, the "chaperone" component (FimC) transports fimbrial subunits to the "usher" component (FimD), which is a large β-barrel OMP responsible for assembling and secreting the pilus (Lillington *et al.*, 2015) (Figure 4.1.2b). The actual fimbrial appendage is comprised of 500-3000 copies of the major fimbrial subunit (FimA) and a single copy of each minor subunit (Hahn *et al.*, 2002). The minor subunits are located at either the base or tip of the shaft. At the base is a terminator subunit that: (i) prevents additional subunit polymerisation and (ii) locks the fimbriae onto the usher (FimI) (Bečárová, 2015); at the tip, adaptor subunits (FimFG) connect the shaft to the tip adhesin (FimH) (Lillington *et al.*, 2015). The tip adhesin contains two functional domains: an immunoglobulin-like fold and a c-type lectin domain (Sauer *et al.*, 2016). While

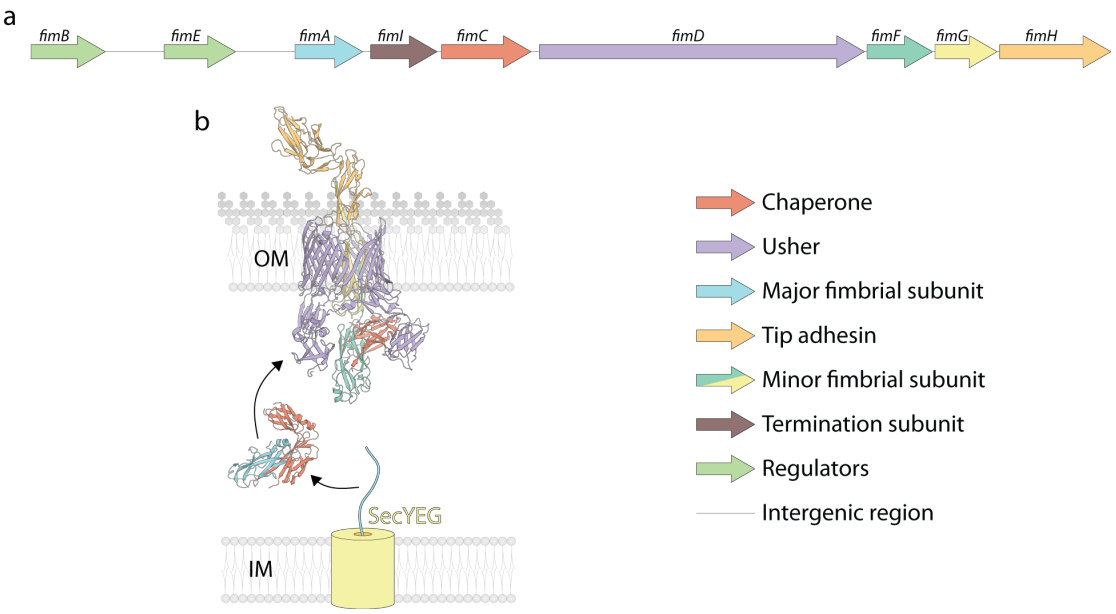


Figure 4.1.2 Type 1 fimbriae.

a, The genetic organisation of the nine *fim* genes from *E. coli* K-12 MG1655 (locus tags b4312-b4320). **b,** Illustration of the chaperone-usher pathway, using biogenesis of type 1 fimbriae as an example. For clarity, fimbrial subunit oxidation by the Dsb proteins was omitted. Following Sec translocation, fimbrial subunits are recognised by fimbrial chaperone (FimC) and transferred to the usher component (FimD) for subsequent secretion. Fimbrial subunits are added sequentially to the growing pilus, starting with the minor subunits and followed by hundreds to thousands of copies of the major fimbrial subunit. Shown are ribbon diagrams of each fimbrial subunit at the initial stages of type 1 fimbrial biogenesis. The pilus is comprised of two subunits where FimD has assembled FimGH and has primed FimF for insertion into the pilus by binding a FimC:FimF complex (PDB: 4J3O for the FimCDFGH complex). Once FimD binds the FimC:FimA complex (PDB: 4DWH): FimF will be inserted into the pilus, FimC will be ejected, and FimA will be primed for insertion.

different tip adhesins contain similar c-type lectin domains, which are responsible for carbohydrate binding, FimH itself recognises α-D-mannosylated glycoproteins, especially the terminal mannose residue of uroplakins Ia and Ib found throughout the luminal surface of the bladder epithelium (Wu *et al.*, 1996; Zhou *et al.*, 2001). All fimbrial subunits are required to produce functional fimbriae: without FimA, no fimbriae are formed; without FimI, fimbriae become too long and are completely excreted from the cell; without the tip adhesin or adaptor subunits, fimbriae can no longer adhere to cell surfaces (Båga *et al.*, 1987; Lindberg *et al.*, 1987; Maurer & Orndorff, 1987).

Structural analysis of each fimbrial subunit reveals they contain an "incomplete" immunoglobulin-like fold (Geibel *et al.*, 2013; Habenstein *et al.*, 2015): two stacked three-stranded β-sheets in a β-sandwich conformation (Figure 4.1.3a). However, "complete" immunoglobulin-like folds are comprised of a 7-9-stranded β-sandwich (Bork *et al.*, 1994). To date, literature reports suggest that the fimbrial subunits adopt a constant 1 type immunoglobulin fold (see for example: Geibel *et al.*, 2013; Hospenthal *et al.*, 2016; Phan *et al.*, 2011; Puorger *et al.*, 2008) (Figure 4.1.3b), but closer inspection of strand 'D' revealed an as yet unreported topology (Figure 4.1.3c). In this "pilin fold", the stretch of amino acids representing strand D is instead split into two β-strands (D1 and D2), where β-strand D1 forms part of the first β-sheet (adjacent to β-strand E) and β-strand D2 forms part of the second β-sheet (adjacent to β-strand C) (Figure 4.1.3d-g).

For each fimbrial subunit: the absence of the seventh β-strand in the "incomplete" immunoglobulin-like domain leaves a series of five solvent-accessible

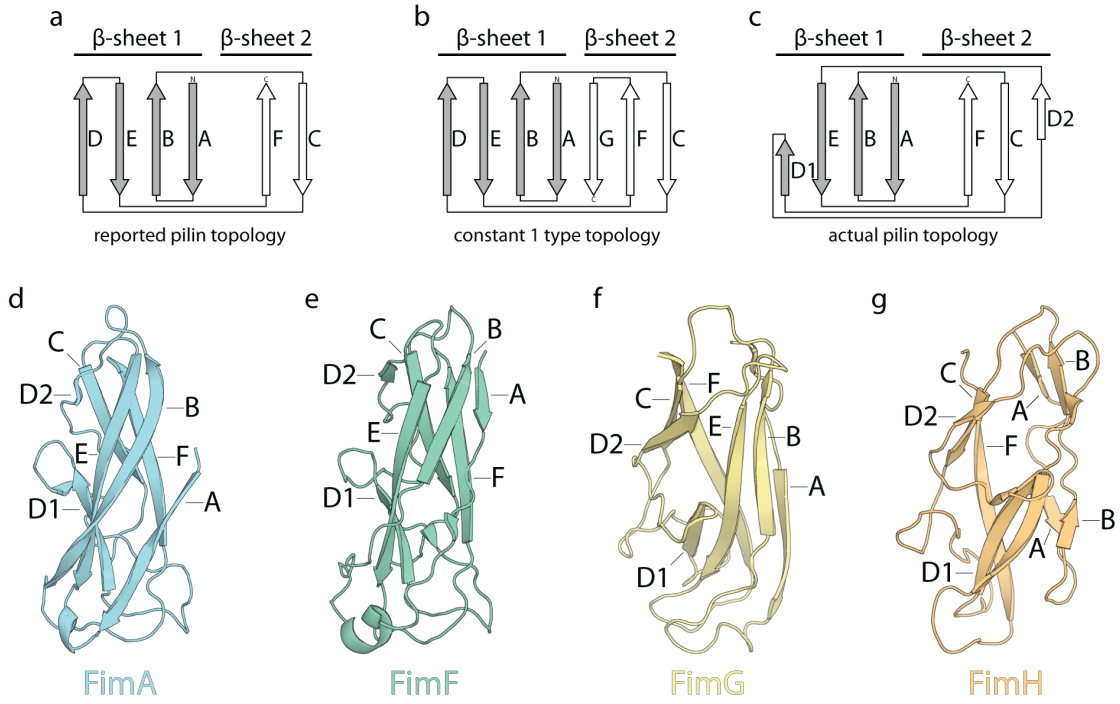


Figure 4.1.3 Fimbrial subunits adopt an unusual immunoglobulin-like fold.

Immunoglobulin-like folds are 7-9-stranded β-sandwich domains important for interactions with other proteins. Illustrations of the immunoglobulin topologies for: **a,** the reported "incomplete" pilin; **b,** the "complete" constant 1 type immunoglobulin fold "**a**" was derived from, and; **c,** the observed "incomplete" pilin fold. Arrows represent β-strands, where the first and second β-sheets are comprised of grey and white arrows, respectively. Ribbon diagrams of the "incomplete" immunoglobulin-like domain observed in type 1 fimbrial subunits, where each β-strand has been labelled as per "**c**": **d,** residues 28-159 of mature FimA (PDB: 2N7H, chain B); **e,** residues 15-154 of mature FimF (PDB: 4J3O, chain F); **f,** residues 16-144 of mature FimG (PDB: 4J3O, chain G), and; **g,** residues 160-279 of mature FimH (PDB: 4J3O, chain H).

hydrophobic pockets (P1-P5) (Busch & Waksman, 2012). Initially, FimC - itself comprised of two immunoglobulin-like domains with switch type topology (Figure 4.1.4) - contributes its G1 β-strand (β-strand G from its first domain) to "complete" a fimbrial subunit's immunoglobulin-like fold in a process known as "donor strand complementation" (Busch & Waksman, 2012) (Figure 4.1.5). In the now "completed" pilin fold, the G1 donor β-strand runs parallel to the acceptor's β-strand F to protect the P1-P4 hydrophobic pockets, but not the P5 pocket, from solvent access (Busch & Waksman, 2012). Leaving the P5 pocket unoccupied is an important aspect for pilus biogenesis during donor strand exchange, where the G1 donor β-strand is subsequently replaced by an N-terminal extension from the incoming fimbrial subunit, which binds to the unoccupied P5 pocket and then sequentially displaces the chaperone's G1 β-strand in a zip-in-zip-out mechanism (Busch & Waksman, 2012).

Once donor strand exchange between the two subunits has completed, the N-terminal extension from the incoming subunit runs anti-parallel to the acceptor subunit's β-strand F (Figure 4.1.5). The energy to drive donor strand exchange is thought to derive from the significantly energetically favourable conformational change that occurs during the transition from a parallel to anti-parallel donor strand-strand F conformation (Puorger *et al.*, 2008; Zavialov *et al.*, 2005). Eventually, the termination subunit (FimI) - which does not contain a P5 pocket for subsequent incorporation of a fimbrial subunit - will be incorporated into the pilus, thereby completing pilus biogenesis (Busch & Waksman, 2012). FimD is a 24-stranded β-barrel with four periplasmic domains that are indispensable for pilus biogenesis: an N-terminal domain (NTD), a plug domain and two C-terminal

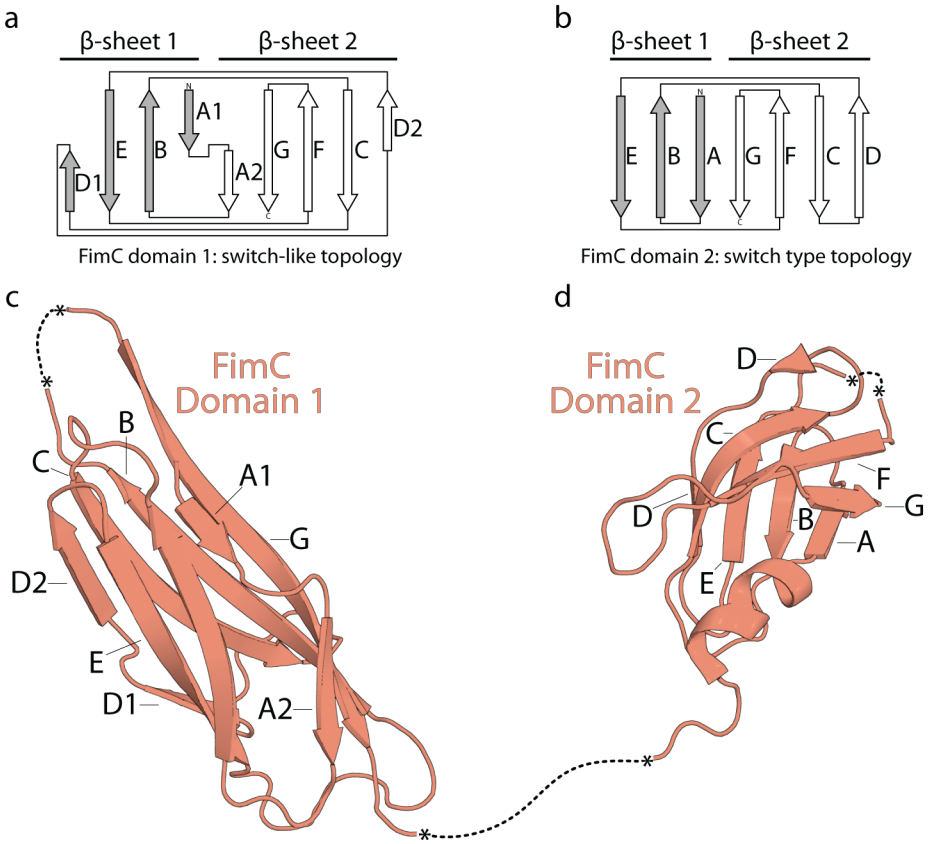


Figure 4.1.4 FimC contains two switch type immunoglobulin domains.

Topology maps (above) and ribbon diagrams (below) of FimC (PDB: 4DWH). Arrows represent β-strands, where the first and second β-sheets are comprised of grey and white arrows, respectively. Asterisks with dashed lines are shown to indicate intentional breaks (between residues 119 and 120) or unstructured regions (residues 94-100 and 179-182) within the depicted structure. **a,** The first FimC domain, which adopts a novel, as yet unreported immunoglobulin-like fold derived from the switch type topology. **b,** The second FimC domain, which adopts a canonical switch type immunoglobulin fold. **c,** Residues 1-119 of mature FimC and **d,** Residues 120-204 of mature FimC.

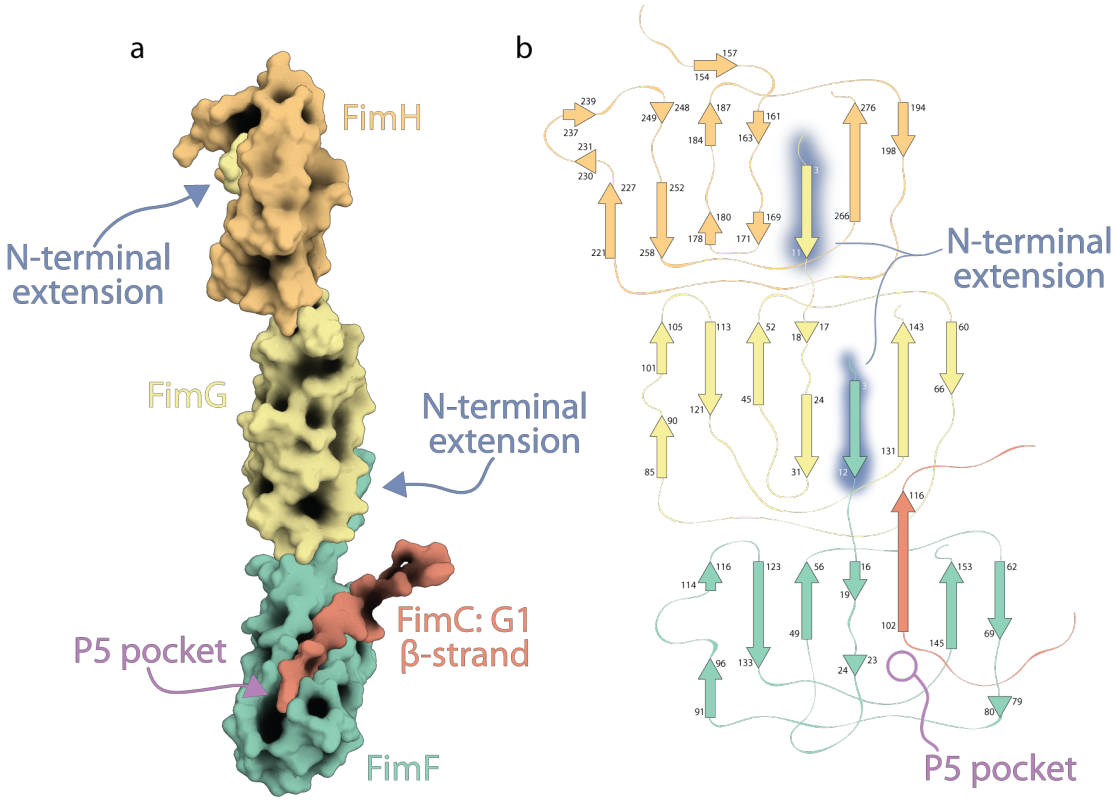


Figure 4.1.5 Donor strand exchange.

Donor strand exchange occurs when an incoming N-terminal extension (blue) interacts with the P5 pocket (purple) of an acceptor subunit, followed by the subsequent displacement of the chaperone donor strand (G1) by that N-terminal extension. **a,** Surface-structure representation of the fimbrial tip during the early stages of pilus biogenesis (PDB: 4J3O), comprising: full-length FimFG, strand G1 of FimC (residues 101-117), and the immunoglobulin-like domain of FimH (residues 153-279). **b,** Derived topology model based on "**a**", where each arrow represents a β-strand with the first and last residues indicated.

domains (CTDs: CTD1 and CTD2) (Figure 4.1.6). The NTD of FimD contains a β-sandwich domain, but rather than the aligned β-sandwich found in an immunoglobulin-like fold, it adopts an orthogonal β-sandwich topology (Figure 4.1.7a). Conversely, the remaining periplasmic domains of FimD - like the fimbrial subunits and chaperone component - contain an immunoglobulin-like topology (Figure 4.1.7b-d) (Bodelón *et al.*, 2013).

FimD exists in: (i) an inactive form, where the plug domain resides within and occludes the lumen of the barrel, and; (ii) an active form - where the plug domain is located in the periplasm where it stabilises the NTD (Phan *et al.*, 2011) (Figure 4.1.8). Interestingly, the kidney-shaped β-barrel becomes rounder on FimD activation, presumably to accommodate fimbrial subunits threading through the barrel lumen (Figure 4.1.8e) (Phan *et al.*, 2011). While the exact activation mechanism remains to be elucidated, it is known that binding of the c-type lectin domain of FimH stimulates this conversion from inactive FimD to active FimD (Munera *et al.*, 2007; Munera *et al.*, 2008).

Both the NTD and CTDs may bind a chaperone-subunit complex, but although the CTDs have a higher binding affinity, the NTD acts as the initial recognition site (Munera *et al.*, 2007; Werneburg *et al.*, 2015). A recent study by Werneburg *et al.* (2015) revealed that, in its inactive form, FimD's plug domain masks the CTDs' binding ability, leaving only the NTD to bind a chaperone-subunit complex. Following activation, the CTDs are "unmasked" so a chaperone-subunit complex will transfer from the NTD to the higher affinity CTDs, thereby once again "masking" the CTDs and freeing the NTD (Werneburg *et al.*, 2015). Using the NTD

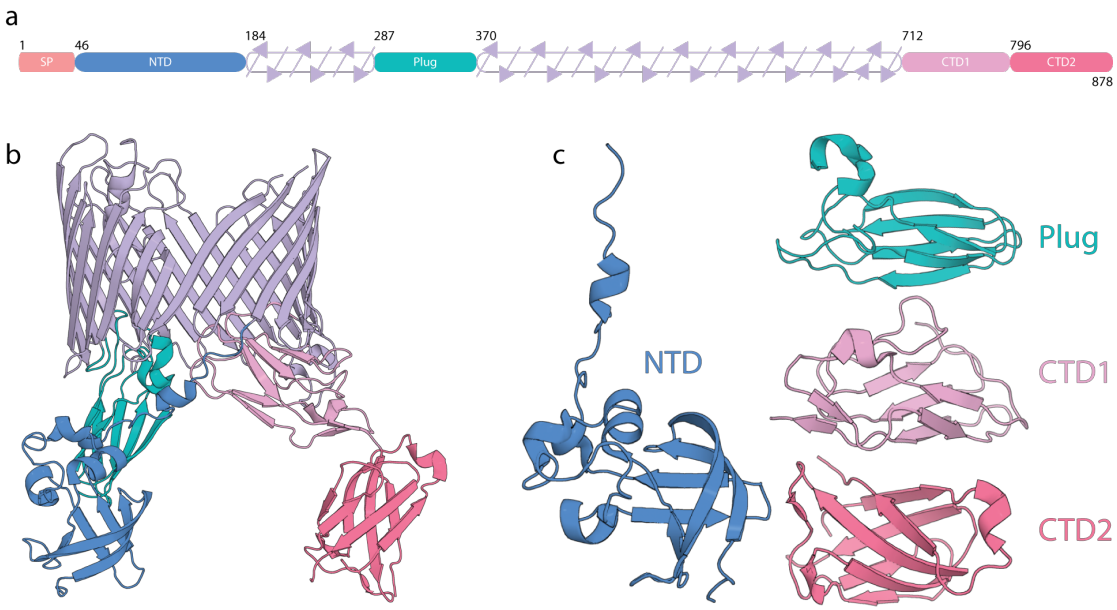


Figure 4.1.6 FimD structure.

Shown here is Figure 3.8.1. **a,** Structural map of FimD, based on its crystal structure (PDB: 3RFZ). **b,** Ribbon diagram of FimD. **c,** Ribbon diagrams of the four periplasmic domains of FimD in isolation, comprising: the N-terminal domain (NTD), plug domain and two C-terminal domains (CTD1 and CTD2).

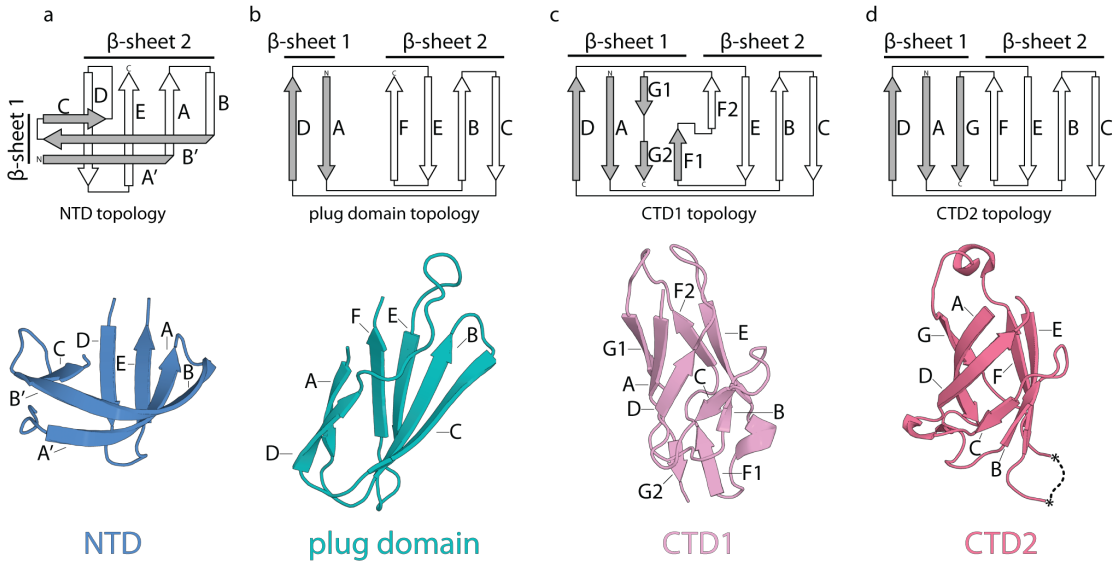


Figure 4.1.7 β-Sandwich domains within FimD.

Topology maps (above) and ribbon diagrams (below) of the orthogonal β-sandwich domain (**a**) and immunoglobulin-like folds (**b**-**d**) found within FimD (PDB: 3RFZ). Arrows represent β-strands, where the first and second β-sheets are comprised of grey and white arrows, respectively. Asterisks with dashed lines are shown to indicate intentional breaks (between residues 64 and 99 of NTD) or unstructured regions (between residues 804 and 808 of CTD2) within the depicted structures. **a,** NTD, comprising residues 26-115 of mature FimD. **b,** The plug domain, comprising residues 255-316 of mature FimD. **c,** CTD1, comprising residues 672-750 of mature FimD, where the secondary structure was altered as per Phan *et al.* (2011, Supplementary Figure S3 therein). **d,** CTD2, comprising residues 751-834 of mature FimD.

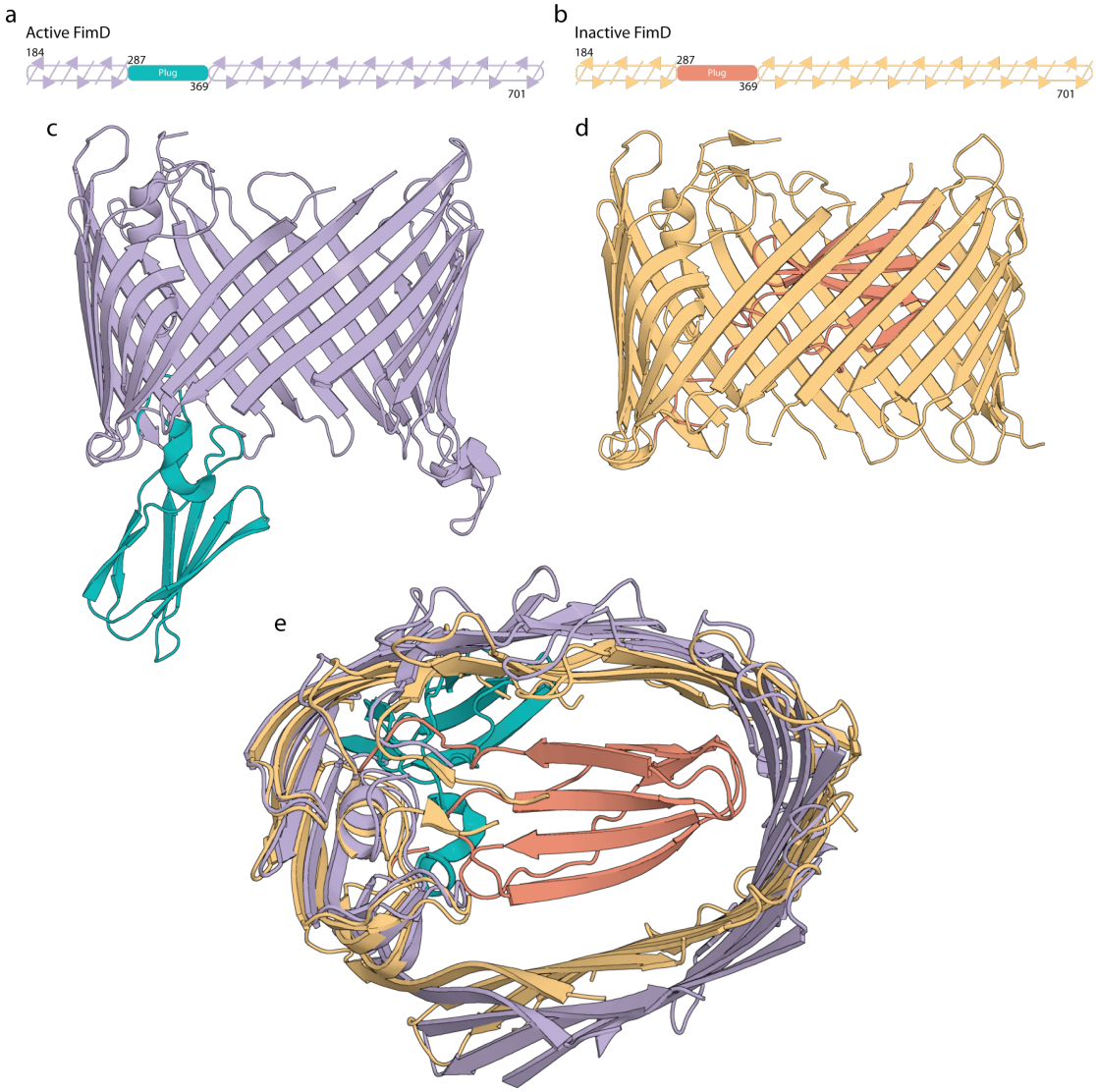


Figure 4.1.8 Active and inactive conformations of FimD.

Structures of the plug and transmembrane (excluding β-strand 23') domains of FimD are shown. **a,** Structural map of active FimD, based on its crystal structure (PDB: 3RFZ). **b,** Structural map of inactive FimD, based on its crystal structure (PDB: 3OHN). **c,** Ribbon diagram of active FimD. **d,** Ribbon diagram of inactive FimD. **e,** Ribbon diagrams of the top views of active and inactive FimD, superimposed using the "align" function of Pymol (with a 2.0 Å cutoff, following 9 cycles).

and CTDs as a platform, donor strand exchange will occur between two chaperone-subunit complexes, where the NTD has bound the "incoming" subunit and the CTDs have bound the "acceptor" subunit (Figure 4.1.9). Once donor strand exchange occurs: (i) FimC dissociates from the CTDs, (ii) the acceptor subunit is transferred from the CTDs to the FimD barrel lumen, (iii) the chaperone-subunit complex is transferred from the NTD to the CTDs, (iv) donor strand exchange repeats until the termination subunit is added (Busch *et al.*, 2015).

One of the more intriguing aspects of pilus biogenesis is that, despite a mixed population of chaperone-subunit complexes, fimbrial subunits are inserted in the same order: FimH, then FimG, then FimF, then multiple copies of FimA, then finally FimI, the termination subunit. This "ordered" assembly occurs due to the differential binding affinities between each FimD complex and the incoming chaperone-subunit complex, where as the each subunit is added, conformational changes occur in FimD that enables it to selectively recruit the next subunit (Busch *et al.*, 2015).

An essential precursor to the assembly of chaperone-usher fimbriae is the biogenesis of the usher component, FimD. Following translocation via the Sec pathway, the nascent FimD polypeptide is transferred to the OM by SurA and is thought to be assembled by the BAM complex (Palomino *et al.*, 2011). Although the steady-state level of FimD is (i) reduced in the absence of BamB or BamE, and; (ii) proportional to the levels of SurA, BamB and BamA (Palomino *et al.*, 2011), the recent finding that TamA is assembled in a BamA-dependent manner now adds a level of uncertainty to these past conclusions (Dunstan *et al.*, 2015).

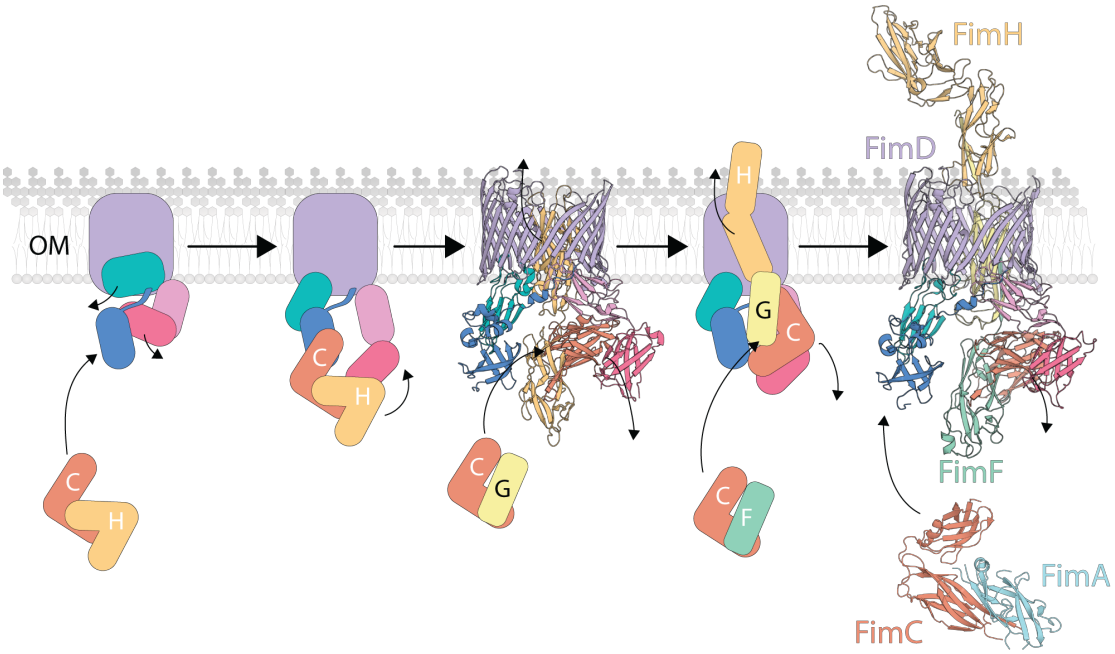


Figure 4.1.9 Biogenesis of type 1 fimbriae.

Schematic of the initial stages of type 1 fimbrial biogenesis, from illustrations of FimD activation by the FimC:FimH complex to the completed pilus tip. Ribbon diagrams are shown of the FimCDH complex (centre, PDB: 3RFZ) and the FimCDFGH complex (right, PDB: 4J3O). Curved arrows indicate subunit movements during donor strand exchange resulting in the subsequent FimD complex immediately to the right. A ribbon diagram of the FimCA complex (PDB: 4DWH) is used to indicate that the next series of steps involve the incorporation of multiple FimA subunits. FimD domains are coloured as per Figure 4.1.6: transmembrane domain (purple), NTD (blue), plug (green), CTD1 (light pink), and CTD2 (dark pink). The remaining fimbrial subunits are coloured as per Figure 4.1.2: FimA (cyan), FimC (red), FimF (light green), FimG (pale yellow) and FimH (orange).

Considering that the TAM is also capable of catalysing OMP assembly (Selkrig *et al.*, 2012), it was important to dissect the relative contributions of the BAM complex and the TAM in catalysing the biogenesis of the fimbrial usher component.

4.2 Assembly of FimD in wildtype cells

In section 3.8, protease shaving of wildtype *E. coli* revealed that surface-exposed FimD was cleaved into two major fragments (A and C). To better monitor the formation of fragments A and C, the pulse chase analysis was repeated over a range of chase temperatures (Figure 4.2.1). A chase temperature range between 20 °C and 30 °C was found to be ideal; whereas FimD assembly was too slow at 15 °C and too rapid at 37 °C. Importantly, protease shaving did not result in the simultaneous appearance of FimD fragments A and C; instead, fragment C appeared before fragment A.

To confirm this observation, the early minutes of FimD assembly were monitored using a 30 °C chase temperature (Figure 4.2.2a). By measuring the density of the fragments from four biological replicates, it can be seen that fragment C appears at the first chase time of 10 seconds, whereas the appearance of fragment A is delayed by up to 10-30 seconds (Figure 4.2.2b). The absence of fragment A at the initial time points indicates that it was completely degraded by proteinase K, suggesting that FimD insertion is initiated from a site within fragment C. The protease-sensitivity of fragment A would then be protected from proteolysis in the latter time points as the region corresponding to fragment A is assembled to generate a mature FimD, presumably by membrane insertion.

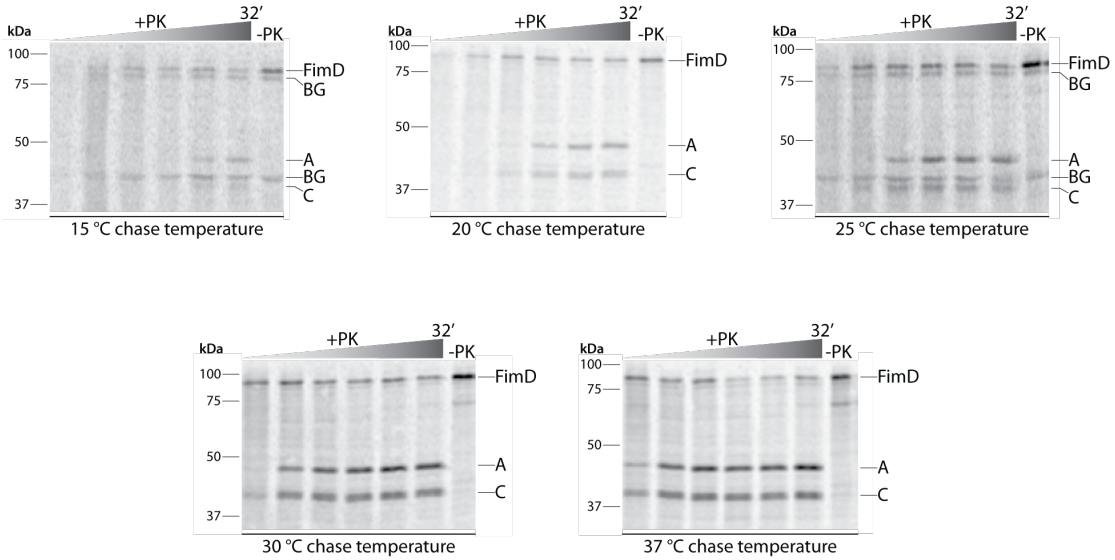


Figure 4.2.1 Chase temperature optimisation for FimD biogenesis.

FimD biogenesis in *E. coli* BL21 Star™ (DE3) wildtype cells harbouring pKS02 (pET-15b containing *fimD*) was assessed by pulse chase analysis using the indicated chase temperatures. Aliquots were taken at 10 seconds, 2, 4, 8, 16 and 32 minutes, and treated with (+PK) or without (-PK, last timepoint only) 50 µg mL^{-1} proteinase K (static, 10 min, on ice). Proteins were then TCA-precipitated, washed with acetone and boiled for 3 minutes in SDS sample buffer. Samples were analysed by SDS-PAGE and storage phosphor-imaging. Sizes in kDa are indicated on the left. Full-length FimD, its fragments A and C, and prominent background (BG) bands are indicated on the right. The time increment is indicated above the autoradiogram as a graded triangle and the chase temperature used is indicated below the autoradiogram. The "30 °C" panel is identical to the "nitrocellulose" panel from Figure 3.8.3c. Analysis at 20 °C, 30 °C and 37 °C was performed simultaneously (i.e. same *E. coli* batch); analysis at 15 °C and 25 °C was also performed simultaneously (i.e. same *E. coli* batch).

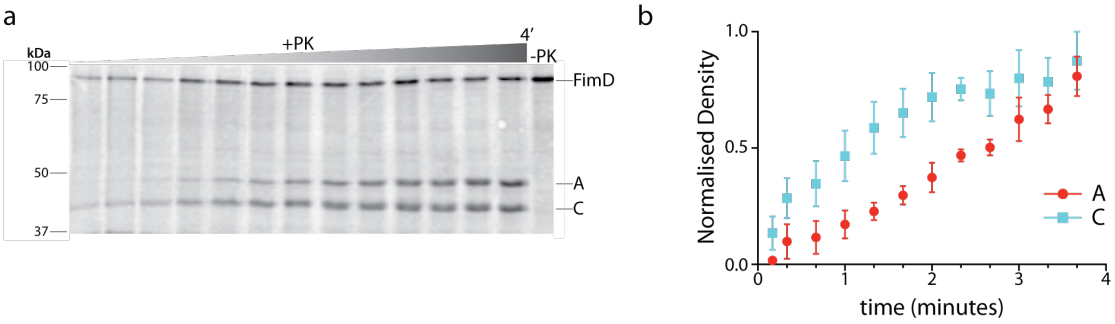


Figure 4.2.2 The early minutes of FimD biogenesis.

FimD biogenesis in *E. coli* BL21 Star™ (DE3) wildtype cells harbouring pKS02 (pET-15b containing *fimD*) was assessed by pulse chase analysis. Aliquots were taken at 10 and 20 seconds, then every 20 seconds up to 4 minutes, and treated with (+PK) or without (-PK, last timepoint only) 50 µg mL^{-1} proteinase K (static, 10 min, on ice). Proteins were then TCA-precipitated, washed with acetone and boiled for 3 minutes in SDS sample buffer. Samples were analysed by SDS-PAGE, storage phosphor-imaging and densitometry. **a,** Representative autoradiogram from four biological replicates. Sizes in kDa are indicated on the left. Full-length FimD and its fragments A and C are indicated on the right. The time increment is indicated above the autoradiogram as a graded triangle. **b,** Normalised band density of fragments A and C plotted versus time, where error bars represent standard error of the mean from four biological replicates.

4.3 The identity of fragments A and C.

The ~50 kDa and ~40 kDa fragments (A and C, respectively) appear to comprise the majority of full-length FimD (about 90 kDa), suggesting FimD may be cleaved at a site either 40 kDa or 50 kDa from the N-terminus. By inspecting the membrane topology of FimD (Figure 4.3.1), there are two extracellular loops where proteolytic cleavage would result in products of the correct sizes. Cleavage at loop 5 would result in fragment C being N-terminal and fragment A being C-terminal (Figure 4.3.1c); however, cleavage at loop 7 would mean fragments A and C respectively correspond to the N- and C-terminus of FimD instead (Figure 4.3.1d).

Loop 7 of FimD was previously shown to be protease-sensitive by Saulino *et al.* (1998). They used exogenously added trypsin to determine the order of pilus assembly through the FimD pore. When the correct fimbrial subunit FimH was co-expressed with FimD (and the FimC chaperone), a ~40 kDa trypsin-protected fragment was formed. N-terminal sequencing revealed the trypsin-protected fragment began at FTDYN, indicating proteolytic cleavage occurred between K469 and F470 of the mature protein (located within loop 7). It should be noted that the authors reported the cleavage site as being between residues 477 and 488, likely due to the fact that the FimD signal peptide was then thought to be 37 residues long, but is now known to be 45 residues long (Ng *et al.*, 2004). When FimD was cleaved with trypsin for structural determination, Phan *et al.* (2011) reported that trypsin removed two 2 kDa fragments: one at the N-terminus and another from within loop 7, presumably leaving ~40 kDa and ~50 kDa fragments that they could use to form crystals. Additionally, the crystal structures of FimD show that the remaining region comprising loop 7 is disordered (Geibel *et al.*, 2013; Phan *et*

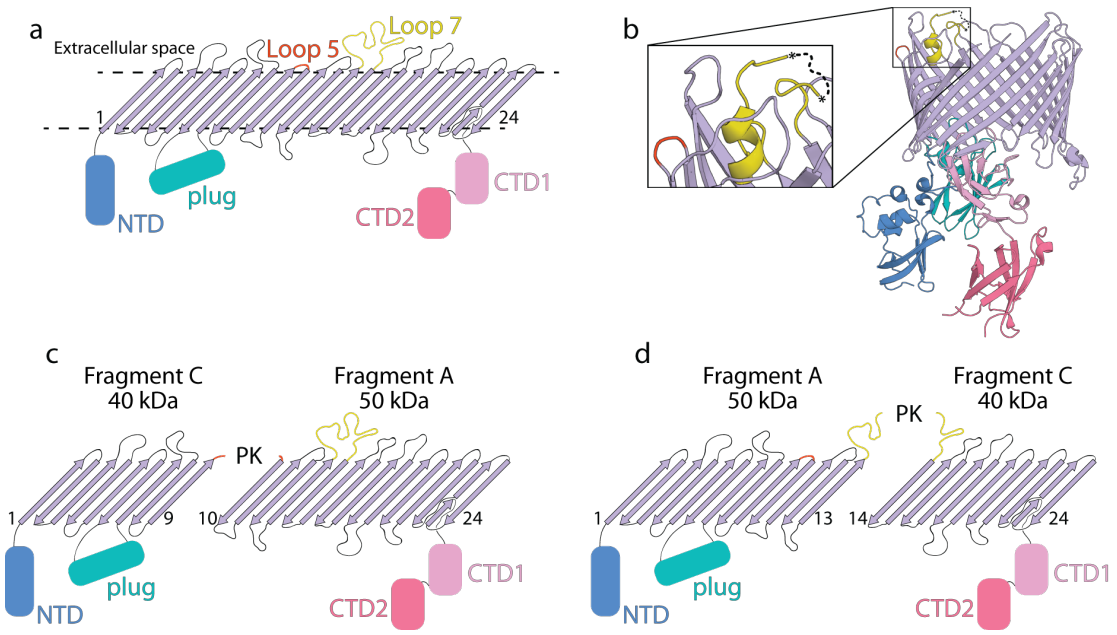


Figure 4.3.1 Predicted cleavage sites.

a, Cartoon depiction of the membrane topology of FimD. The first (1) and last (24) β-strands are indicated, as are loops 5 (red) and 7 (yellow). **b,** Ribbon diagram of active FimD (PDB: 3RFZ) coloured as in "**a**". Asterisks with dashed lines are shown to indicate the unstructured region within loop 7 (between residues 453 and 474 of mature FimD). **c**-**d,** The putative identities of fragments A and C are indicated where proteolytic cleavage occurs within loop 5 (**c**) or loop 7 (**d**).

al., 2011), and may potentially adopt a protease-sensitive structure.

From this data, it can be seen that loop 7 is the best candidate cleavage site for proteinase K. To test this experimentally, modified FimD constructs were used: N_{strep}-FimD, which contains an N-terminal *strep*-tag II (ASWSHPQFEK), or TEV-FimD-C_{strep}, which includes a tobacco etch virus (TEV) cleavage site (ENLYFQG) within loop 7 and a C-terminal *strep*-tag II (Figure 4.3.2). Protease shaving and pulse chase analysis of N_{strep}-FimD and TEV-FimD-C_{strep} revealed the modifications do not affect the biogenesis of FimD, as judged by the similar fragmentation profile to the unmodified FimD (Figure 4.3.3a).

The addition of *strep*-tag II will add 1.5 kDa to the protein (Figure 4.3.2b), which at the range of 40-50 kDa is a size that could be resolved by SDS-PAGE analysis (Figure 4.3.3b). Inspection of the relative fragment sizes showed that fragment A from N_{strep}-FimD and fragment C from TEV-FimD-C_{strep} migrated slower than the corresponding native fragments, suggesting these bands contained *strep*-tag II. The presence of *strep*-tag II on these respective fragments was then confirmed by detection with the highly specific streptavidin-HRP conjugate probe (Figure 4.3.3c). Full length and fragment A from N_{strep}-FimD, but not fragment C, were bound by streptavidin-HRP. Likewise, full length and fragment C from TEV-FimD-C_{strep}, but not fragment A, were bound by streptavidin-HRP. This indicates that fragment A corresponds to the N-terminal 50 kDa region of FimD and fragment C represents the remaining 40 kDa C-terminal portion of FimD.

FimD contains a suboptimal TEV cleavage site located near the end of β-strand 5

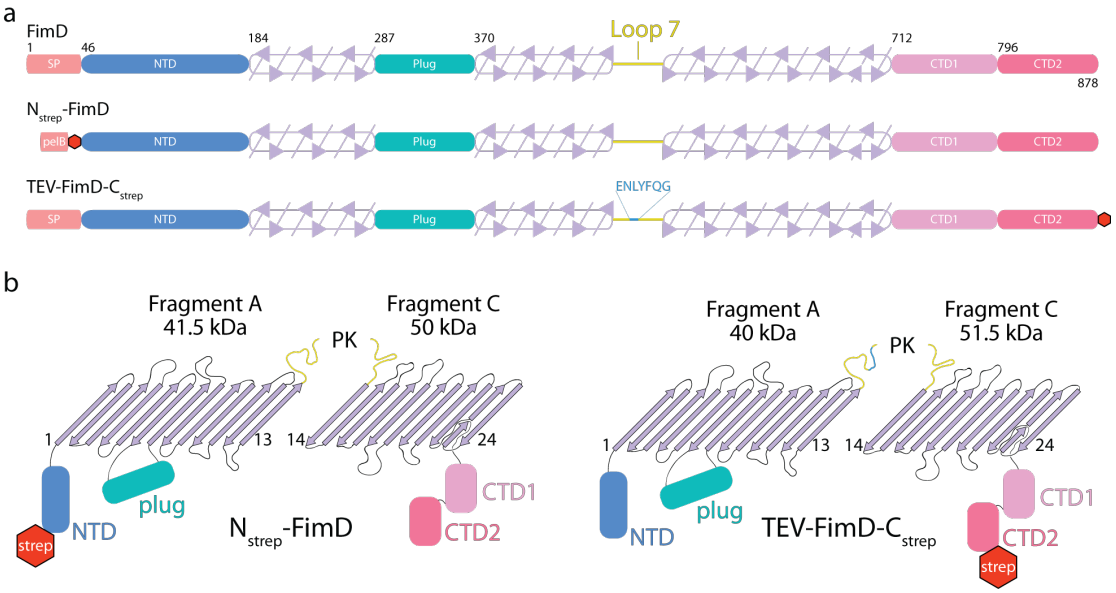


Figure 4.3.2 Structural maps and topologies of FimD modified to include a *strep*-tag II.

a, Structural map of unmodified and modified pre-FimD based on the crystal structure (PDB: 3RFZ). The signal peptides (SP) used are from FimD (native and TEV-FimD-C_{strep}) or pelB (N_{strep}-FimD). The blue region within loop 7 (ENLYFQG) indicates the TEV site and red hexagons represent *strep*-tag II. Amino acid positions are indicated at the start of each domain for the preprotein. **b,** The putative identities of fragments A and C are indicated where proteolytic cleavage occurs within loop 7.

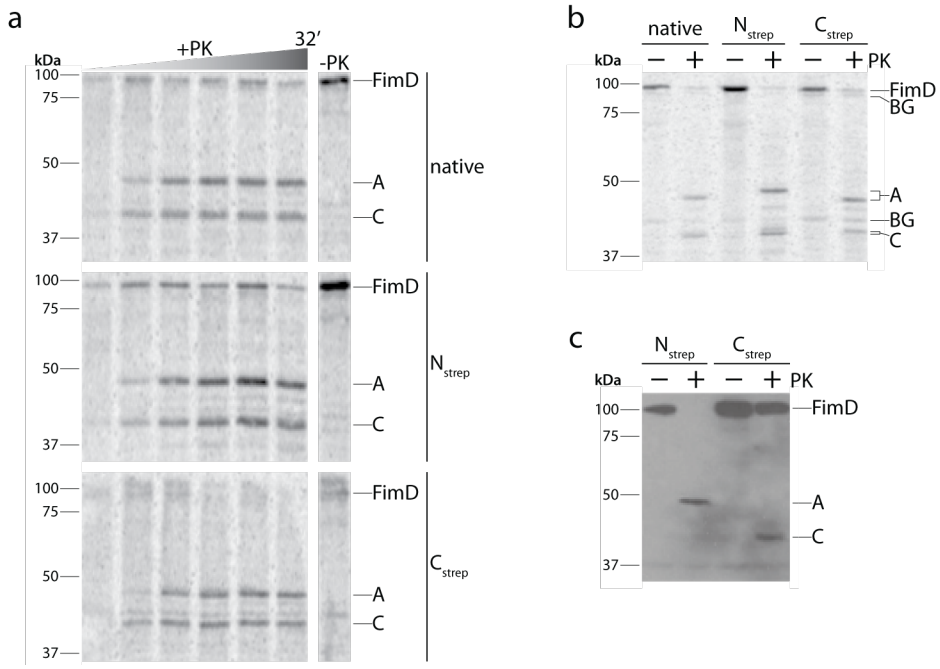


Figure 4.3.3 The identities of fragments A and C.

Shown are cropped versions of **a,** Figure 4.5.2; **b,** Figure 4.5.3a, and; **c,** Figure 4.5.3b. **a**-**c,** FimD biogenesis was assessed by pulse chase analysis in *E. coli* BL21 Star™ (DE3) wildtype cells harbouring either: (i) pKS02 (pET-15b containing *fimD*), (ii) pCJS51 (pET-22b(+) containing N_{strep}-*fimD*) or (iii) pCJS29 (pET-15b containing *TEV-fimD-*C_{strep}). Aliquots were taken at either: **a,** 10 seconds, 2, 4, 8, 16 and 32 minutes; **b,** 32 minutes only, or; **c,** 16 minutes only. Aliquots were treated with (+PK) or without (-PK, last timepoint only) 50 μg mL^{-1} proteinase K (static, 10 min, on ice). Proteins were then TCA-precipitated, washed with acetone and boiled for 3 minutes in SDS sample buffer. Samples were analysed by SDS-PAGE and either storage phosphor-imaging (**a** and **b**) or Western blotting (**c**) using Precision Protein™ StrepTactin-HRP conjugate (Bio-Rad) for detection. Sizes in kDa are indicated on the left. Full-length FimD, its fragments A and C, and prominent background (BG) bands are indicated on the right. The time increment is indicated above the autoradiogram as a graded triangle and whether native or modified FimD biogenesis was assessed is indicated above or to the right of the autoradiogram.

and within the start of loop 3: GDGYTQG (Figure 4.3.4a). Loop 3 is unusual in that it points perpendicularly toward the barrel lumen where it plays a structurally important role in stabilising the plug domain of inactive FimD (Volkan *et al.*, 2013) (Figure 4.3.4b). Interestingly, an α-helical region within loop 7 also plays a structurally significant role in stabilising the plug domain of inactive FimD (Volkan *et al.*, 2013) (Figure 4.3.4b). Protease shaving analysis of FimD using TEV protease resulted in the formation of the predicted 66.5 kDa fragment (Fragment A*, Figure 4.3.4c), indicating that TEV protease may enter the lumen of FimD to cleave within loop 3, or that loop 3 sometimes has partial occupancy at the cell surface.

TEV protease should therefore cleave TEV-FimD-C_{strep} in loops 3 and 7 to produce fragment C and two 25 kDa fragments corresponding to fragment A. Fragment C was produced as expected, but was found to migrate slower when TEV protease was used instead of proteinase K, indicating that protease-sensitive loop 7 likely contains multiple proteinase K cleavage sites (Figure 4.3.4c). Surprisingly however, protease shaving analysis of TEV-FimD-C_{strep} biogenesis using TEV protease or proteinase K generated an equal amount of fragment A, and no fragment A* (Figure 4.3.4c). Perhaps once loop 7 was cleaved, it acted to protect loop 3 from further proteolysis. Considering that loop 7 contains the consensus TEV cleavage site, whereas the site in loop 3 is suboptimal, loop 7 must always be cleaved first, thus precluding the formation of fragment A*.

To identify the specific proteinase K cleavage site in loop 7, LC-MS/MS analysis was performed. Non-radiolabelled full-length FimD, and fragments A and C were excised from polyacrylamide gels, and subjected to in-gel tryptic digestion (Phu *et*

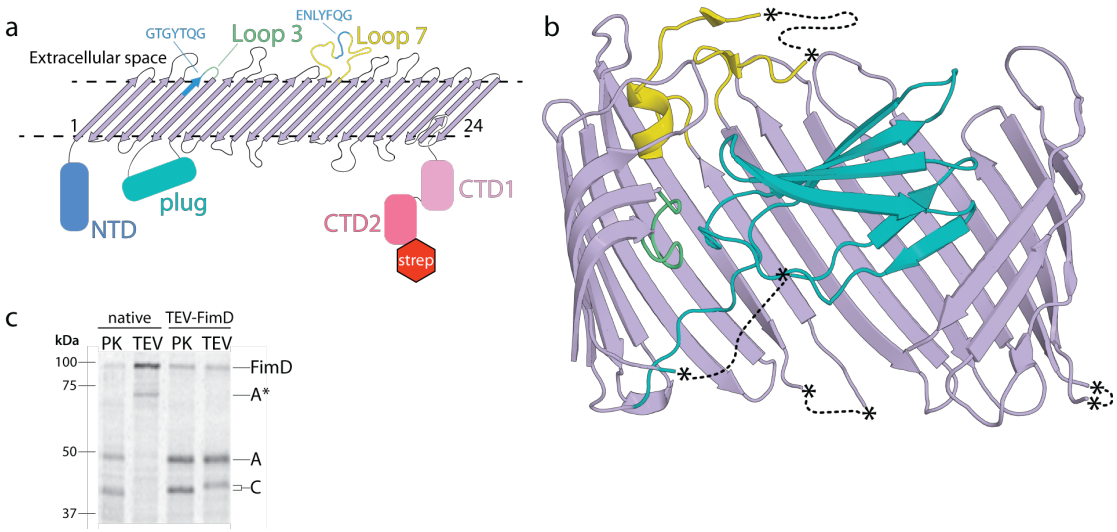


Figure 4.3.4 Protease shaving analysis of FimD biogenesis using proteinase K or TEV proteases.

a, TEV-FimD-C_{strep} topology with the sub-optimal (GTGYTQG) and optimal (ENLYFQG) TEV cleavage sites indicated. **b,** Ribbon diagram of inactive FimD (PDB: 3OHN). For clarity, residues 139-212 and 630-656 were omitted. Asterisks with dashed lines are shown to indicate several unstructured regions, for residues: 244-253 (plug domain), 423-425 (turn 6), 455-470 (loop 7) and 558-571 (turn 9). **c,** FimD biogenesis was assessed by pulse chase analysis in *E. coli* BL21 Star™ (DE3) wildtype cells harbouring either: (i) pKS02 (pET-15b containing *fimD*), (ii) pCJS29 (pET-15b containing *TEV-fimD-*C_{strep}). Aliquots were taken at 32 minutes and treated with either 50 μg mL^{-1} proteinase K (PK) (static, 10 min, on ice) or 100 μg mL^{-1} TEV protease and 1 mM DTT (static, 30 min, 30 °). Proteins were then TCA-precipitated, washed with acetone and boiled for 3 minutes in SDS sample buffer. Samples were analysed by SDS-PAGE and storage phosphor-imaging. Sizes in kDa are indicated on the left. Full-length FimD, its fragments A, A* and C, and prominent background (BG) bands are indicated on the right.

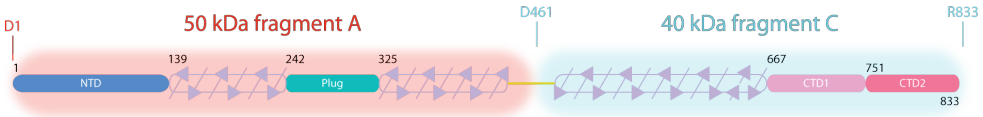


Figure 4.3.5 Boundaries of fragments A and C.

FimD structural map based on the crystal structure (PDB: 3RFZ) with the boundaries of fragments A (red) and C (blue) - determined from *strep*-tag II, TEV cleavage and LC-MS/MS analysis, indicated.

al., 2011), which shortens the proteins into analysable peptides. Trypsin cleaves the peptide bond following a lysine or arginine residue, whereas proteinase K has a much broader specificity (Keil, 1992). Proteinase K has been shown to preferentially cleave peptide bonds following amino acids with aliphatic or aromatic side chains (Keil, 1992), but a recent study by Qasim (2014) revealed that proteinase K also has a strong tendency to cleave the peptide bond following threonine and glutamine, which have polar side chains.

From the LC-MS/MS data, the most prominent non-tryptic peptide that was unique to fragment C (i.e. not observed from tryptic digestion of full-length FimD or fragment A) was DQVIQVKPK, which corresponds to a cleavage site between Q460 and D461 of the mature protein. Although there may be additional cleavage sites within loop 7, fragment A roughly corresponds to a 50.7 kDa band comprised of the NTD, the plug domain and the first 13 transmembrane β-strands (typically D1-Q460), whereas fragment C is 40.7 kDa and represents the remaining 11 transmembrane β-strands, CTD1 and CTD2 (typically D461-R833) (Figure 4.3.5).

4.4 FimD requires the TAM for efficient assembly

When FimD assembly was monitored by pulse chase analysis, exogenously added proteinase K cleaved FimD into 50 kDa N-terminal and 40 kDa C-terminal fragments. In the absence of the TAM however, a 45 kDa fragment B was generated as the major FimD proteolytic cleavage product (Figure 4.4.1). When Δ*tamA* or Δ*tamB* strains were then complemented, FimD biogenesis levels were restored to wild-type levels, as judged by the significant (i) increase in fragments A and C, and (ii) concomitant decrease in fragment B (Figure 4.4.2).

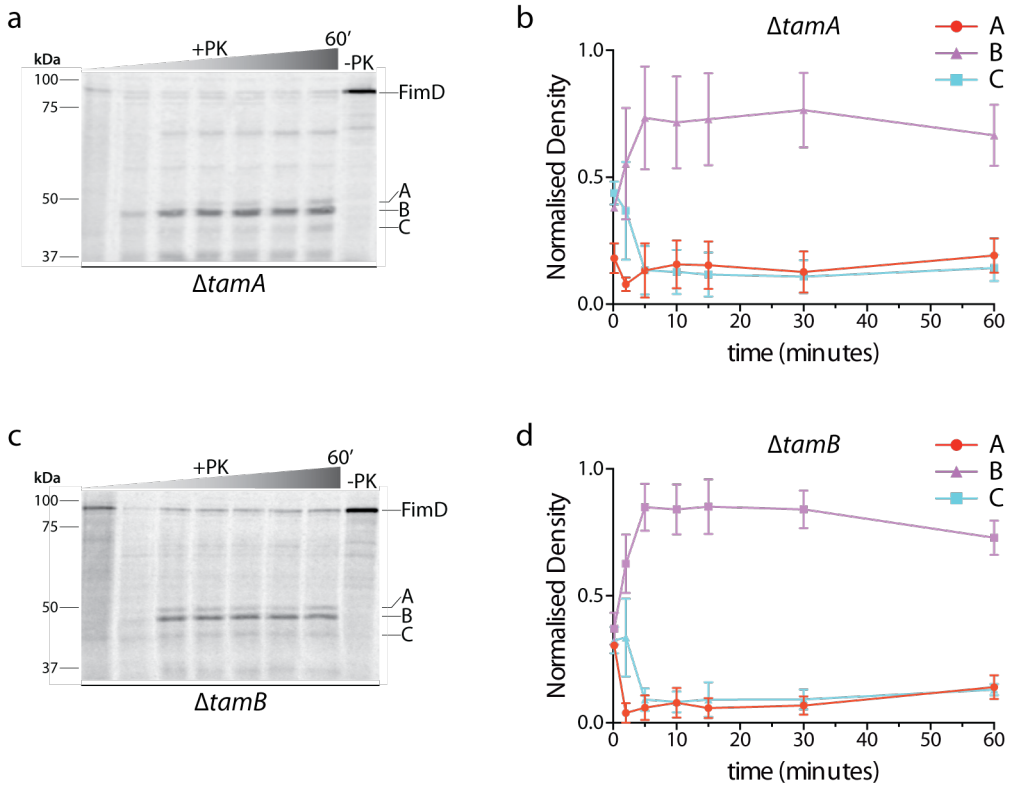


Figure 4.4.1 FimD requires the TAM for assembly.

FimD biogenesis was assessed by pulse chase analysis in the indicated strains of *E. coli* BL21 Star™ (DE3) harbouring pKS02 (pET-15b containing *fimD*). Aliquots were taken at 10 seconds, 2, 5, 10, 15, 30 and 60 minutes, and treated with (+PK) or without (-PK, last timepoint only) 50 µg mL^{-1} proteinase K (static, 10 min, on ice). Proteins were then TCA-precipitated, washed with acetone and boiled for 3 minutes in SDS sample buffer. Samples were analysed by SDS-PAGE, storage phosphor-imaging and densitometry. **a** and **c,** Representative autoradiogram from four biological replicates. Sizes in kDa are indicated on the left. Full-length FimD and its fragments A and C are indicated on the right. The time increment is indicated above the autoradiogram as a graded triangle and the strain is indicated below. **b** and **d,** Normalised band density of fragments A, B and C plotted versus time, where error bars represent standard error of the mean from four biological replicates and the strain is indicated above the graph.

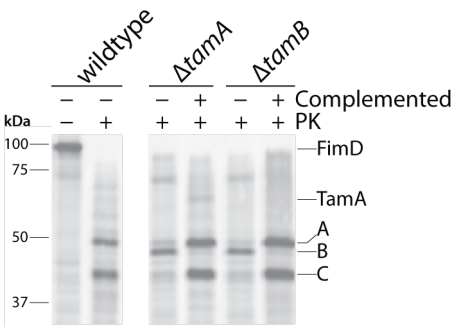


Figure 4.4.2 Complementation restores "wildtype" FimD biogenesis.

Shown is the "cropped" version of Figure 4.4.8. FimD biogenesis was assessed by pulse chase analysis in the indicated strains of *E. coli* BL21 Star™ (DE3) harbouring pKS02 (pET-15b *fimD*) in conjunction with either the pACYCDuet-1 base vector or a complementation plasmid (indicated by "-" or "+" complemented, respectively): pCJS69 (pACYCDuet-1 containing *tamA*) or pCJS72 (pACYCDuet-1 containing *tamB*). Aliquots were taken at 16 minutes only and treated with (+PK) or without (-PK, "wildtype" only) 50 µg mL^{-1} proteinase K (static, 10 min, on ice). Proteins were then TCA-precipitated, washed with acetone and boiled for 3 minutes in SDS sample buffer. Samples were analysed by SDS-PAGE and storage phosphor-imaging. Sizes in kDa are indicated on the left. TamA, full-length FimD and its fragments A, B and C are indicated on the right. The strain identities and whether they harbour a complementation plasmid are indicated above the autoradiogram.

In addition to fragment B, protease shaving analysis revealed that a small proportion of fragments A and C were generated during FimD biogenesis in the absence of the TAM (Figure 4.4.1), indicating that FimD can still assemble correctly, albeit less efficiently. To extend this observation and determine whether it reflects a slower assembly rate, FimD biogenesis was allowed to proceed for up 256 minutes in Δ*tamA* or Δ*tamB* mutants (Figure 4.4.3). Pulse chase analysis revealed that at 2 minutes, fragment B is generated and from 32 minutes: fragments A and C are also generated, with approximately equivalent amounts of fragments A, B and C at 256 minutes. Although the function of FimD assembled in the absence of the TAM cannot be determined during these experiments, fimbriae assembled in the absence of the TAM were subsequently shown to be functional (but significantly depleted) in a parallel study by I. D. Hay (Stubenrauch *et al.*, 2016) (discussed in Section 4.8).

The BAM complex likely fills the role of catalysing FimD biogenesis in the absence of the TAM, hinting that even in the presence of the TAM, FimD biogenesis could theoretically also proceed via the slower TAM-independent pathway. To confirm this hypothesis, FimD assembly in wildtype cells was allowed to proceed for up to 60 minutes (Figure 4.4.4). Subsequent protease shaving and densitometrical analysis revealed that, in addition to the rapid generation of fragments A and C, a small proportion of fragment B was also produced, thereby confirming that FimD may assemble independently of the TAM (albeit less efficiently), despite the presence of the TAM.

Dunstan *et al.* (2015) previously showed that depletion of BamA results in a

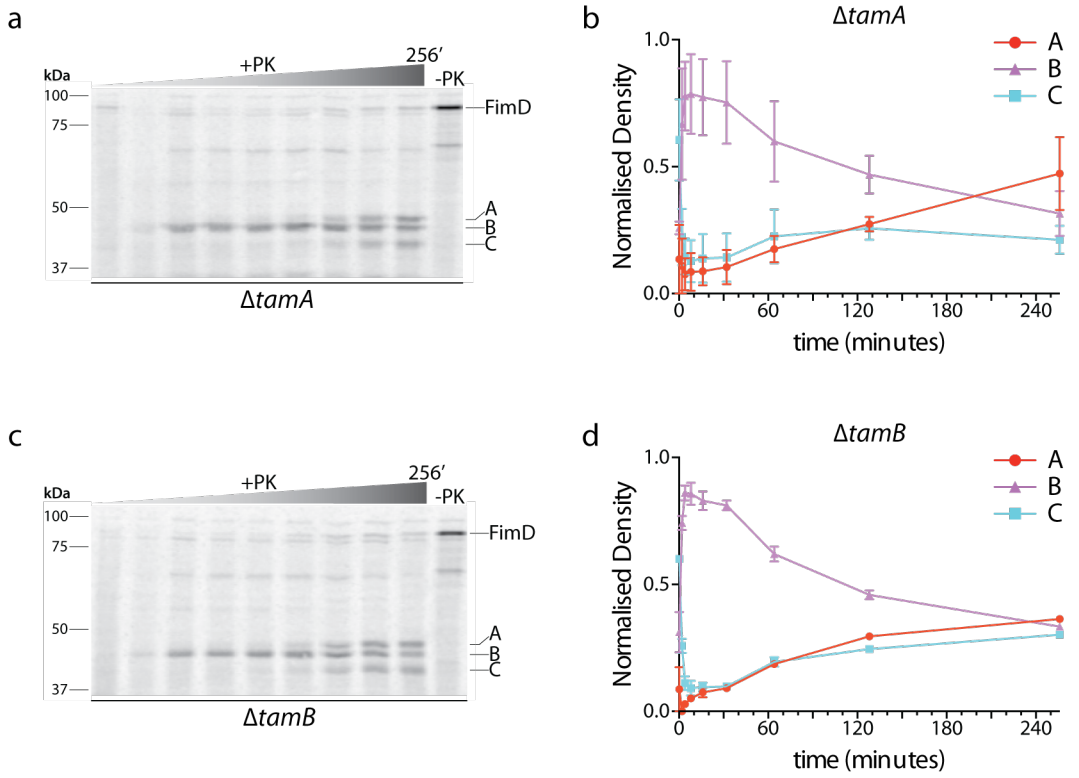


Figure 4.4.3 Correct FimD biogenesis is delayed in the absence of the TAM.

FimD biogenesis in the indicated strains of *E. coli* BL21 Star™ (DE3) harbouring pKS02 (pET-15b containing *fimD*) was assessed by pulse chase analysis. Aliquots were taken at 10 seconds, 2, 4, 8, 16, 32, 64, 128 and 256 minutes, and treated with (+PK) or without (-PK, last timepoint only) 50 μg mL^{-1} proteinase K (static, 10 min, on ice). Proteins were then TCA-precipitated, washed with acetone and boiled for 3 minutes in SDS sample buffer. Samples were analysed by SDS-PAGE, storage phosphor-imaging and densitometry. **a** and **c,** Representative autoradiogram from four biological replicates. Sizes in kDa are indicated on the left. Full-length FimD and its fragments A and C are indicated on the right. The time increment is indicated above the autoradiogram as a graded triangle and the strain is indicated below. **b** and **d,** Normalised band density of fragments A, B and C plotted versus time, where error bars represent standard error of the mean from four biological replicates and the strain is indicated above the graph.

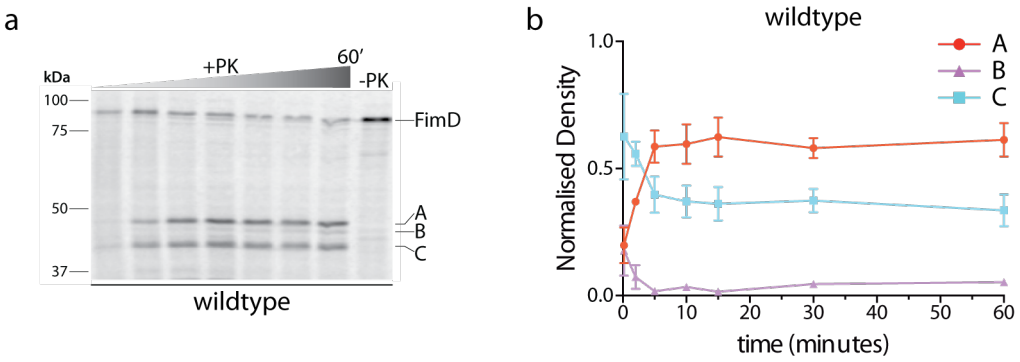


Figure 4.4.4 FimD biogenesis occurs predominantly via the TAM-dependent pathway.

FimD biogenesis was assessed by pulse chase analysis in *E. coli* BL21 Star™ (DE3) wildtype strains harbouring pKS02 (pET-15b containing *fimD*). Aliquots were taken at 10 seconds, 2, 5, 10, 15, 30 and 60 minutes, and treated with (+PK) or without (-PK, last timepoint only) 50 µg mL^{-1} proteinase K (static, 10 min, on ice). Proteins were then TCA-precipitated, washed with acetone and boiled for 3 minutes in SDS sample buffer. Samples were analysed by SDS-PAGE, storage phosphor-imaging and densitometry. **a,** Representative autoradiogram from four biological replicates. Sizes in kDa are indicated on the left. Full-length FimD and its fragments A and C are indicated on the right. The time increment is indicated above the autoradiogram as a graded triangle and the strain is indicated below. **b,** Normalised band density of fragments A, B and C plotted versus time, where error bars represent standard error of the mean from four biological replicates and the strain is indicated above the graph.

concomitant decrease in the levels of TamA, indicating that the TAM itself needs the BAM complex for TamA biogenesis. Considering the BAM lipoproteins facilitate substrate assembly, it was hypothesised that deletion of the non-essential BAM lipoproteins may reduce the levels of TamA, and therefore the TAM. Subsequent BN-PAGE and SDS-PAGE analysis of *E. coli* membranes confirmed that, while TamA levels are not affected by deletion of *bamC*, the amount of TamA was somewhat reduced in either Δ*bamB* or Δ*bamE* mutants (Figure 4.4.5). Assuming that the BAM lipoproteins themselves are dispensable for FimD biogenesis, the proportion of FimD assembled independently of the TAM should: (i) remain the same in the absence of BamC, and; (ii) increase in the absence of either BamB or BamE.

Analysis of FimD biogenesis in Δ*bamC* mutants indicated that FimD displayed a wildtype preference for the TAM-independent assembly pathway (Figure 4.4.6), as judged by the rapid generation of fragments A and C and the slow formation of low levels of fragment B. In contrast, analysis of FimD biogenesis in Δ*bamB* (Figure 4.4.7a-b) or Δ*bamE* (Figure 4.4.7c-d) mutants revealed that in addition to the rapid formation of fragments A and C, from about 4 minutes: a large proportion of FimD was also proteolytically degraded into fragment B. Complementation of either Δ*bamB* or Δ*bamE* mutants restored wildtype levels of FimD TAM-dependent biogenesis (Figure 4.4.8), thereby confirming that by reducing TAM levels, FimD biogenesis is shifted toward the less efficient TAM-independent assembly pathway.

Because TamA levels (but not TamB levels) are decreased in Δ*bamB* mutants

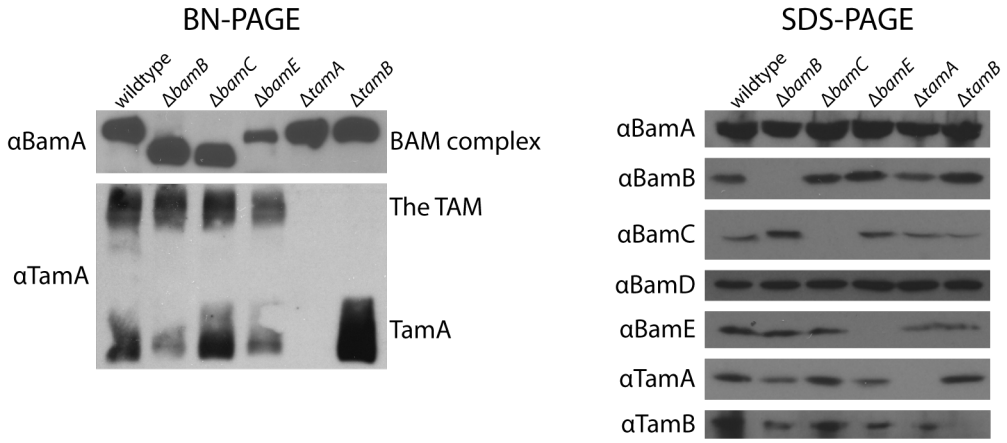


Figure 4.4.5 TamA levels are lower in Δ*bamB* or Δ*bamE* mutants.

E. coli BL21 Star™ (DE3) membranes were isolated from the indicated strains of *E. coli* and analysed by Western blotting - following BN-PAGE or SDS-PAGE analysis - using mouse antibodies specific for BAM complex subunits or rabbit antibodies specific for the TAM subunits. The PAGE method and strains are indicated above the immunoblots, antibodies are indicated to the left, and the native complex detected is indicated to the right of the BN-PAGE immunoblots.

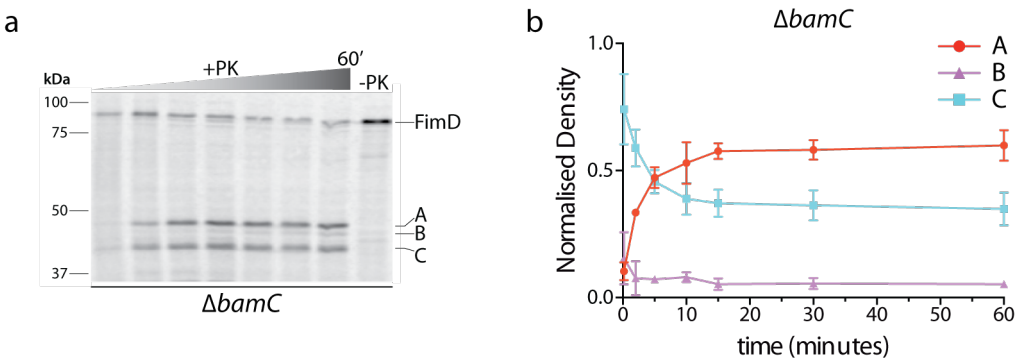


Figure 4.4.6 FimD biogenesis is not affected by BamC.

FimD biogenesis was assessed by pulse chase analysis in *E. coli* BL21 Star™ (DE3) Δ*bamC* strains harbouring pKS02 (pET-15b containing *fimD*). Aliquots were taken at 10 seconds, 2, 5, 10, 15, 30 and 60 minutes, and treated with (+PK) or without (-PK, last timepoint only) 50 μg mL^{-1} proteinase K (static, 10 min, on ice). Proteins were then TCA-precipitated, washed with acetone and boiled for 3 minutes in SDS sample buffer. Samples were analysed by SDS-PAGE, storage phosphor-imaging and densitometry. **a,** Representative autoradiogram from four biological replicates. Sizes in kDa are indicated on the left. Full-length FimD and its fragments A and C are indicated on the right. The time increment is indicated above the autoradiogram as a graded triangle and the strain is indicated below. **b,** Normalised band density of fragments A, B and C plotted versus time, where error bars represent standard error of the mean from four biological replicates and the strain is indicated above the graph.

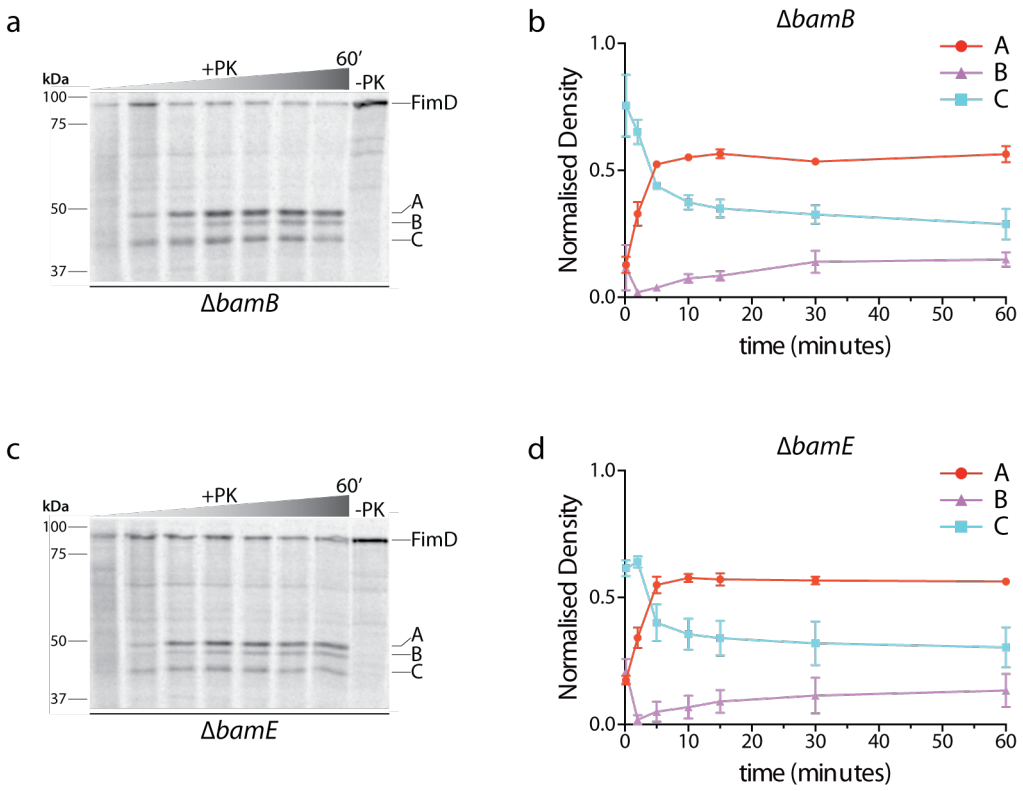


Figure 4.4.7 FimD biogenesis is affected by BamB and BamE levels.

FimD biogenesis was assessed by pulse chase analysis in the indicated strains of *E. coli* BL21 Star™ (DE3) harbouring pKS02 (pET-15b containing *fimD*). Aliquots were taken at 10 seconds, 2, 5, 10, 15, 30 and 60 minutes, and treated with (+PK) or without (-PK, last timepoint only) 50 μg mL^{-1} proteinase K (static, 10 min, on ice). Proteins were then TCA-precipitated, washed with acetone and boiled for 3 minutes in SDS sample buffer. Samples were analysed by SDS-PAGE, storage phosphor-imaging and densitometry. **a** and **c,** Representative autoradiogram from four biological replicates. Sizes in kDa are indicated on the left. Full-length FimD and its fragments A and C are indicated on the right. The time increment is indicated above the autoradiogram as a graded triangle and the strain is indicated below. **b** and **d,** Normalised band density of fragments A, B and C plotted versus time, where error bars represent standard error of the mean from four biological replicates and the strain is indicated above the graph.

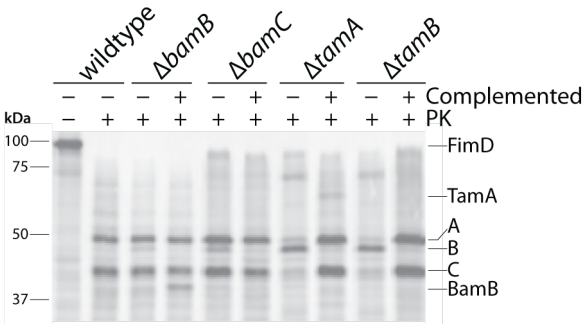


Figure 4.4.8 Complementation analysis of FimD biogenesis in TAM or non-essential BAM lipoprotein null-mutants.

Shown is the "uncropped" version of Figure 4.4.2. FimD biogenesis was assessed by pulse chase analysis in the indicated strains of *E. coli* BL21 Star™ (DE3) harbouring pKS02 (pET-15b containing *fimD*) in conjunction with either the pACYCDuet-1 base vector or a complementation plasmid (indicated by "-" or "+" complemented, respectively): pCJS69 (pACYCDuet-1 containing *tamA*), pCJS72 (pACYCDuet-1 containing *tamB*), pCJS73 (pACYCDuet-1 containing *bamB*), or pCJS74 (pACYCDuet-1 containing *bamE*). Aliquots were taken at 16 minutes and treated with (+PK) or without (-PK) 50 μg mL^{-1} proteinase K (static, 10 min, on ice). Proteins were then TCA-precipitated, washed with acetone and boiled for 3 minutes in SDS sample buffer. Samples were analysed by SDS-PAGE and storage phosphor-imaging. Sizes in kDa are indicated on the left. TamA, BamB, full-length FimD, and its fragments A, B and C are indicated on the right. The strain identities and whether they harbour a complementation plasmid are indicated above the autoradiogram.

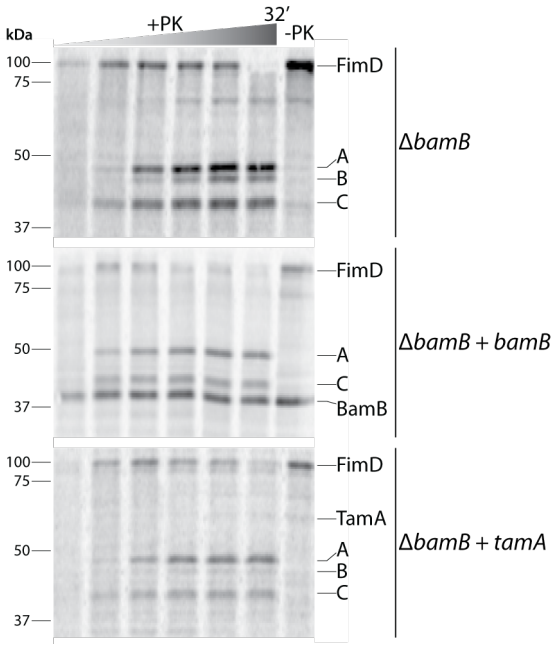


Figure 4.4.9 Correct FimD biogenesis is restored on restoration of TamA in Δ*bamB* mutants.

FimD biogenesis was assessed by pulse chase analysis in BL21 Star™ (DE3) Δ*bamB* cells harbouring either pCJS52 (pETDuet-1 containing *fimD* in MCS1) (indicated by "Δ*bamB*", top panel); pCJS60 (pCJS52 containing *bamB* in MCS2) (indicated by " Δ*bamB* + *bamB*", middle panel), or; pCJS58 (pCJS52 containing *tamA* in MCS2) (indicated by " Δ*bamB* + *tamA*", bottom panel). Aliquots were taken at 10 seconds, 2, 4, 8, 16 and 32 minutes, and treated with (+PK) or without (-PK, last timepoint only) 50 μg mL^{-1} proteinase K (static, 10 min, on ice). Proteins were then TCA-precipitated, washed with acetone and boiled for 3 minutes in SDS sample buffer. Samples were analysed by SDS-PAGE and storage phosphor-imaging. Sizes in kDa are indicated on the left. TamA, BamB, full-length FimD, its fragments A, B and C, the strain identities and whether they harbour a complementation plasmid are indicated to the right of the autoradiogram. The time increment is indicated above the autoradiogram as a graded triangle.

(Figure 4.4.5), it was hypothesised that overexpressing *tamA* in a Δ*bamB* background might similarly lead to a restoration of TAM-dependent assembly, similar to complementation of Δ*bamB* with *bamB*. As expected, FimD biogenesis in a Δ*bamB* mutant complemented with either *bamB* or *tamA* lead to a significant decrease in the generation fragment B (Figure 4.4.9). Taken together, by artificially increasing or reducing TAM levels, there is a concomitant increase or decrease in TAM-catalysed FimD assembly, respectively.

4.5 The identity of fragment B

When FimD is assembled in the absence of the TAM, extracellular shaving with proteinase K resulted in the presence of a ~45 kDa fragment B. It was tempting to speculate that this corresponds to FimD (~90 kDa) following proteolytic cleavage into two equal-sized fragments. By inspecting the membrane topology of FimD (Figure 4.5.1), cleavage in extracellular loop 6 would result in two fragments of roughly equal size. To determine if fragment B corresponds to either (or both) of the putative 45 kDa terminal regions of FimD, *strep*-tag II analysis of the cleavage products generated during N_{strep}-FimD and TEV-FimD-C_{strep} biogenesis was performed.

Fragment B is generated during protease shaving and pulse chase analysis of N_{strep}-FimD and TEV-FimD-C_{strep} in Δ*tamA* cells, indicating that the modifications do not affect TAM-independent biogenesis of FimD (Figure 4.5.2). However, inspection of the relative fragment sizes did not suggest that fragment B contains the ~1.5 kDa *strep*-tag II (Figure 4.5.3a) and following streptavidin blotting, *strep*-tag II was not detected within fragment B (Figure 4.5.3b), despite detection of

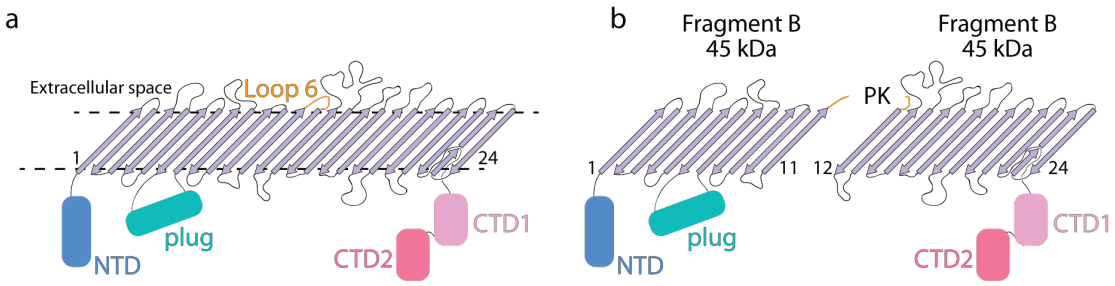


Figure 4.5.1 Predicted identity of fragment B.

a, FimD topology where loop 6 has been highlighted to demonstrate **b,** the putative identities of fragment B following proteolytically cleavage.

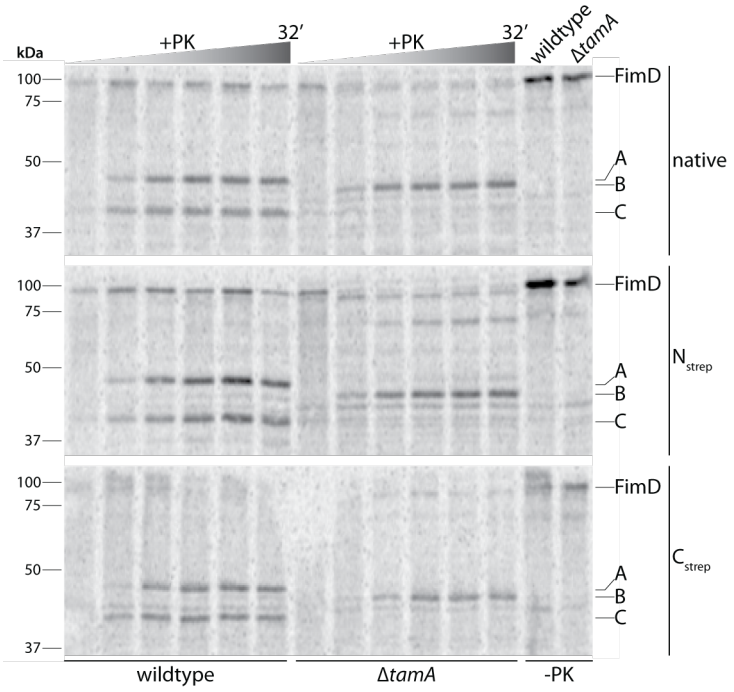


Figure 4.5.2 FimD biogenesis is not affected by the incorporation of *strep*-tag II.

Shown is the "uncropped" version of Figure 4.3.3a. FimD biogenesis was assessed by pulse chase analysis in the indicated strains of *E. coli* BL21 Star™ (DE3) harbouring either: (i) pKS02 (pET-15b containing *fimD*), (ii) pCJS51 (pET-22b(+) containing N_{strep}-*fimD*) or (iii) pCJS29 (pET-15b containing *TEV-fimD-*C_{strep}). Aliquots were taken at 10 seconds, 2, 4, 8, 16 and 32 minutes and treated with (+PK) or without (-PK, last timepoint only) 50 µg mL^{-1} proteinase K (static, 10 min, on ice). Proteins were then TCA-precipitated, washed with acetone and boiled for 3 minutes in SDS sample buffer. Samples were analysed by SDS-PAGE and storage phosphor-imaging. Sizes in kDa are indicated on the left. Full-length FimD and its fragments A, B and C, and the type of FimD analysed are indicated on the right. The time increment is indicated above the autoradiogram as a graded triangle.

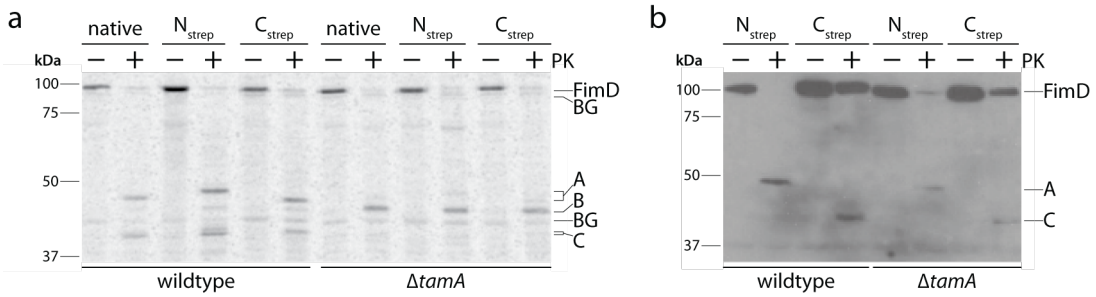


Figure 4.5.3 Fragment B does not contain the N- or C-terminus of FimD.

Shown is the "uncropped" versions of **a,** Figure 4.3.3b and **b,** Figure 4.3.3c. FimD biogenesis was assessed by pulse chase analysis in the indicated strains of *E. coli* BL21 Star™ (DE3) harbouring either: (i) pKS02 (pET-15b containing *fimD*), (ii) pCJS51 (pET-22b(+) containing N_{strep}-*fimD*) or (iii) pCJS29 (pET-15b containing *TEV-fimD-*C_{strep}) as indicated. Aliquots were taken at either 32 (**a**) or 16 (**b**) minutes and treated with (+PK) or without (-PK) 50 µg mL^{-1} proteinase K (static, 10 min, on ice). Proteins were then TCA-precipitated, washed with acetone and boiled for 3 minutes in SDS sample buffer. Samples were analysed by SDS-PAGE and either storage phosphor-imaging (**a**) or Western blotting (**b**) using Precision Protein™ StrepTactin-HRP conjugate (Bio-Rad) for detection. Sizes in kDa are indicated on the left. Full-length FimD, its fragments A and C, and prominent background (BG) bands are indicated on the right. The type of FimD analysed is indicated above the autoradiograms and the strain background is indicated below the autoradiograms.

fragments A or C in the same "Δ*tamA*" lanes. Taken together, fragment B does not map to either termini of FimD, indicating that proteinase K must cleave FimD more than once to generate a protease-resistant internal region corresponding to fragment B.

In order to generate the protease-resistant fragment B, FimD must at least be partially assembled into the outer membrane. To determine whether the known protease-sensitive loops 3 and 7 of FimD were accessible to exogenously added protease, TEV protease shaving analysis of FimD and TEV-FimD-C_{strep} was performed (Figure 4.5.4a). Because TEV protease can only cleave FimD once (within loop 3) and TEV-FimD-C_{strep} once (within loop 7, but not loop 3) during TAM-dependent biogenesis, if either loop is not exposed for proteolysis during TAM-independent usher biogenesis, this will be reflected by the absence of (or significant reduction in) corresponding cleavage products. However, regardless of whether usher biogenesis occurred in the presence or absence of the TAM, TEV proteolysis generated: (i) equal amounts of the 66.5 kDa fragment A* during FimD biogenesis, and; (ii) equal amounts of fragments A and C during TEV-FimD-C_{strep} biogenesis. This data therefore indicates that both loops 3 and 7 are surface exposed during TAM-independent usher biogenesis, indicating that fragment B contains the 24.5 kDa region between loops 3 and 7 (corresponding to residues G225-I457 and about 56 % of fragment B) (Figure 4.5.4b).

Because it is now clear that fragment B represents a central portion of FimD (where the termini of FimD have been cleaved), FimD was internally immunolabelled. A hexahistidine tag (HHHHHHSG) was inserted into three

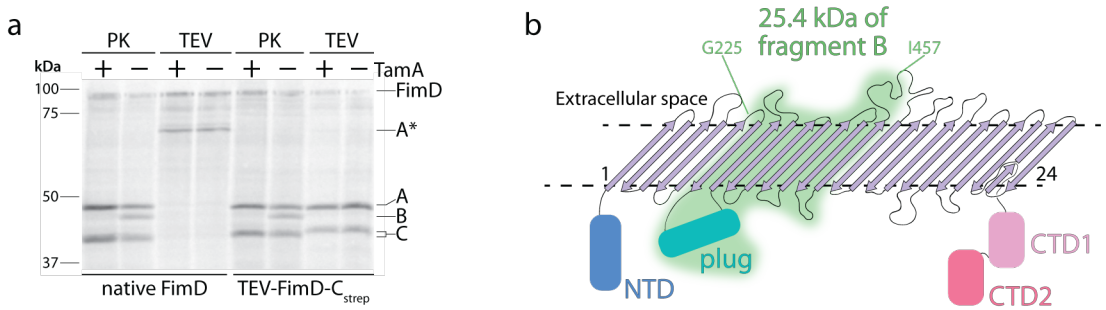


Figure 4.5.4 Fragment B represents a "central" portion of FimD.

a, FimD biogenesis was assessed by pulse chase analysis in *E. coli* BL21 Star™ (DE3) wildtype ("+ TamA") or Δ*tamA* ("- TamA") cells harbouring either: (i) pKS02 (pET-15b containing *fimD*), (ii) pCJS29 (pET-15b containing *TEV-fimD-C*$_{strep}$). Aliquots were taken at 32 minutes and treated with either 50 μg mL^{-1} proteinase K (static, 10 min, on ice) or 100 μg mL^{-1} TEV protease and 1 mM DTT (static, 30 min, 30 °). Proteins were then TCA-precipitated, washed with acetone and boiled for 3 minutes in SDS sample buffer. Samples were analysed by SDS-PAGE and storage phosphor-imaging. Sizes in kDa are indicated on the left. Full-length FimD and its fragments A, A*, B and C are indicated on the right. The strains are indicated above the autoradiogram, and the type of FimD analysed is indicated below the autoradiogram. **b,** FimD topology where the known structured region of fragment B is highlighted in orange.

positions within FimD that would ensure they reside topologically within the periplasm (Figure 4.5.5). $His_{135/136}$-FimD contains a hexahistidine tag at the C-terminal region of the NTD, between P135 and G136 of the mature protein. $His_{211/212}$-FimD contains a hexahistidine tag between P211 and L212, located within periplasmic turn 2 (between β-strands 4 and 5). $His_{563/564}$-FimD contains a hexahistidine tag between S563 and D564, which is located within periplasmic turn 9 (between β-strands 18 and 19).

The fragmentation profile observed following protease shaving and pulse chase analysis of each hexahistidine-tagged construct was comparable to that of native FimD (Figure 4.5.6), indicating that the hexahistidine tag in any of the three positions does not affect FimD biogenesis. Immunoblotting revealed the presence of the hexahistidine tag within: (i) fragment A generated from $His_{135/136}$-FimD and $His_{211/212}$-FimD but not $His_{563/564}$-FimD as expected, and (ii) fragment C generated from $His_{563/564}$-FimD but not $His_{135/136}$-FimD nor $His_{211/212}$-FimD, also as expected (Figure 4.5.7a, and see Figure 4.5.5b for expected results). During usher biogenesis in the absence of the TAM, the hexahistidine tag was detected within fragment B generated from $His_{135/136}$-FimD and $His_{211/212}$-FimD, but not $His_{563/564}$-FimD despite the detection of fragment C in the "$His_{563/564}$-FimD, Δ*tamA*" lane (Figure 4.5.7a).

Taken together, fragment B must represent a portion of FimD that precedes S563, but includes the region from G136 to L212. This extends the known fragment B boundaries determined during TEV protease analysis to G136 and I457 of the mature protein, corresponding to a 35.5 kDa (79 %) portion of fragment

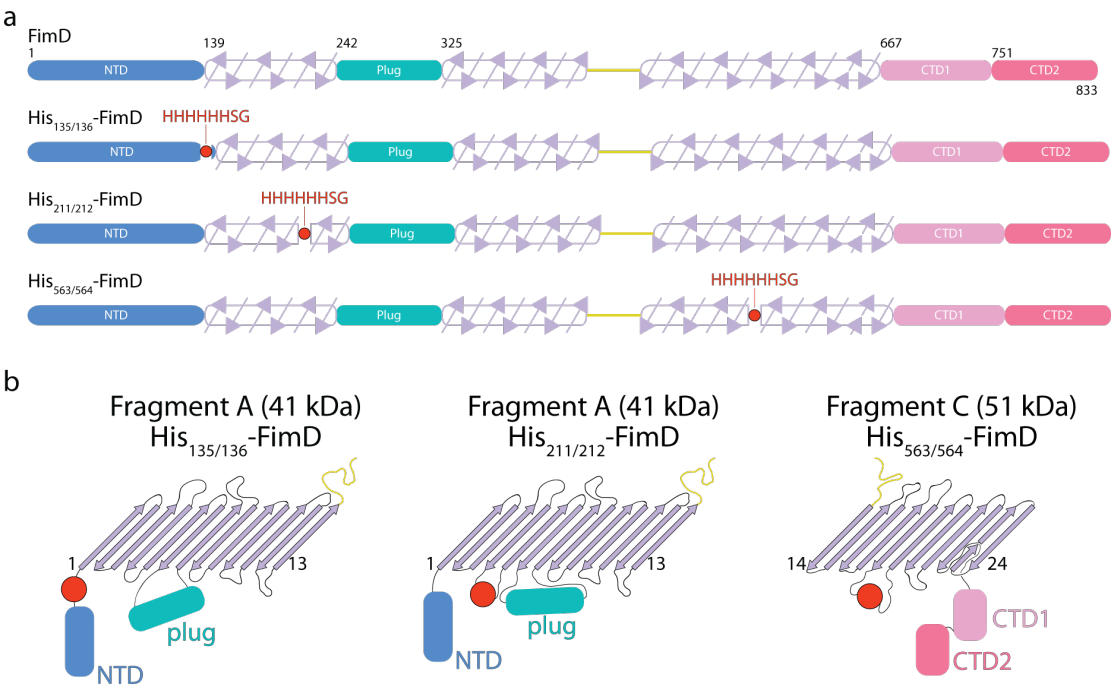


Figure 4.5.5 Structural maps and topologies of FimD modified to include a hexahistidine tag. **a,** Structural maps of unmodified and modified FimD based on the crystal structure (PDB: 3RFZ), where red circles represent the hexahistidine tag. **b,** Topologies of fragments A or C that contain the hexahistidine tag.

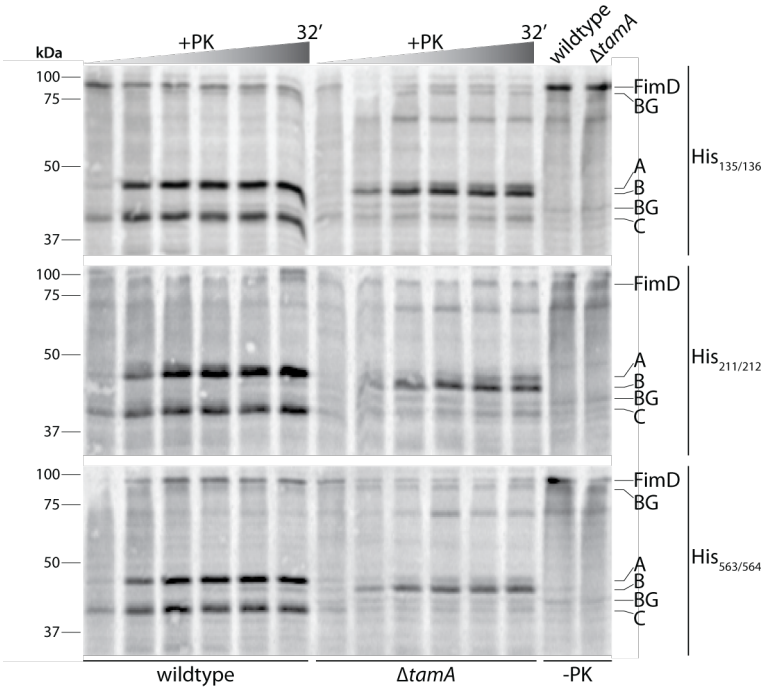


Figure 4.5.6 FimD biogenesis is not affected by the incorporation of hexahistidine tag.

FimD biogenesis was assessed by pulse chase analysis in the indicated strains of *E. coli* BL21 Star™ (DE3) harbouring either: (i) pCJS64 (pET-15b containing *His$_{135/136}$-fimD*), (ii) pCJS65 (pET-15b containing *His$_{211/212}$-fimD*) or (iii) pCJS66 (pET-15b containing *His$_{563/564}$-fimD*). Aliquots were taken at 10 seconds, 2, 4, 8, 16 and 32 minutes and treated with (+PK) or without (-PK, last timepoint only) 50 μg mL^{-1} proteinase K (static, 10 min, on ice). Proteins were then TCA-precipitated, washed with acetone and boiled for 3 minutes in SDS sample buffer. Samples were analysed by SDS-PAGE and storage phosphor-imaging. Sizes in kDa are indicated on the left. Full-length FimD, its fragments A, B and C, prominent background (BG) bands and the type of FimD analysed are indicated on the right. The time increment is indicated above the autoradiogram as a graded triangle. The strains are indicated above and below the autoradiogram.

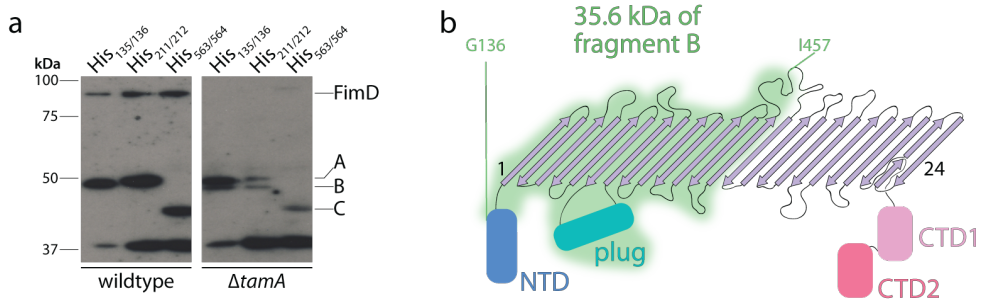


Figure 4.5.7 The boundaries of fragment B.

a, FimD biogenesis was assessed by pulse chase analysis in the indicated strains of *E. coli* BL21 Star™ (DE3) harbouring either: (i) pCJS64 (pET-15b containing *His*$_{135/136}$*-fimD*), (ii) pCJS65 (pET-15b containing *His*$_{211/212}$*-fimD*) or (iii) pCJS66 (pET-15b containing *His*$_{563/564}$*-fimD*). After 16 minutes, samples were treated with 50 µg mL^{-1} proteinase K (static, 10 min, on ice). Proteins were then TCA-precipitated, washed with acetone and boiled for 3 minutes in SDS sample buffer. Samples were analysed by SDS-PAGE and Western blotting using anti-hexahistidine antibodies. Sizes in kDa are indicated on the left. Full-length FimD and its fragments A, B and C are indicated on the right. The type of FimD is indicated above the autoradiogram and the strains are indicated below the autoradiogram. **b,** FimD topology where the known structured region of fragment B is highlighted in orange.

B (Figure 4.5.7b). Interestingly, this places the second cleavage site for proteinase K within the NTD, but structural analysis does not provide any clues as to the relative location of the NTD cleavage site. Indeed, while the NTD from active FimD resides within the periplasm (Phan *et al.*, 2011), the precise topology of the NTD for inactive FimD and, more importantly, the pre-mature FimD assembly intermediates remains unknown.

By taking G136 and I457 to be the N- and C-terminal extremities of fragment B respectively, a 45 kDa fragment would be generated between G136 and ~A540 or between ~R47 and I457, indicating that fragment B resides within a ~54.5 kDa region approximately between residues R47 and A540. To delineate the boundaries of fragment B, non-radiolabelled fragment B was excised from polyacrylamide gels, subjected to in-gel tryptic digestion and analysed by LC-MS/MS. Unfortunately, the boundaries of fragment B could not be further extended, because no significant fragments outside the range of G136-I457 were identified. Inspection of the normalised density plots for the proteolytic fragments generated during FimD biogenesis in each of the six strains (Figures 4.4.3b,d, 4.4.4b, 4.4.6b, 4.4.7b,d) revealed that rather than an equal distribution of fragments A and C, fragment A levels were always greater than those of fragment C.

Perhaps during TAM-independent FimD biogenesis, two major assembly intermediates are formed before the final correct usher conformation is adopted (Figure 4.5.8a). The first assembly intermediate has a protease-sensitive N- and C- terminus, and so generates fragment B. Following a relatively slow assembly

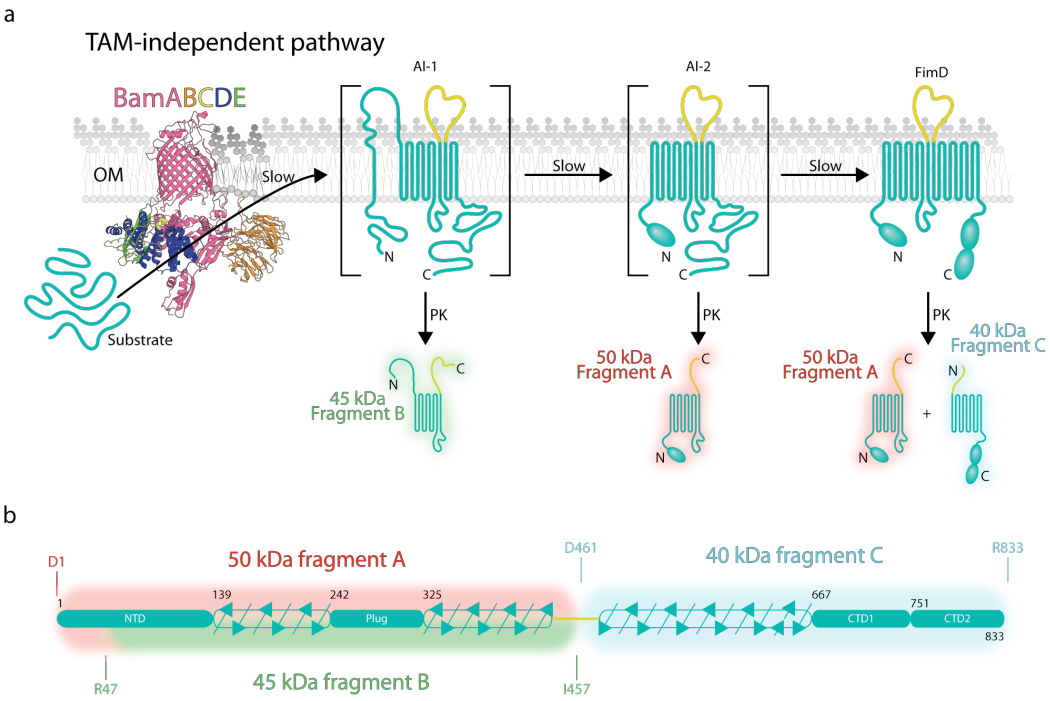


Figure 4.5.8 TAM-independent FimD biogenesis.

a, Cartoon depiction of the TAM-independent FimD biogenesis pathway. A ribbon diagram of the BAM complex is shown (PDB: 5D0O), coloured as in Figure 1.9.3, where the partial OM destability (darker regions of the bilayer) caused by the BAM complex is also shown. FimD is depicted (not to scale) in teal with loop 7 coloured yellow, with its N- and C-termini indicated. Two FimD assembly intermediates (AI) are shown in square brackets, where each of fragments A, B and/or C are indicated below the relevant structures they may be generated from following addition of exogenous proteinase K (PK). **b,** FimD structural map based on the crystal structure (PDB: 3RFZ) and coloured as in "**a**". The boundaries of fragments A (red) and C (blue), and the likely boundaries of fragment B (green), are shown.

process, the remaining N-terminal portion is assembled correctly, and so fragment A may be generated, thereby explaining the greater proportion of fragment A than fragment C. Following a second relatively slow assembly process, the final correct structure is adopted, which generates both fragments A and C. If this hypothesis is true, then fragment B is comprised of the 45 kDa C-terminal region of fragment A (approximately residues R47-I457), where the 5 kDa N-terminal region was removed by exogenous protease (Figure 4.5.8b).

4.6 Other fimbrial ushers require the TAM for assembly

To determine if other fimbrial ushers also require the TAM for efficient assembly, pulse chase analysis of six other fimbrial ushers (Figure 4.6.1) was performed. P fimbriae (π clade) are encoded by the *pap* gene cluster (pyelonephritis-associated pili) and, along with type 1 fimbriae, are the best-characterised and most-studied chaperone-usher systems (Figure 4.6.2). There are three known tip adhesin (*papG*) alleles with products that mediate nearly identical binding to a Gal-α(1-4)β-Gal moiety of glycolipids, but have different host specificity (Strömberg *et al.*, 1990).

Despite the presence of multiple *papG* alleles, antigenic diversity among P fimbriae is attributable to the major fimbrial subunit (PapA), where 11 *papA* alleles are known (Johnson *et al.*, 2000). Interestingly, it is not unusual for bacteria to harbour more than one copy of the *pap* gene cluster, where each cluster may contain different *papA* and/or *papG* alleles (Johnson *et al.*, 2000). Uropathogenic *E. coli* CFT073 contains ten chaperone-usher fimbrial clusters, including two *pap* clusters on separate pathogenicity-associated islands (PAI): PAI-1 and

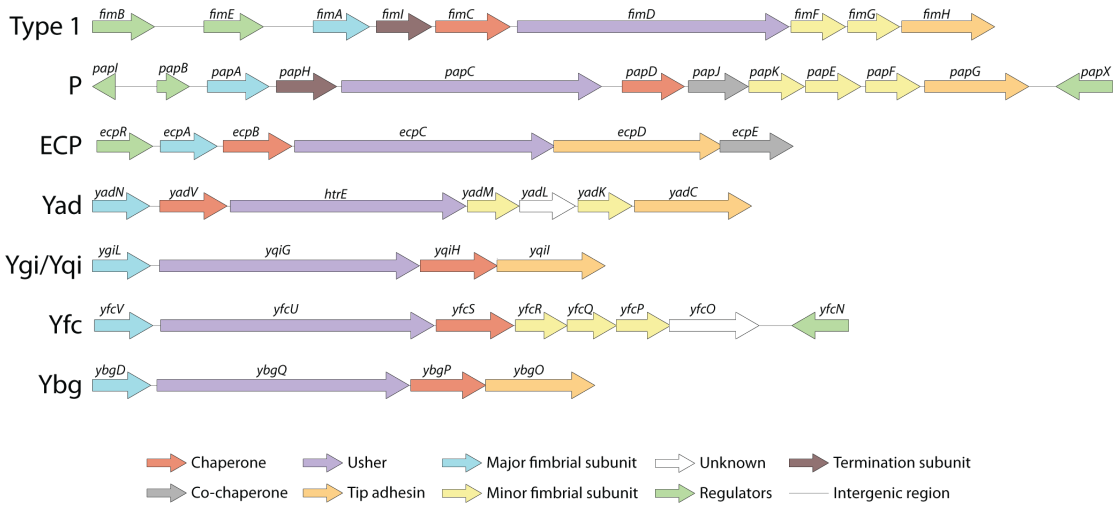


Figure 4.6.1 Genetic organisation of various chaperone-usher fimbrial loci.

Seven fimbrial clusters are shown, which are derived from the same strain where the usher that was analysed in subsequent pulse chase experiments originated from. Fimbrial types indicated on the left and genetic organisation is as per *E. coli* K-12 MG1655 (Type 1 and Ybg fimbriae), enteropathogenic *E. coli* E2348/69 (ECP fimbriae), or uropathogenic *E. coli* CFT073 (P, Yad, Ygi/Yqi, and Yfc fimbriae). Locus tags for each loci are: b4312-20 (Type 1 fimbriae), b0716-19 (Ybg fimbriae), E2348C_RS01300-25 (ECP fimbriae), c3582-93 where *papB* was later manually added by Luo *et al.* (2009) (P fimbriae), c0166-72 (Yad fimbriae), c3791-4 (Ygi/Yqi fimbriae) and c2877-84 (Yfc fimbriae). Coloured arrows represent genes, and arrow length is proportional to the size of each gene.

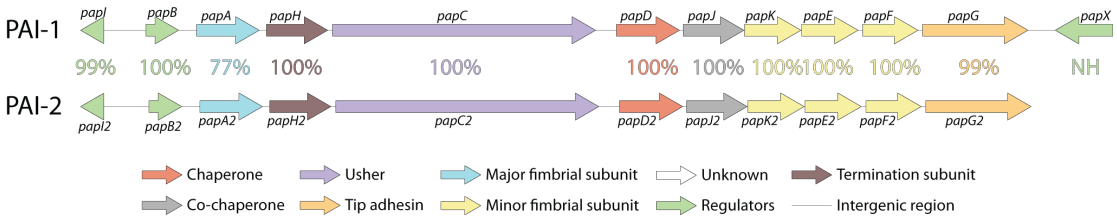


Figure 4.6.2 Uropathogenic *E. coli* CFT073 harbours two copies of the *pap* fimbrial cluster.

Genetic organisation of the *pap* clusters from both pathogenicity-associated islands (PAI-1 and PAI-2) with amino acid identity shown between homologous genes, except for *papX* in PAI-1, which has no homologue (NH) in PAI-2. Amino acid sequence identities were calculated using the MUSCLE algorithm. Locus tags for the *pap* cluster in PAI-1 are c3582-93 and in PAI-2 are c5179-89. The original annotation of uropathogenic *E. coli* CFT073 (Welch *et al.*, 2002) did not include *papB* in either PAI, and so were later manually added by Luo *et al.* (2009): the locus tag for *papB* in PAI-1 is c0212r and in PAI-2 is c0290r. Coloured arrows represent genes, and arrow length is proportional to the size of each gene.

PAI-2 (Welch *et al.*, 2002) (Figure 4.6.2a). While both *papG* are allele II, *papA* from PAI-1 is allele F7-2 and *papA* from PAI-2 is allele F7-1 (Johnson *et al.*, 2000). Despite these differences, both *pap* clusters contain identical *papC* (Figure 4.6.2a), so there was no need to clarify which PapC was used in subsequent pulse chase analysis experiments.

In contrast to FimD biogenesis in wildtype cells, where protease shaving analysis generated fragments A and C, the 89 kDa full-length PapC was relatively protease-resistant in wildtype cells (Figure 4.6.3a). FimD's proteolytic degradation was attributed to the protease-sensitivity of its 42-residue loop 7, but the equivalent loop in PapC is only 22-residues long and does not appear to contain any unstructured, protease-sensitive region (Figure 4.6.3b-c). During PapC biogenesis in the absence of the TAM however, protease shaving analysis revealed that a protease-sensitive PapC assembly intermediate was produced that was rapidly proteolytically degraded into a ~60 kDa fragment D, indicating that PapC biogenesis requires the TAM.

The ECP fimbriae (α clade) are involved in host cell adhesion (specifically HeLa, HTB-4 and HEp-2 epithelial cells), invasion and biofilm formation (Lehti *et al.*, 2010; Rendón *et al.*, 2007; Saldaña *et al.*, 2014). They were originally known as Yag fimbriae, before Pouttu *et al.* (2001) renamed them to Mat (Meningitis-associated and temperature-regulated) fimbriae based on their findings that only distinct *E. coli* pathotypes produced Mat fimbriae. However, Rendón *et al.* (2007) demonstrated that Mat fimbriae may be produced by the majority of *E. coli* commensal and pathogenic strains and therefore renamed them to ECP (*E. coli*

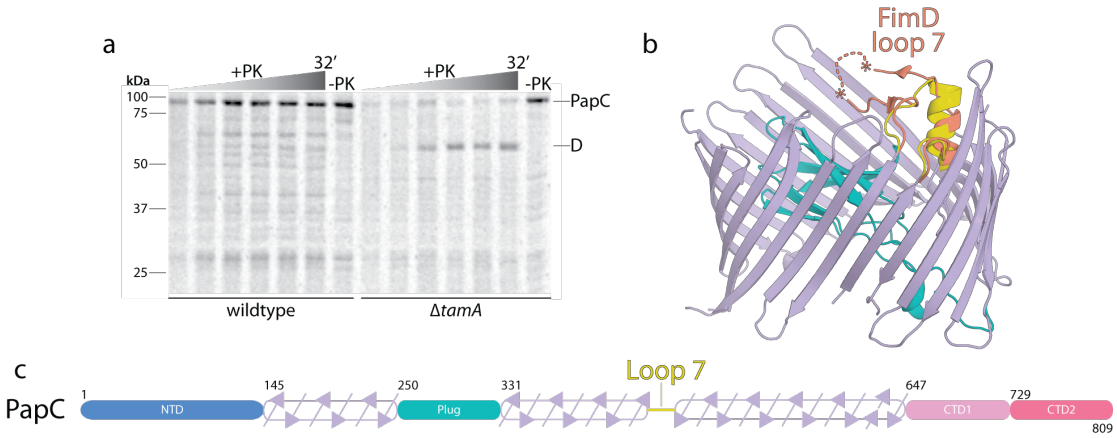


Figure 4.6.3 PapC requires the TAM for assembly.

a, Using *papC* from uropathogenic *E. coli* CFT073, PapC biogenesis was assessed by pulse chase analysis in the indicated strains of *E. coli* BL21 Star™ (DE3) harbouring pCJS32 (pET-22b(+) containing *papC*). Aliquots were taken at 10 seconds, 2, 4, 8, 16 and 32 minutes, and treated with (+PK) or without (-PK, last timepoint only) 50 μg mL^{-1} proteinase K (static, 10 min, on ice). Proteins were then TCA-precipitated, washed with acetone and boiled for 3 minutes in SDS sample buffer. Samples were analysed by SDS-PAGE and storage phosphor-imaging. Sizes in kDa are indicated on the left. Full-length PapC and its fragment D are indicated on the right. The time increment is indicated above the autoradiogram as a graded triangle and the strain is indicated below. **b,** Ribbon diagram of the transmembrane domain from inactive PapC (PDB: 2VQI) and loop 7 from inactive FimD (PDB: 3OHN), superimposed using the "align" function of Pymol (with a 4.0 Å cutoff, following 9 cycles). Asterisks indicate the unstructured region of FimD's loop 7, between residues 454 and 471. **c,** Structural map of mature PapC with amino acid positions indicated at the start and/or end of each domain.

common pilus) fimbriae. Pouttu *et al.* (2001) reported that these fimbriae were only produced at low temperatures (20 °C), and Rendón *et al.* (2007) confirmed the low temperature requirement under aerobic and anaerobic growth, but found that low (26 °C) or high temperatures (37 °C) growth in an atmosphere containing 5 % CO_2 greatly enhances *ecp* expression.

Using *ecpC* from enteropathogenic *E. coli* E2348/69, pulse chase analysis of EcpC biogenesis in wildtype cells revealed that EcpC migrated as two distinct bands: the full-length 88 kDa EcpC from the earliest timepoint, and a ~70 kDa fragment F from about 8 minutes (Figure 4.6.4a). Intriguingly, the presence of fragment F is independent of the presence of exogenously added protease, indicating that: (i) fragment F is a natural degradation product of EcpC, and (ii) full-length EcpC is relatively protease-resistant. During EcpC biogenesis in the absence of the TAM, protease shaving analysis indicated that a full-length EcpC assembly intermediate was degraded by exogenously added protease into smaller fragments, including the ~75 kDa fragment E, rather than the natural degradation into fragment F indicative of correct assembly, indicating that EcpC likely requires the TAM for efficient assembly.

Yad fimbriae (γ4 clade) are important for biofilm formation and adherence to abiotic surfaces (Korea *et al.*, 2010) or human bladder epithelium (Spurbeck *et al.*, 2011) and appear to be expressed under acid-stress (pH 3.0) (Chingcuanco *et al.*, 2012) or high temperatures (41 °C) (Verma *et al.*, 2016). The Yad fimbrial chaperone was named EcpD based on its homology to the *E. coli* PapD chaperone (Raina *et al.*, 1993) and has only recently been renamed to YadV to avoid confusion

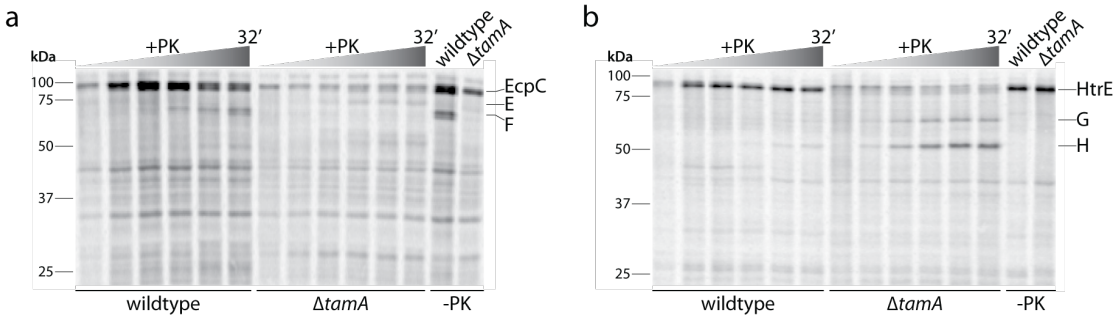


Figure 4.6.4 EcpC and HtrE require the TAM for assembly.

Using *ecpC* from enteropathogenic *E. coli* E2348/69 or *htrE* from uropathogenic *E. coli* CFT073, biogenesis of EcpC (**a**) or HtrE (**b**) was assessed by pulse chase analysis in the indicated strains of *E. coli* BL21 Star™ (DE3) harbouring either **a,** pCJS37 (pETDuet-1 containing *ecpC* in MCS2) or **b,** pCJS49 (pET-15b containing *htrE*). Aliquots were taken at 10 seconds, 2, 4, 8, 16 and 32 minutes, and treated with (+PK) or without (-PK, last timepoint only) 50 μg mL^{-1} proteinase K (static, 10 min, on ice). Proteins were then TCA-precipitated, washed with acetone and boiled for 3 minutes in SDS sample buffer. Samples were analysed by SDS-PAGE and storage phosphor-imaging. Sizes in kDa are indicated on the left. The time increment is indicated above the autoradiogram as a graded triangle and the strain is indicated above and below the autoradiogram. **a,** Full-length EcpC and its fragments E and F are indicated on the right. **b,** Full-length HtrE and its fragments G and H are indicated on the right.

with the ECP fimbrial tip, EcpD. Using *htrE* (High temperature requirement gene E) from uropathogenic *E. coli* CFT073, pulse chase analysis of HtrE biogenesis in wildtype cells indicated that the 92 kDa full-length HtrE was relatively protease-resistant (Figure 4.6.4b). In contrast, protease shaving analysis revealed that in a Δ*tamA* mutant background, HtrE biogenesis produced a protease-sensitive assembly intermediate that was rapidly proteolytically degraded into ~65 kDa fragment G or ~55 kDa fragment H (Figure 4.6.4b). Based on the differential protease-sensitivity, this data indicates that HtrE likely requires the TAM for efficient assembly.

Yqi (or Ygi) fimbriae (π clade) mediate adherence to human kidney epithelium and are important for biofilm formation and colonisation of chicken lungs (Antão *et al.*, 2009; Spurbeck *et al.*, 2011). The YqiG usher from some strains are disrupted by an insertion sequence (*IS21*) element, including lab strain *E. coli* K-12 MG1655 (Figure 4.6.5a). This IS element appears to interrupt the region comprising the signal sequence of *yqiG* and presumably prevents YqiG biogenesis. However, Mohd Yusoff *et al.* (2013) reported that deletion of *yqiG* reduces glucose utilisation when grown anaerobically in glucose minimal media, and abolishes biohydrogen production from glucose fermentation, indicating that presence of a *yqiG* "pseudogene" product may be important for glucose metabolism.

Using *yqiG* from uropathogenic *E. coli* CFT073, pulse chase analysis of YqiG biogenesis in wildtype cells revealed that the 90 kDa full-length YqiG remained mostly protease-resistant, except from 16 minutes where it was partially proteolytically degraded into one of ~45 kDa, ~60 kDa or ~65 kDa fragments. In

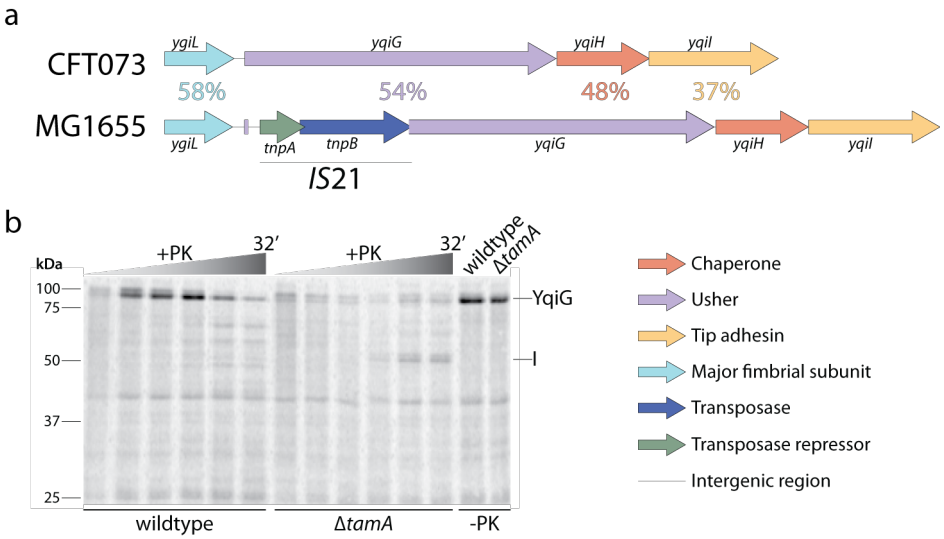


Figure 4.6.5 YqiG requires the TAM for assembly.

a, Genetic organisation of the *ygi/yqi* clusters from *E. coli* K-12 MG1655 and uropathogenic *E. coli* CFT073 with amino acid identity shown between homologous genes, calculated using the MUSCLE algorithm. YqiG from either side of the *IS*21 was combined to calculate identity. Locus tags for the *ygi/yqi* cluster in the MG1655 strain are b3043-8 and in the CFT073 strain are c3791-4. **b,** Using *yqiG* from uropathogenic *E. coli* CFT073, YqiG biogenesis was assessed by pulse chase analysis in the indicated strains of *E. coli* BL21 Star™ (DE3) harbouring pCJS31 (pET-15b containing *yqiG*). Aliquots were taken at 10 seconds, 2, 4, 8, 16 and 32 minutes, and treated with (+PK) or without (-PK, last timepoint only) 50 µg mL^{-1} proteinase K (static, 10 min, on ice). Proteins were then TCA-precipitated, washed with acetone and boiled for 3 minutes in SDS sample buffer. Samples were analysed by SDS-PAGE and storage phosphor-imaging. Sizes in kDa are indicated on the left. Full-length YqiG and its fragment I are indicated on the right. The time increment is indicated above the autoradiogram as a graded triangle and the strains are indicated above and below the autoradiogram.

the absence of the TAM however, protease shaving analysis indicated that YqiG was rapidly proteolytically degraded into smaller fragments at all timepoints, including a ~50 kDa fragment I from about 8 minutes (Figure 4.6.5b). This data indicates that in the absence of the TAM, YqiG assembles into a protease-sensitive intermediate, suggesting that YqiG requires the TAM for efficient assembly.

Yfc fimbriae (π clade) are widely distributed among pathogenic and non-pathogenic *E. coli* (Wurpel *et al.*, 2013), but the YfcU usher of some strains, including the lab strain K-12 MG1655, contains an in-frame stop codon within the coding sequence (Figure 4.6.6a) and is presumably non-functional. However, Korea *et al.* (2010) reported that, despite the premature stop codon, expression of the *yfc* operon mediated adherence to T24 human bladder cells, but whether this alternate YfcU remains functional via translational read-through or another fimbrial usher substitutes to assemble the Yfc fimbriae remains to be determined. Interestingly, Schneider *et al.* (2000) reported that independently disrupting *yfcU* with an insertion sequence improves *E. coli* B strain fitness in glucose minimal media, indicating that the premature stop codon may have been selected for to improve glucose metabolism.

Using *yfcU* from uropathogenic *E. coli* CFT073, pulse chase analysis of YfcU biogenesis in wildtype cells indicated that the full-length ~94 kDa YfcU was relatively protease-resistant (Figure 4.6.6b). During its biogenesis in the absence of the TAM however, the majority of YfcU adopted a protease-sensitive conformation that was rapidly proteolytically degraded into either a ~65 kDa fragment J or ~60 kDa fragment K. Because the protease-sensitive intermediate

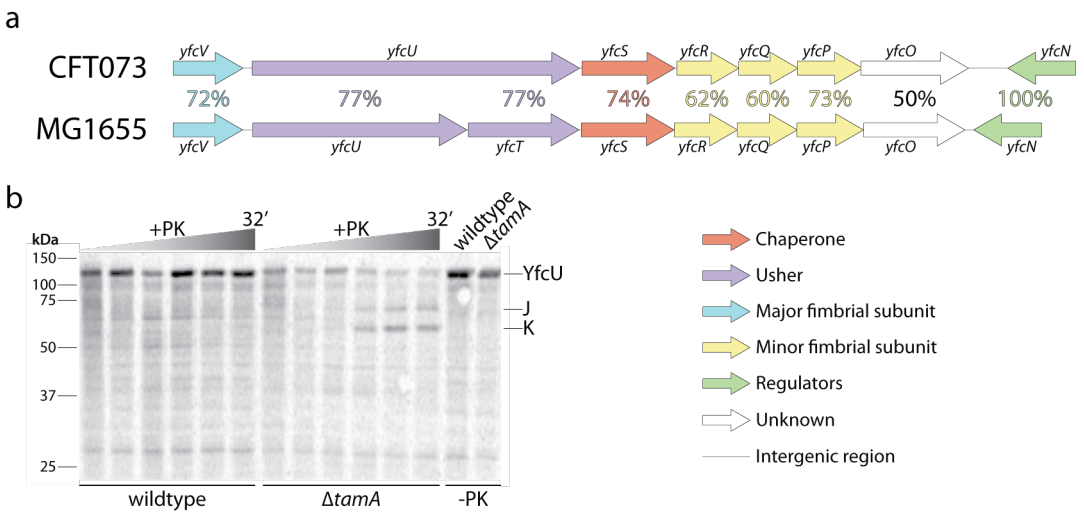


Figure 4.6.6 YfcU requires the TAM for assembly.

a, Genetic organisation of the *yfc* clusters from *E. coli* K-12 MG1655 and uropathogenic *E. coli* CFT073 with amino acid identity shown between homologous genes, calculated using the MUSCLE algorithm. Locus tags for the *yfc* cluster in the MG1655 strain are b2332-9 and in the CFT073 strain are c2877-84. **b,** Using *yfcU* from uropathogenic *E. coli* CFT073, YfcU biogenesis was assessed by pulse chase analysis in the indicated strains of *E. coli* BL21 Star™ (DE3) harbouring pCJS33 (pET-15b containing *yfcU*). Aliquots were taken at 10 seconds, 2, 4, 8, 16 and 32 minutes, and treated with (+PK) or without (-PK, last timepoint only) 50 µg mL^{-1} proteinase K (static, 10 min, on ice). Proteins were then TCA-precipitated, washed with acetone and boiled for 3 minutes in SDS sample buffer. Samples were analysed by SDS-PAGE and storage phosphor-imaging. Sizes in kDa are indicated on the left. Full-length YfcU and its fragment J and K are indicated on the right. The time increment is indicated above the autoradiogram as a graded triangle and the strains are indicated above and below the autoradiogram.

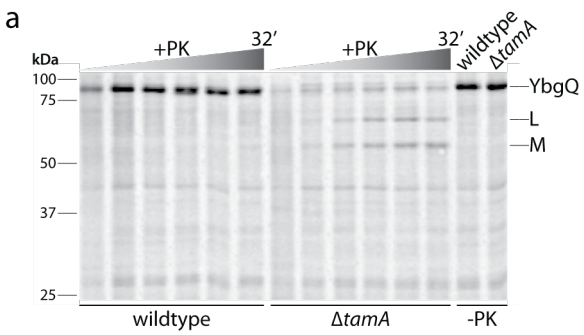


Figure 4.6.7 YbgQ requires the TAM for assembly.

Using *ybgQ* from *E. coli* K-12 MG1655, YbgQ biogenesis was assessed by pulse chase analysis in the indicated strains of *E. coli* BL21 Star™ (DE3) harbouring pCJS40 (pET-15b containing *ybgQ*). Aliquots were taken at 10 seconds, 2, 4, 8, 16 and 32 minutes, and treated with (+PK) or without (-PK, last timepoint only) 50 μg mL^{-1} proteinase K (static, 10 min, on ice). Proteins were then TCA-precipitated, washed with acetone and boiled for 3 minutes in SDS sample buffer. Samples were analysed by SDS-PAGE and storage phosphor-imaging. Sizes in kDa are indicated on the left. Full-length YbgQ and its fragment L and M are indicated on the right. The time increment is indicated above the autoradiogram as a graded triangle and the strains are indicated above and below the autoradiogram.

was only observed during TAM-deficient YfcU assembly, YfcU likely requires the TAM to ensure the correct protease-resistant conformation is adopted efficiently.

Ybg fimbriae (π clade) are also widely distributed among pathogenic and non-pathogenic *E. coli* (Wurpel *et al.*, 2013), but attempts to identify Ybg surface receptors were unsuccessful (Korea *et al.*, 2010). Using *ybgQ* from *E. coli* K-12 MG1655, pulse chase analysis of YbgQ biogenesis in wildtype cells indicated that the 88 kDa full-length YbgQ remained protease-resistant throughout its assembly (Figure 4.6.7). However, in the absence of the TAM, YbgQ assembled into a protease-sensitive intermediate that was rapidly proteolytically degraded into ~70 kDa fragment L or ~60 kDa fragment M (Figure 4.6.7). This data therefore suggests that YbgQ also likely requires the TAM for efficient assembly.

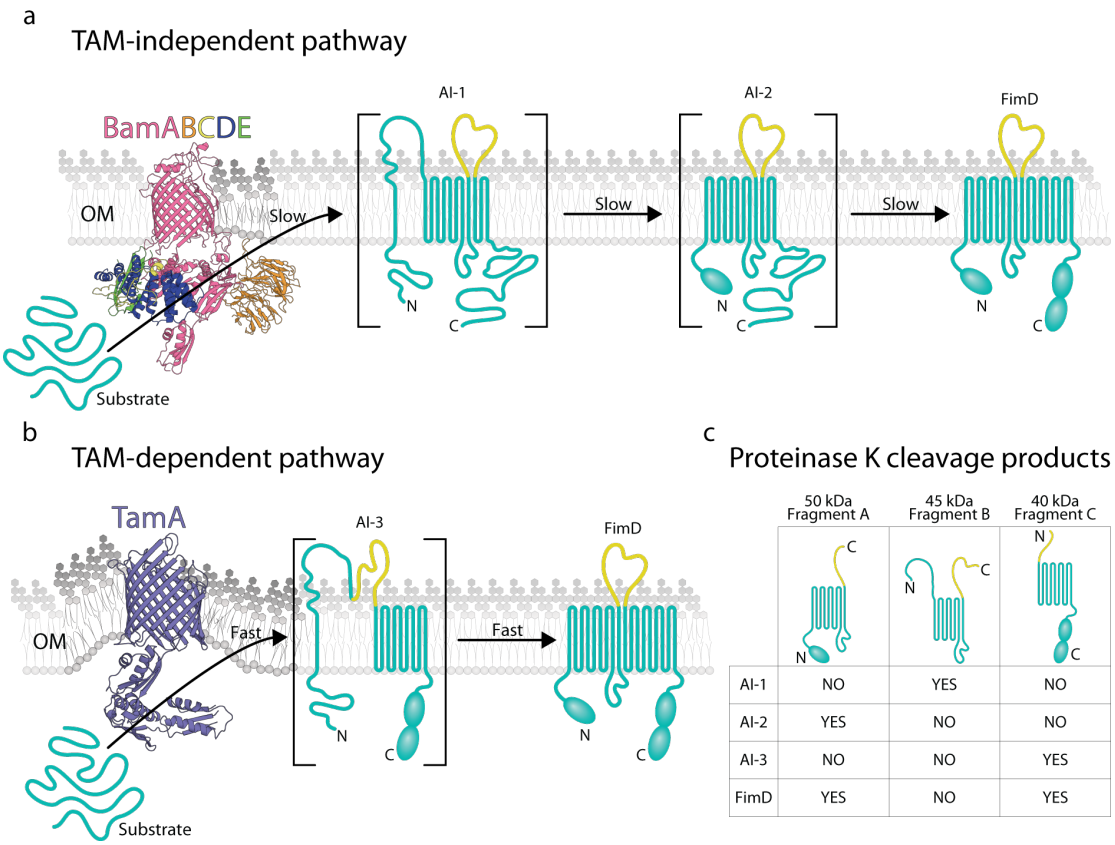
4.7 Discussion

Pulse chase analysis of seven fimbrial ushers (FimD, PapC, EcpC, HtrE, YqiG, YfcU and YbgQ) revealed that they each require the TAM for efficient biogenesis. Unlike the majority of fimbrial ushers that adopt a protease-resistant final conformation, FimD contains a large extracellular loop that is readily cleaved by exogenously added protease to generate two protease-resistant cleavage products: an N-terminal, 50 kDa fragment A and a C-terminal, 40 kDa fragment C. Using this structural characteristic of FimD, the early minutes of FimD biogenesis were probed and the C-terminal fragment was found to form earlier than the N-terminal fragment. This data is consistent with the hypothesis that for the Omp85 insertase family, β-barrel assembly is initiated within the transmembrane domain from the C-terminal β-strands and proceeds toward the N-terminal β-strands in a strand-

by-strand fashion (Estrada Mallarino *et al.*, 2015; Gu *et al.*, 2016; Höhr *et al.*, 2015).

When usher biogenesis was monitored in the absence of the TAM however, the seven fimbrial ushers under investigation were found to adopt a protease-sensitive assembly intermediate. Following protease shaving analysis, these assembly intermediates were rapidly proteolytically degraded into smaller fragments that included at least one prominent cleavage product (i.e. fragments B, D, E, and G-M). Inspection of the 45 kDa fragment B generated from the FimD assembly intermediate revealed that it was comprised of a central portion of FimD approximately corresponding to residues R47-I457. This indicated that assembly in the absence of the TAM was likely initiated from a central β-strand, rather than a C-terminal strand of the transmembrane domain. In accordance with the strand-by-strand assembly hypothesis, once the "final" N-terminal β-strand is assembled, the β-barrel is still not complete. Because assembly appears to initiate from a central region, it is likely that the 7-10 C-terminal transmembrane β-strands were "skipped". Nevertheless, the correct FimD tertiary structure is eventually formed, albeit slowly, but whether the BAM complex is involved in this process or has "disengaged" following the partial β-barrel assembly remains to be seen.

Taken together, the fimbrial ushers assemble via one of two pathways (Figure 4.7.1). In the TAM-dependent pathway, the TAM catalyses the rapid biogenesis of fimbrial ushers where β-barrel assembly is initiated from a C-terminal β-strand. In the TAM-independent pathway, the BAM complex instead catalyses fimbrial usher biogenesis, but β-barrel assembly is initiated from a central β-strand and the correct usher tertiary structure is much slower to form.



	50 kDa Fragment A	45 kDa Fragment B	40 kDa Fragment C
AI-1	NO	YES	NO
AI-2	YES	NO	NO
AI-3	NO	NO	YES
FimD	YES	NO	YES

Figure 4.7.1 Proposed mechanism of fimbrial usher assembly via TAM-dependent and TAM-independent pathways.

Fimbrial usher biogenesis is shown, using FimD biogenesis as an example. Ribbon diagrams of the BAM complex (PDB: 5D0O) and TamA (PDB: 4C00) are coloured as in Figure 1.9.3. FimD is depicted (not to scale) in teal with loop 7 coloured yellow, and its N- and C-termini indicated. Three FimD assembly intermediates (AI) are shown in square brackets. **a,** Cartoon depiction of the TAM-independent FimD biogenesis pathway, where the partial OM destability caused by the BAM complex is shown as darker regions of the bilayer. **b,** Cartoon depiction of the TAM-dependent FimD biogenesis pathway. Although TamB has been omitted for clarity, the increased destability of the OM caused by the TAM is shown as darker regions of the bilayer. **c,** A table showing the cleavage products generated when any of the indicated structures are subjected to exogenous proteinase K treatment.

Using fragments A, B and C generated during FimD biogenesis as the indicator for TAM-dependent and TAM-independent assembly, FimD was shown to assemble via both pathways if both the BAM complex and the TAM were present (wildtype, Δ*bamB*, Δ*bamC*, and Δ*bamE* cells). Furthermore, if the proportion of the TAM was lowered (as in Δ*bamB* or Δ*bamE*) or increased (via overexpression of TamA in a Δ*bamB* background), the proportion of FimD assembled via the TAM was similarly lowered or increased. If FimD has an equal chance of assembling via the TAM-dependent and TAM-independent pathways, the TAM must be significantly more abundant than the BAM complex. However, Wiśniewski and Rakus (2014) reported that the average copy number of BamA is about 14 times greater than that of TamA, indicating that perhaps: (i) the TAM's affinity for FimD is greater than the BAM complex's and/or (ii) the TAM has privileged access to the cargo of periplasmic chaperones.

In the course of this work on the assembly of FimD, I. D. Hay investigated the induction of *fimD* expression in the presence or absence of the TAM (Stubenrauch *et al.*, 2016). Uropathogenic *E. coli* strains CFT073 and UTI89 were modified to contain an inducible type 1 fimbrial expression system. On induction, both strains were capable of adhering to mannosylated surfaces, a feature that was significantly attenuated in isogenic Δ*tamAB* mutants within a physiologically relevant timeframe (20 minutes). On restoration of *tamAB in trans*, the complemented strains showed complementation of adherence, thereby establishing that the TAM is important for efficiently assembling the fimbrial ushers used to generate type 1 fimbriae in a physiologically relevant timeframe.

Uropathogenic *E. coli* are the main etiological agent responsible for urinary tract infections, causing up to 85 % of community-acquired urinary tract infections (Ronald, 2002; Zhang & Foxman, 2003). Other etiological agents include *K. pneumoniae* and *P. mirabilis*, which are also especially prevalent during (i) nosocomial catheter-associated and (ii) complicated urinary tract infections, respectively (Ronald, 2002; Zhang & Foxman, 2003). It is known that chaperone-usher fimbriae from *K. pneumoniae* (type 1 fimbriae) and *P. mirabilis* (MRP fimbriae) are important for urinary tract infections (Murphy *et al.*, 2013; Zunino *et al.*, 2003), and that TAM-negative mutants of *K. pneumoniae* and *P. mirabilis* are significantly outcompeted by their respective isogenic, TAM-positive parent strains (Burall *et al.*, 2004; Struve *et al.*, 2003). Taken together, it is tempting to speculate that the TAM is also responsible for assembling the chaperone-usher fimbriae in a physiologically relevant timeframe for these uropathogens. Interestingly, Burall *et al.* (2004) did show that the major fimbrial subunit (MrpA) was still produced in *P. mirabilis* TAM mutants, but the authors only analysed overnight suspensions incubated *in vitro*, which is not reflective of the *in vivo* conditions required for the initiation of a urinary tract infection.

Similarly to Ag43, which also requires the TAM for biogenesis (Selkrig *et al.*, 2012), chaperone usher fimbriae are tightly regulated by phase variation (Henderson *et al.*, 1999; van der Woude & Bäumler, 2004). While expression of the type 1 fimbriae are usually turned "off", following phase switching to the "on" position, each bacterium rapidly produces 100-500 type 1 fimbriae to ensure adherence to host cells (Gally *et al.*, 1993; Lowe *et al.*, 1987). The evolution of the TAM may therefore have been an important evolutionary requirement for the rapid

assembly of the various fimbrial ushers and AIDA-like autotransporters when expression is turned on, that could otherwise be assembled by the BAM complex, but outside of an appropriate physiological timeframe.

Chapter 5 – Other substrates of the TAM

5.1 Introduction

From the models that have been proposed for β-barrel assembly catalysed by the BAM complex (Gu *et al.*, 2016; Noinaj *et al.*, 2015), it is not immediately obvious how an OMP with globular domains or quaternary structure is assembled. For two families of proteins, AIDA-like ATs (Selkrig *et al.*, 2012) and fimbrial ushers (Section 4), it is clear that the TAM is also involved in assembling OMPs that contain extracellular or periplasmic globular domains. To address the extent to which the TAM plays a general role in assembling proteins that diverge from the classic β-barrel architecture, pulse chase analysis of further candidate TAM substrates was investigated (Figure 5.1.1).

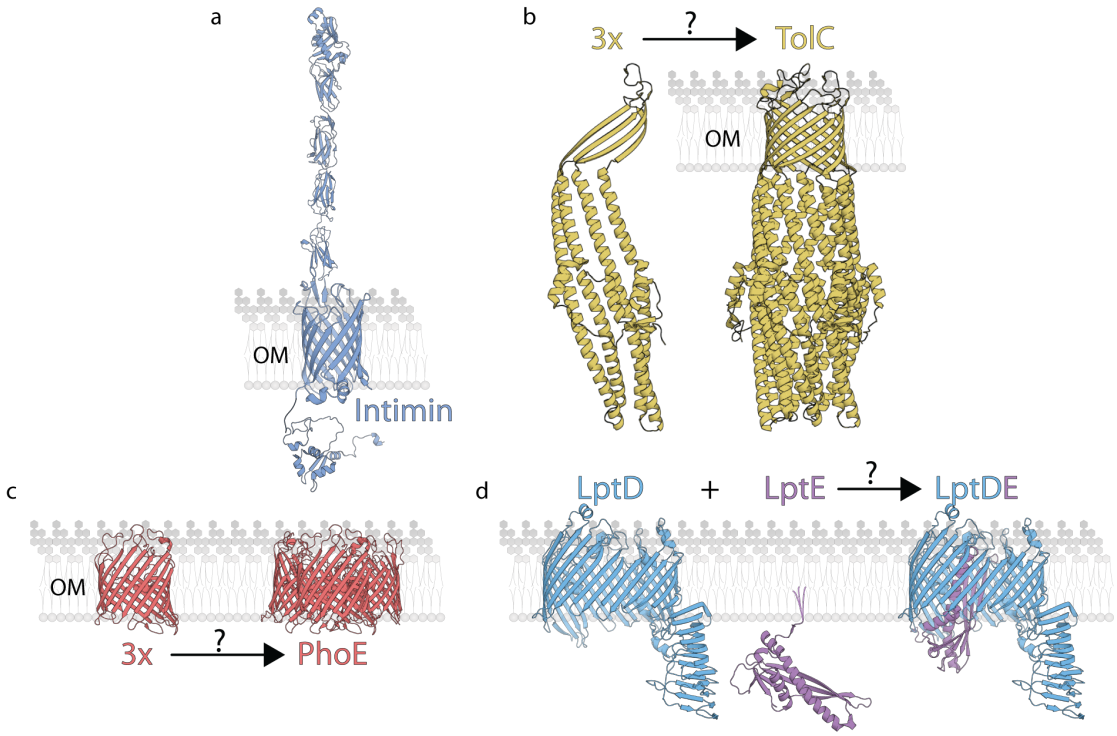


Figure 5.1.1 The remaining candidate TAM substrates.

Shown here are the complex domain architectures of the remaining candidate TAM substrates as per Figure 3.2.1. **a,** Intimin (PDB: 2MPW, 4E1S, 1F02 and residues 455-653 modelled after PDB: 1CWV using Phyre2). **b,** TolC (PDB: 1EK9). **c,** PhoE (PDB: 1PHO). **d,** LptDE (PDB: 4Q35).

5.2 Inverse autotransporter family

The inverse autotransporters are a large family of proteins that assemble via the T5SS pathway. They contain a large extracellular globular domain, usually involved in cell adhesion, but sometimes host-cell invasion (Leo *et al.*, 2015). The N-terminal β-barrel domain typifies the inverse autotransporter family and is highly conserved at the tertiary structure level, whereas their C-terminal extracellular passenger domain is more structurally diverse (Heinz *et al.*, 2016) (Figures 5.1.1a and 5.2.1). Despite this, many members harbour immunoglobulin-like folds within their passenger domain (Bodelón *et al.*, 2013), a motif also observed in the TAM-dependent fimbrial ushers (Section 4.1). In a functional sense, the two best-studied examples are intimin from pathogenic strains of *E. coli* and invasin from *Yersinia* species, which have both been shown to depend on the BAM complex for insertion into the OM (Bodelón *et al.*, 2009; Oberhettinger *et al.*, 2012). However, in light of the recent observation that TamA itself is also dependent on the BAM complex for OM insertion (Dunstan *et al.*, 2015), it is important to dissect the contribution the TAM might have on the biogenesis of inverse autotransporter proteins.

Intimin is one of the major virulence factors associated with diarrhoeagenic *E. coli* pathotypes, specifically enteropathogenic and enterohaemorrhagic *E. coli* strains; *C. rodentium*, and; *Hafnia alvei* (Leo *et al.*, 2015). Intimin binds to Tir, which is the translocated intimin receptor injected into a host cell via a T3SS. During disease, intimin is directly responsible for generating "attaching and effacing" lesions in mammalian intestinal epithelia, which are an histopathological feature involving actin pedestal formation and subsequent intimate host-cell

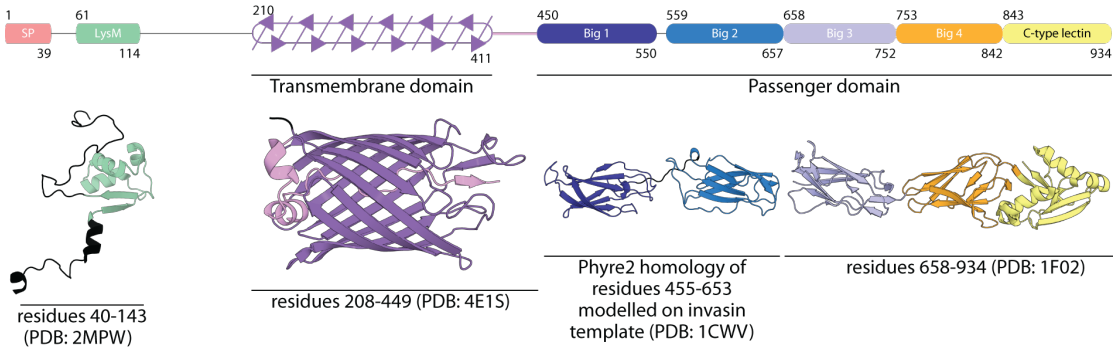


Figure 5.2.1 Domain architecture of intimin.

A structural map of intimin with domain boundaries as per the corresponding structures depicted below. Identities of each structure with PDB reference codes are also indicated, where bacterial immunoglobulin-like (Big) domains are numbered sequentially starting from the N-terminal Big domain.

attachment to promote disease persistence (Lai *et al.*, 2013). Using intimin from enteropathogenic *E. coli* E2348/69, intimin biogenesis was assessed by pulse chase analysis in the presence or absence of TamA (Figure 5.2.2). Interestingly, the ~98 kDa mature intimin forms a doublet that includes a ~115 kDa species, which is too large to correspond to the pre-intimin precursor. Intimin is known to fold into a functional dimer *in vivo* (Touzé *et al.*, 2004), but in addition to the ~115 kDa intimin species being too small to correspond to an intimin dimer, dimerisation should be abolished under the denaturing conditions of SDS-PAGE. The 115 kDa species could correspond to the fully reduced intimin species as observed and suggested by Bodelón *et al.* (2009, see Figure 2 therein), indicating that the smaller intimin band corresponded to the oxidised intimin species that may have been insufficiently reduced by DTT (or was capable of regenerating cystine following reduction).

Regardless, following its biogenesis in the presence of the TAM, intimin becomes relatively protease-resistant; whereas in a Δ*tamA* background, intimin is degraded into a doublet of proteolytic fragments (Figure 5.2.2a). This doublet likely arises from the mature intimin doublets, where the reduced intimin species had degraded into fragment O and the oxidised intimin species was degraded into fragment P. To confirm this defect was due to the absence of the TAM, TamA was reintroduced using a complementation vector (pCJS69) (Figure 5.2.2b) and found to be capable of restoring the protease-resistance of intimin. This was adjudged by the loss of fragments O and P and restoration of full-length intimin levels, indicating that the TAM indeed contributes to the biogenesis of intimin.

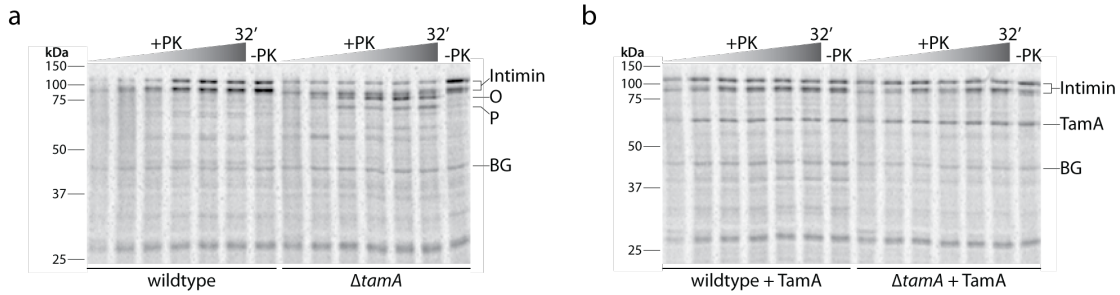


Figure 5.2.2 Intimin biogenesis requires the TAM.

Intimin (EaeA) biogenesis was assessed by pulse chase analysis in the indicated *E. coli* Star™ (DE3) strains harbouring pCJS30 (pET-15b with *eaeA* encoded) in addition to either **a,** the pACYCDuet-1 base vector, or **b,** the pCJS69 (pACYCDuet-1 with *tamA* encoded) complementation vector (indicated by "+ TamA"). At 10 seconds, 2, 4, 8, 16 and 32 minutes, aliquots were taken and treated with (+PK) or without (-PK, last timepoint only) 50 μg mL^{-1} proteinase K (static, 10 min, on ice). Proteins were then TCA-precipitated, washed with acetone and boiled for 3 minutes in SDS sample buffer. Analysis was by SDS-PAGE and storage phosphorimaging. Sizes in kDa are indicated on the left. TamA, the full-length intimin doublet, its fragments O and P, and the prominent background (BG) bands are indicated on the right. The time increment is indicated above the autoradiograms as a graded triangle and the strain is indicated below the autoradiograms.

A comprehensive bioinformatic analysis of the inverse autotransporter family revealed that, rather than the historical archetypes (intimin and invasin), most members of the family resembled FdeC (Heinz *et al.*, 2016). FdeC has a similar architecture to intimin, except that it contains ten immunoglobulin-like domains (Figure 5.2.3) whereas intimin has four (Heinz *et al.*, 2016) (Figure 5.2.1). Unlike intimin however, which is restricted to relatively few *E. coli* pathotypes, FdeC is broadly conserved in *E. coli* and is important for a variety of diseases. In uropathogenic *E. coli*, FdeC was shown to contribute to urinary tract infections (Nesta *et al.*, 2012), and in enterotoxigenic *E. coli*, the direct FdeC homologue - more commonly known as EaeH in these strains - is important for host cell attachment and facilitates heat-labile toxin delivery (Sheikh *et al.*, 2014).

To ascertain whether FdeC also requires the TAM, the biogenesis of FdeC (derived from uropathogenic *E. coli* CFT073) was also assessed by pulse chase analysis in the presence or absence of TamA (Figure 5.2.4). Mature FdeC (150 kDa) was found to be strongly protease-sensitive in both wildtype and Δ*tamA* backgrounds (Figure 5.2.4a). In wildtype cells, FdeC was degraded into a series of proteolytic fragments, including two prominent bands at ~100 kDa and ~50 kDa, respectively referred to as fragments Q and R. Subsequent LC-MS/MS analysis revealed that fragment R included the N-terminal transmembrane domain, whereas fragment Q represented the remaining C-terminal passenger domain. Although in a Δ*tamA* background FdeC was similarly degraded into fragments Q and R, fragment R was much less prominent and the overall proteolytic fragmentation pattern was not comparable to wildtype. This indicated that FdeC was slower to adopt the correct tertiary structure in the absence of the TAM. To confirm this, TamA was

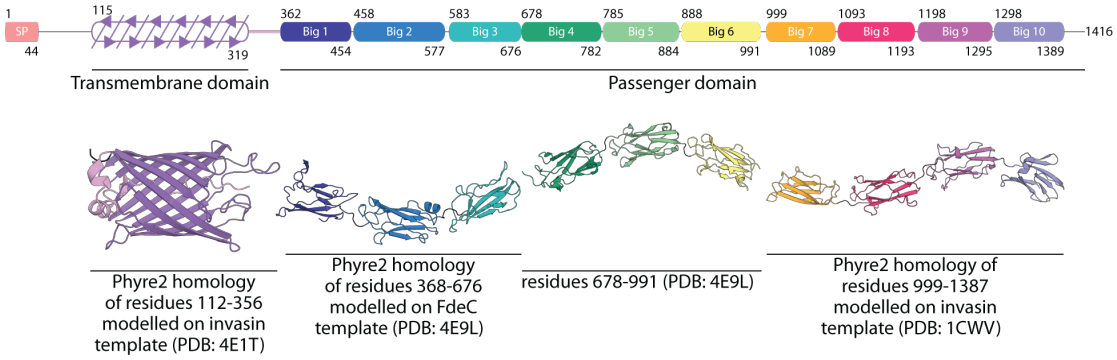


Figure 5.2.3 Domain architecture of FdeC.

A structural map of FdeC with domain boundaries as per the corresponding structures depicted below. Identities of each structure with PDB reference codes are also indicated, where bacterial immunoglobulin-like (Big) domains are numbered sequentially starting from the N-terminal Big domain.

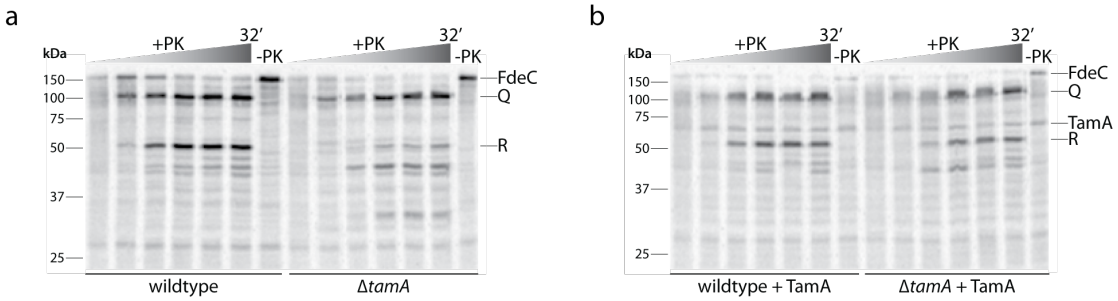


Figure 5.2.4 FdeC biogenesis requires the TAM.

FdeC biogenesis was assessed by pulse chase analysis in the indicated *E. coli* Star™ (DE3) strains harbouring pCJS50 (pET-15b with *fdeC* encoded) in addition to either **a,** the pACYCDuet-1 base vector, or **b,** the pCJS69 (pACYCDuet-1 with *tamA* encoded) complementation vector (indicated by "+ TamA"). At 10 seconds, 2, 4, 8, 16 and 32 minutes, aliquots were taken and treated with (+PK) or without (-PK, last timepoint only) 50 μg mL^{-1} proteinase K (static, 10 min, on ice). Proteins were then TCA-precipitated, washed with acetone and boiled for 3 minutes in SDS sample buffer. Analysis was by SDS-PAGE and storage phosphorimaging. Sizes in kDa are indicated on the left. TamA, full-length FdeC and its fragments Q and R are indicated on the right. The time increment is indicated above the autoradiograms as a graded triangle and the strain is indicated below the autoradiograms.

reintroduced by complementation (Figure 5.2.4b) and found to be capable of restoring the wildtype proteolytic fragmentation pattern following FdeC protease shaving analysis, thereby confirming that both intimin and FdeC require the TAM for proper biogenesis.

5.3 TolC

TolC is the OMP component for a range of T1SSs where it plays a role in exporting toxins, such as α-haemolysin and colicin V, through HlyAB-TolC orCvaBA-TolC apparatus, respectively (Koronakis *et al.*, 2004). Additionally, TolC is the OMP component in a range of multi-drug efflux pumps, including AcrAB-TolC, AcrEF-TolC and EmrAB-TolC efflux pumps (Koronakis *et al.*, 2004). TolC adopts an unusual and complicated quaternary structure, whereby three TolC monomers each contribute four β-strands and four α-helices that assemble into a 12-stranded β-barrel with a 12-helical periplasmic tunnel important for the exportation of toxins and antimicrobials (Figure 5.1.1b).

The assembly of the TolC trimer is reported to be BAM-dependent, where depletion of either BamD (Malinverni *et al.*, 2006, BamD is YfiO therein) or BamA (Bennion *et al.*, 2010; Werner & Misra, 2005, BamA is YaeT therein) impacts on the steady-state level of TolC trimers. Given this, it is curious that Charlson *et al.* (2006) showed TolC levels almost doubled in response to deletion of *bamB* (*yfgL* therein), and anecdotally report that TolC levels remain unaffected by deletion of *bamC* (*nlpB* therein). The levels of BamB (but not BamC) are also important for TAM biogenesis (Section 4.6), and analysis of the data reported by Nichols *et al.* (2011) indicated that *E. coli* Δ*tamA* or Δ*tamB* strains have significantly increased

resistance to bleomycin (Appendix 2), which is a drug exported via AcrAB-TolC efflux pump (Cha *et al.*, 2014). Taken together, these reports highlight a possible interaction between TolC and the TAM. It was therefore important to determine whether TolC levels will similarly increase in the absence of the TAM.

Because TolC harbours a homotrimeric quaternary structure, the assembly of TolC from monomers to trimers can be monitored following pulse chase analysis by SN-PAGE. Wildtype assembly of TolC was first optimised by modifying the chase temperatures (Figure 5.3.1). While few (if any) bands were observed at 4 °C, TolC trimerisation was detected partially at 16 °C and prominently at 25 °C. Subsequent pulse chase analysis using the 25 °C chase temperature revealed a similar trimerisation phenotype for wildtype and each deletion mutant tested (Figure 5.3.2).

To confirm whether TolC levels were affected, the observed rate constant of the monomer band was determined for each TolC replicate by M. J. Belousoff and reported by Stubenrauch *et al.* (2016). The observed rate constant of the monomer band corresponds to how rapidly it assembles into trimeric TolC. In line with Charlson *et al.* (2006), TolC monomers assembled into trimers at wildtype levels in the absence of BamC, but the rate almost doubles if *bamB* is deleted instead. Deletion of *bamE* does not appear to affect trimerisation of TolC, which was also observed for the TolC homologue in Δ*bamE* mutants of *Caulobacter crescentus* (Toporowski *et al.*, 2004), and the absence of *tamA* or *tamB* similarly does not appear to affect TolC trimerisation.

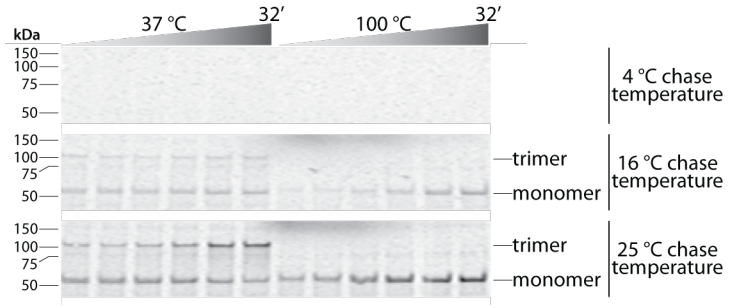


Figure 5.3.1 Chase temperature optimisation for pulse chase analysis of TolC assembly.

TolC assembly was assessed by pulse chase analysis in *E. coli* Star™ (DE3) wildtype cells harbouring pMB11 (pET-15b with *tolC* encoded), using the chase temperatures indicated to the right of the autoradiogram. At 10 seconds, 2, 4, 8, 16 and 32 minutes, duplicate aliquots were transferred to SN sample buffer and incubated for 10 minutes at either 37 °C or 100 °C as indicated above the autoradiogram. Analysis was by SN-PAGE and storage phosphorimaging. Sizes in kDa are indicated on the left. TolC monomers and trimers are indicated on the right. The time increment is indicated above the autoradiograms as a graded triangle.

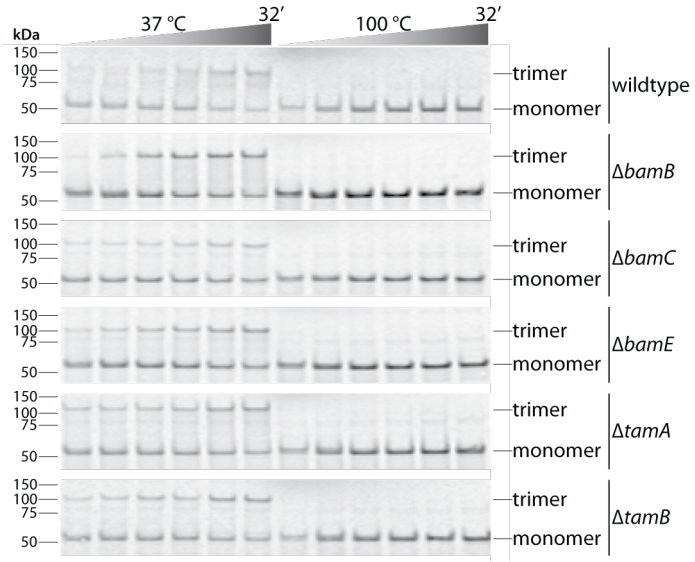


Figure 5.3.2 Pulse chase analysis of TolC assembly.

TolC assembly was assessed by pulse chase analysis in the indicated *E. coli* Star™ (DE3) strains harbouring pMB11 (pET-15b with *tolC* encoded) using a 25 °C chase temperature. At 10 seconds, 2, 4, 8, 16 and 32 minutes, duplicate aliquots were transferred to SN sample buffer and incubated for 10 minutes at either 37 °C or 100 °C as indicated above the autoradiogram. Analysis was by SN-PAGE and storage phosphorimaging. Sizes in kDa are indicated on the left. TolC monomers and trimers. The time increment is indicated above the autoradiograms as a graded triangle and the strain is indicated to the right of the autoradiograms.

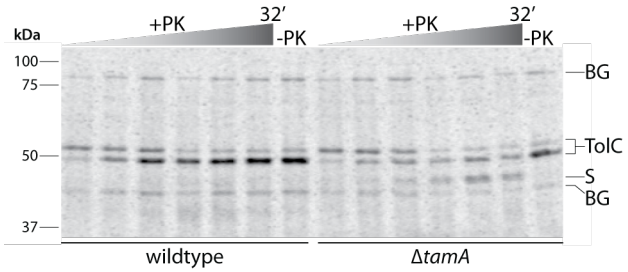


Figure 5.3.3 Efficient TolC biogenesis may require the TAM.

TolC biogenesis was assessed by pulse chase analysis in the indicated *E. coli* Star™ (DE3) strains harbouring pMB11 (pET-15b with *tolC* encoded). At 10 seconds, 2, 4, 8, 16 and 32 minutes, aliquots were taken and treated with (+PK) or without (-PK, last timepoint only) 50 μg mL^{-1} proteinase K (static, 10 min, on ice). Proteins were then TCA-precipitated, washed with acetone and boiled for 3 minutes in SDS sample buffer. Analysis was by SDS-PAGE and storage phosphorimaging. Sizes in kDa are indicated on the left. The monomeric TolC doublet, its fragment S, and the prominent background (BG) bands are indicated on the right. The time increment is indicated above the autoradiograms as a graded triangle and the strain is indicated below the autoradiograms.

Although the TAM may not influence the rate at which TolC trimerises, it may be possible the TAM plays a role in assembling a “functional” trimer, rather than the folding process itself. It has previously been shown that TolC is insensitive to extracellular protease (Werner *et al.*, 2003), but a defective trimer may harbour protease-sensitive loops that can be assayed via protease shaving. Preliminary protease shaving and pulse chase analysis of TolC (Figure 5.3.3) demonstrated that the pre-TolC precursor (top band of the "TolC" doublet) and its processed mature form (bottom band of the "TolC" doublet) are indeed protease-resistant when assembled in the presence of the BAM complex and the TAM. However, during assembly in the absence of TamA, TolC was degraded (into fragment S) following protease shaving analysis. This data suggested the TAM plays a role in the final assembly of active TolC trimers, a role that will be worth investigating further.

5.4 Trimeric porins

The trimeric porins are involved in passive diffusion of small solutes across the OM (Nikaido, 2003). When crystallised, the purified porin adopts a quaternary structure comprising three β-barrel monomers arranged in a triangular pattern (Figure 5.1.1c), but according to Rassam *et al.* (2015), a higher order of amorphous oligomers are observed in the native OM of *E. coli*. While a number of studies have discussed the role of the BAM complex in assembling a porin β-barrel (Selkrig *et al.*, 2014), it is unclear whether the BAM complex can also contribute to the assembly of the trimeric or higher order oligomeric structures. Several groups have observed that deep-rough LPS mutants are severely depleted of these porins, whereby sequentially increasing the number of sugar moieties in rough LPS

mutants (i.e. increasing the outer core oligosaccharide length) resulted in a concomitant increase in porin biogenesis (Hagge *et al.*, 2002; Laird *et al.*, 1994; Ried *et al.*, 1990; Sen & Nikaido, 1991). However, whether an additional protein factor, such as the TAM plays a role in the assembly of trimeric or higher order oligomeric porins is unknown.

Using PhoE (phosphorin) as a model trimeric porin (Figure 5.1.1c), its assembly may be monitored by pulse chase analysis through SN-PAGE. Wildtype assembly of PhoE species was first optimised by modifying the chase temperature during pulse chase analysis (Figure 5.4.1). Although distinct PhoE dimers and trimers were not observed at any timepoint for the chase temperatures tested, a smear corresponding to an oligomeric species was observed (Figure 5.4.1, "oligomers"). This oligomeric ladder has been observed for another trimeric porin, LamB (Misra *et al.*, 1991; Wu *et al.*, 2005), where they were speculated to: (i) be caused by the differential binding of the porin to LPS and/or (ii) correspond to various stable and meta-stable oligomeric intermediates. The oligomeric ladder was not observed during 16 °C chase; however, pulse chase analysis at 25 °C, 30 °C and 37 °C revealed no obvious difference in the intensity of monomeric and oligomeric bands.

Subsequent pulse chase analysis using the 37 °C chase temperature revealed that PhoE oligomers formed in wildtype *E. coli* and each non-essential BAM lipoprotein and TAM mutant (Figure 5.4.2). The observed rate constants of monomeric PhoE as it assembled into an oligomeric species in each of the six strains was determined by M. J. Belousoff as reported by Stubenrauch *et al.* (2016), and was

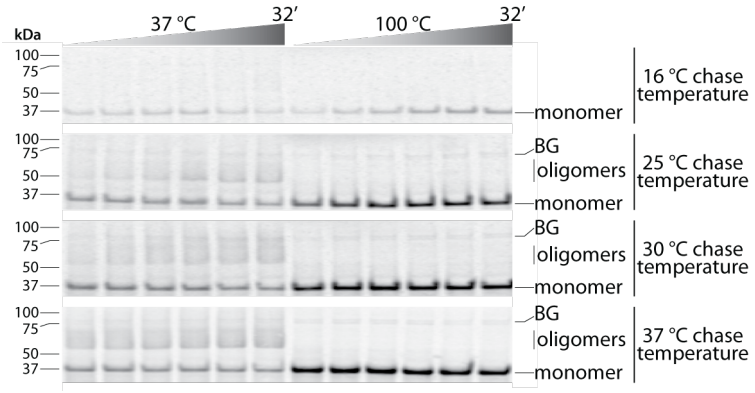


Figure 5.4.1 Chase temperature optimisation for pulse chase analysis of PhoE assembly.

PhoE assembly was assessed by pulse chase analysis in *E. coli* Star™ (DE3) wildtype cells harbouring pKS07 (pET-15b with *phoE* encoded), using the chase temperatures indicated to the right of the autoradiogram. At 10 seconds, 2, 4, 8, 16 and 32 minutes, duplicate aliquots were transferred to SN sample buffer and incubated for 10 minutes at either 37 °C or 100 °C as indicated above the autoradiogram. Analysis was by SN-PAGE and storage phosphorimaging. Sizes in kDa are indicated on the left. PhoE monomers, the ladder corresponding to PhoE oligomers, and the prominent background (BG) bands are indicated on the right. The time increment is indicated above the autoradiograms as a graded triangle.

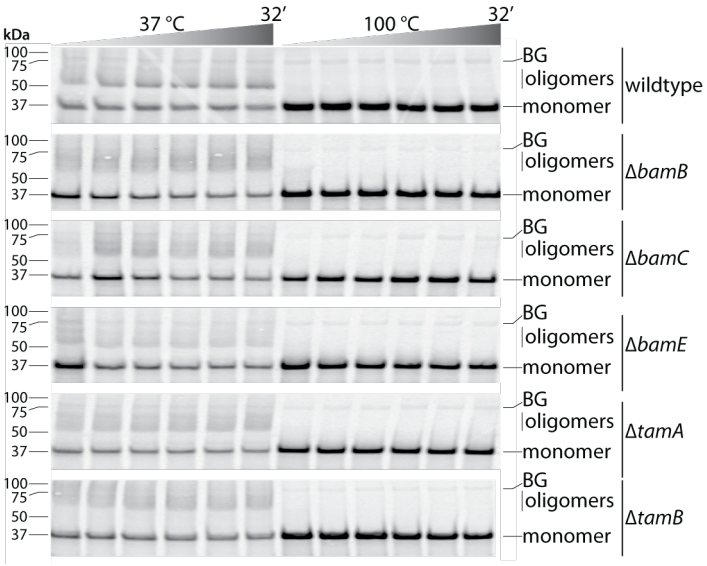


Figure 5.4.2 Pulse chase analysis of PhoE assembly.

PhoE assembly was assessed by pulse chase analysis in the indicated *E. coli* Star™ (DE3) strains harbouring pKS07 (pET-15b with *phoE* encoded) using a 37 °C chase temperature. At 10 seconds, 2, 4, 8, 16 and 32 minutes, duplicate aliquots were transferred to SN sample buffer and incubated for 10 minutes at either 37 °C or 100 °C as indicated above the autoradiogram. Analysis was by SN-PAGE and storage phosphorimaging. Sizes in kDa are indicated on the left. PhoE monomers, the ladder corresponding to the PhoE oligomers, and the prominent background (BG) bands are indicated on the right. The time increment is indicated above the autoradiograms as a graded triangle and the strain is indicated to the right of the autoradiograms.

found to be similar regardless of the strain background. Although this could indicate that neither the TAM nor the non-essential BAM lipoproteins are required for efficient PhoE assembly, because the identity of the oligomeric species is unknown, it is difficult to rule out the possibility that alternate oligomeric intermediates were formed during pulse chase analysis of the various strains.

5.5 LptDE

One of the final steps in LPS biogenesis is its transfer into the OM outer leaflet by LptDE. LptDE has a novel quaternary structure comprised of a 26-stranded β-barrel, LptD, with the lipoprotein LptE docked into the lumen of the LptD β-barrel (Dong *et al.*, 2014; Qiao *et al.*, 2014) (Figure 5.1.1d). It is clear that LptE is targeted to the OM via the Lol pathway (Okuda & Tokuda, 2011), and that LptD biogenesis proceeds via the BAM pathway (Workman *et al.*, 2012; Wu *et al.*, 2006). However, it is not clear how the final quaternary structure is assembled. Do the components meet by chance in the OM? Or are the Omp85 proteins also involved in establishing the quaternary structure of LptDE?

Babu *et al.* (2011) reported that TamB forms an "aggravating" interaction with LptD, where *E. coli* containing an *lptD* hypomorphic allele in a Δ*tamB* (Δ*ytfN* therein) background had a significant OM permeability and growth defect. Considering the defects were not observed in the presence of TamB, is it possible the TAM is required (but dispensable) for LptDE assembly? Unfortunately, Babu *et al.* (2011) did not report the primary structure of the *lptD* hypomorph (except to say that it contains a fusion tag), so experiments along those lines could not be reproduced. Instead, LptDE assembly was directly assessed by pulse chase

analysis in *tam* mutants (Figure 5.5.1).

By monitoring the assembly of LptDE through SN-PAGE analysis, it was found that LptDE displays a "heat-modifiability" phenotype (Figure 5.5.1a). Heat-modifiability is a structural characteristic of a number of folded β-barrel proteins, whereby the hydrogen bonding network within the β-barrel tertiary structure allows the protein to resist SDS-denaturation, provided that the sample is not subjected to heating (Nakamura & Mizushima, 1976). When analysed by SN-PAGE, the compact tertiary structure of the unheated protein causes it to migrate faster through the gel matrix than its heat-denatured counterpart.

Native LptDE was found to migrate as if it were a ~65 kDa band in both wildtype and Δ*tamA* extracts. By comparison, the heat-denatured sample showed an ~90 kDa LptD band and an ~20 kDa LptE band (Figure 5.5.1a). The absence of this native LptDE complex from cells overexpressing *lptD* or *lptE* alone indicated that the residual levels of the second component were not sufficient to reconstitute the quaternary structure for SN-PAGE analysis (Figure 5.5.1b). Of note, this also revealed that LptD alone does not display the same "heat-modifiability" phenotype as an LptDE complex, demonstrating the well-established paradigm that LptD requires LptE for proper biogenesis of a stable β-barrel (Chimalakonda *et al.*, 2011; Freinkman *et al.*, 2011; Wu *et al.*, 2006).

During SDS-PAGE analysis of LptDE assembly, LptDE is denatured into its constituent components, which were observed as a pair of LptE and LptD doublet bands (Figure 5.5.2a). The 90 kDa doublet corresponded to the pre-LptD

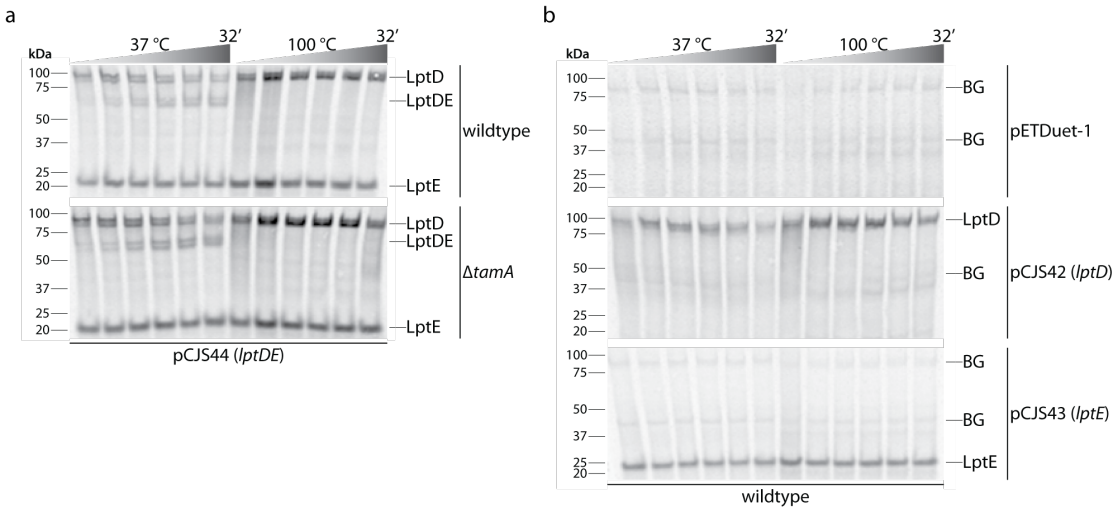


Figure 5.5.1 Pulse chase analysis of LptDE assembly.

LptDE assembly was monitored by pulse chase analysis in the indicated *E. coli* Star™ (DE3) strains harbouring either **a,** pCJS44 (pETDuet-1 with *lptD* and *lptE* encoded in separate MCSs), or; **b,** one of the following control vectors: pETDuet-1, pCJS42 (pETDuet-1 with *lptD* encoded) or pCJS43 (pETDuet-1 with *lptE* encoded). At 10 seconds, 2, 4, 8, 16 and 32 minutes, duplicate aliquots were transferred to SN sample buffer and incubated for 10 minutes at either 37 °C or 100 °C as indicated above the autoradiogram. Analysis was by SN-PAGE and storage phosphorimaging. Sizes in kDa are indicated on the left. LptD, LptE, the LptDE complex, and prominent background (BG) bands are indicated on the right. The time increment is indicated above the autoradiograms as a graded triangle and the strain and plasmid details are indicated below and to the right of the autoradiograms.

precursor (top band of the "LptD" doublet) and its processed mature form (bottom band of the "LptD" doublet). In wildtype cells, a ladder of proteolytic fragments was observed during LptDE assembly (Figure 5.5.2a, top panel), which corresponded to the same proteolytic ladder observed during LptD biogenesis alone (Figure 5.5.2b, top panel). Rather than the native LptDE complex, which is relatively protease-resistant, the proportion of LptD that had not yet established its LptE-dependent quaternary structure (i.e., the LptD band observed during SN-PAGE analysis of LptDE assembly at the permissive 37 °C: Figure 5.5.1a) was responsible for this proteolytic ladder.

In contrast however, when TamA was absent, LptDE was much more sensitive to exogenously added proteinase K, as indicated by: (i) the presence of the ~70 kDa fragment U, and; (ii) the decrease in LptD and LptE levels over time correlating with the concomitant increase in fragment U (Figure 5.5.2a, middle panel). The absence of fragment U following protease-shaving analysis of cells overexpressing *lptD* or *lptE* alone demonstrated that it was the protease-sensitivity of an LptDE complex (rather than that of its constituents) that resulted in fragment U formation when TamA was absent (Figure 5.5.2b).

LptD was found to be significantly more protease-sensitive during LptD biogenesis when the TAM was absent (Figure 5.5.2b, top panel). Additionally, when TamA was reintroduced to complement the LptDE protease-sensitivity wildtype protease-resistance levels of LptDE (Figure 5.5.2a, bottom panel). Taken together, this data corroborates and extends the data reported by Babu *et al.* (2011), by indicating that the TAM may be required for stabilising the LptD

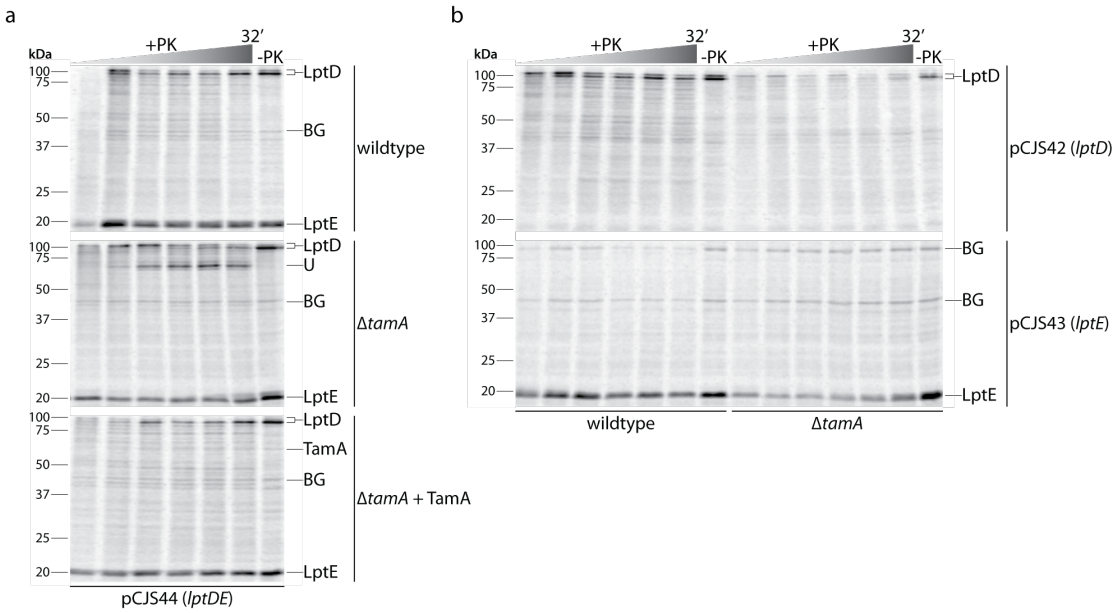


Figure 5.5.2 The TAM is involved in an LptDE assembly pathway.

LptDE assembly was monitored by pulse chase analysis in the indicated *E. coli* Star™ (DE3) strains harbouring either **a,** pCJS44 (pETDuet-1 with *lptD* and *lptE* encoded in separate MCSs) in conjunction with the control pACYCDuet-1 base vector or the pCJS69 (pACYCDuet-1 with *tamA* encoded) complementation vector (indicated as "+ TamA"), or; **b,** pCJS42 (pETDuet-1 with *lptD* encoded) or pCJS43 (pETDuet-1 with *lptE* encoded). At 10 seconds, 2, 4, 8, 16 and 32 minutes, aliquots were taken and treated with (+PK) or without (-PK, last timepoint only) 50 μg mL^{-1} proteinase K (static, 10 min, on ice). Proteins were then TCA-precipitated, washed with acetone and boiled for 3 minutes in SDS sample buffer. Analysis was by SDS-PAGE - using either 10 % (**a**) or 12 % (**b**) polyacrylamide gels - and by storage phosphorimaging. Sizes in kDa are indicated on the left. TamA, LptE, the LptD doublet, fragment U, and the prominent background (BG) bands are indicated on the right. The time increment is indicated above the autoradiograms as a graded triangle and the strain and plasmid details are indicated below and to the right of the autoradiograms.

tertiary structure, thereby facilitating the formation of the native, protease-resistant LptDE complex.

5.7 Discussion

The BAM complex is thought to insert β-barrel proteins within non-polar regions - predominantly mid-cell - of a bacterial cell, where newly assembled β-barrel proteins "push" pre-existing OMPs toward the cell poles (Rassam *et al.*, 2015; Ursell *et al.*, 2012). Interestingly, Rassam *et al.* (2015) observed a dynamic interplay between various OMPs - including BamA - that allowed their co-migration as distinct OMP islands toward cell poles. The promiscuous protein-protein interactions within each OMP island, mediated by various hydrophobic and aromatic residues, prevented the free diffusion of individual OMPs away from the cell poles, resulting in a binary distribution of "old" OMPs within daughter cells following septation. For both BtuB and CirA, β-barrel assembly cannot occur at the cell poles (Rassam *et al.*, 2015), indicating that the polar OMP islands may somehow inactivate the BAM complex.

As more proteins are "pushed" to the cell poles, the BAM complex may be forced to interact with neighbouring β-barrels such that they obstruct periplasmic chaperone interactions and/or "restore" the distorted protein-lipid interface that is necessary for β-barrel insertion. Considering that some autotransporters are targeted directly to the cell poles, Kleanthous *et al.* (2015) therefore hypothesised that: whereas the BAM complex assembles OMPs at non-polar regions within the cell, the TAM may assemble OMPs at the cell poles. However, this hypothesis does not align well with the observations that chaperone-usher fimbriae are

peritrichously distributed across the cell surface (Thanassi *et al.*, 2012).

Following the initial discovery of the TAM, Selkrig *et al.* (2012) identified three AIDA-like autotransporters that required the TAM for assembly: Ag43 and EhaA from *E. coli* and P1121 from *C. rodentium*. Based on the complex architecture of these AIDA-like autotransporters, an alternative hypothesis was developed whereby the TAM may be responsible for assembling particularly difficult OMPs that diverge from the classic, and relatively simple, β-barrel structure. In support of this hypothesis, pulse chase analysis of seven fimbrial ushers - each containing four periplasmic globular domains, in addition to a large transmembrane domain - revealed that they were dependent on the TAM for proper and efficient biogenesis (Chapter 4). Using FimD as the model fimbrial usher, I. D. Hay - in a parallel study reported by Stubenrauch *et al.* (2016) - subsequently showed that FimD requires the TAM to ensure type 1 fimbrial biogenesis occurs in a physiologically relevant timeframe.

In contrast to this hypothesis however, two SPATEs - which are structurally related to the AIDA-like autotransporters - apparently do not require the TAM for biogenesis. Sauri *et al.* (2009) showed that the steady-state level of Hbp is unchanged in the absence of TamA, based on western blotting analysis of mature Hbp. Additionally, Kang'ethe and Bernstein (2013) reported, albeit anecdotally, that EspP does not require the TAM for assembly. However, no functional assays were performed, and in light of the recent finding that the TAM is important for assembling FimD to ensure that type I fimbrial biogenesis occurs in a physiologically relevant timeframe, the conclusion that SPATEs do not require the

TAM should be re-examined.

To extend the hypothesis that OMPs with complicated topologies require the TAM for biogenesis, two inverse autotransporters - intimin and FdeC - were assessed for their ability to assemble in the presence or absence of the TAM and were both found to be TAM dependent. Intimin is normally protease resistant, but in the absence of the TAM, intimin biogenesis produces a protease-sensitive assembly intermediate that is readily proteolytically degraded. In contrast, the surface-exposed domains of FdeC are normally protease-sensitive and are therefore hydrolysed into a series of proteolytic fragments following the addition of exogenous proteinase K. In the presence of the TAM, two prominent proteolytic fragments are generated: ~100 kDa fragment Q corresponding to the passenger domain, and ~50 kDa fragment R corresponding to the transmembrane domain. When the TAM is absent however, fragment R is further degraded indicating that the transmembrane domain adopted by the FdeC assembly intermediate is more protease-sensitive. This is consistent with the fimbrial usher biogenesis mechanism, where in the absence of the TAM, β-barrel assembly is initiated from a central β-strand (rather than the C-terminus), resulting in a more protease-sensitive transmembrane domain.

The mouse pathogen, *C. rodentium*, is commonly used to model the equivalent intimin-dependent human diarrhoeal disease caused by pathogenic *E. coli* strains (Bhinder *et al.*, 2013), where it is known that intimin from enteropathogenic *E. coli* can complement a *C. rodentium* intimin-null mutant (Frankel *et al.*, 1996). Considering that *C. rodentium* TAM mutants were significantly attenuated in their

ability to infect mice (Kelly *et al.*, 2006; Selkrig *et al.*, 2012), it is tempting to speculate that in the absence of the TAM, intimin was not efficiently assembled, which therefore enabled wildtype *C. rodentium* to outcompete the isogenic TAM mutant strains. In *S.* Typhimurium, TAM mutants were found to be significantly more sensitive to complement-mediated killing by human serum (Selkrig *et al.*, 2012). The molecular explanation for this sensitivity is not yet clear, but *S.* Typhimurium are capable of producing a range of AIDA-like autotransporters, chaperone-usher fimbriae, and inverse autotransporters that likely generally contribute to disease severity and host-immune avoidance (Celik *et al.*, 2012; Heinz *et al.*, 2016; Humphries *et al.*, 2001), where at least some of these might require the action of the TAM in the course of their biogenesis.

In contrast, *A. fischeri* do not encode chaperone-usher fimbriae, AIDA-like or inverse autotransporters (Bongrand *et al.*, 2016; Ruby *et al.*, 2005), but *A. fischeri* Δ*tamB* mutants are significantly attenuated for their ability to colonise their symbiotic host (Brooks *et al.*, 2014). This suggests that in addition to the three known TAM substrate families, there are still more proteins requiring the TAM for efficient assembly. In *A. fischeri*, symbiotic colonisation is mediated by a number of factors, including numerous type 4 pili and an outer membrane protein called OmpU (Aeckersberg *et al.*, 2001; Bongrand *et al.*, 2016; Stabb & Ruby, 2003), and whether either (or both) of these colonisation factors are TAM substrates would be worth further investigation.

There are subtle structural differences within the Omp85 insertase family that are consistent with the hypothesis that the TAM assembles OMPs with complicated

topologies. The BamA exit pore that is occluded by loop 6 is also further stabilised by loops 4 and 7 (Noinaj *et al.*, 2015), so considering that the exit pore is thought to partially open during substrate biogenesis to facilitate passage of extracellular loops and/or domains (Gruss *et al.*, 2013; Noinaj *et al.,* 2015), the TamA exit pore that is only occluded by loop 6 may more readily adopt an "open" conformation than that of BamA. In addition to the lopsided aromatic girdle and lateral gate characteristic of Omp85 insertases (Noinaj *et al.*, 2014; Noinaj *et al.*, 2013), TamA contains rigid POTRA domains that allow the TAM to function by further destabilising the OM lipid-protein interface (Selkrig *et al.*, 2015; Shen *et al.*, 2014), which would theoretically lower the activation energy required for the insertion of complex OMPs. In light of these structural and functional advantages displayed by TamA however, it is not immediately clear why the BAM complex is the preferred (and only essential) β-barrel insertase.

The observation by Rassam *et al.* (2015) that the BAM complex is relatively mobile within the OM and capable of clustering promiscuously within OMP islands could provide a clue as to why the BAM complex is the preferred β-barrel insertase. Kleanthous *et al.* (2015) hypothesised that the dynamic interplay between OMPs within the OMP islands is especially important for the formation of β-barrels with quaternary structure, but how this would occur mechanistically was not discussed. Is this hypothesis correct in excluding the TAM from similarly migrating within OMP islands to facilitate quaternary complex assembly though? Considering the TAM forms an intermembrane-spanning complex, TamB must permeate the peptidoglycan layer. Regardless of whether TamB is capable of penetrating the peptidoglycan layer with or without the aid of the enzymes

responsible for degrading and remodelling peptidoglycan, it is unlikely that the TAM is capable of similarly migrating as a quaternary structure like the BAM complex. However, this does not preclude the formation of a TAM following independent diffusion events of TamA and TamB, especially considering that each of the five BAM components must assemble together in a similar fashion.

To determine whether the BAM complex alone, or in conjunction with the TAM, assembles β-barrel proteins that contain quaternary structure, pulse chase analysis of TolC, PhoE and LptDE assembly was performed. While the results for PhoE were inconclusive, the TAM was found to be required for LptDE assembly, and preliminary data indicated that TolC may similarly interact with the TAM. LptD alone is partially protease-sensitive, but in conjunction with LptE, it assembles into the protease-resistant LptDE. In a Δ*tamA* background however, protease shaving and pulse chase analysis indicated that LptD alone and in complex with LptE were significantly more protease-sensitive. Indeed, while 90 kDa LptD outside of the LptDE complex was proteolytically degraded almost completely, a ~70 kDa LptD fragment U was generated from the LptDE assembly intermediate. However, considering that LptDE is essential, the TAM likely participates as an insertase that is subsidiary to the BAM complex.

Previous work by Babu *et al.* (2011) demonstrated that *E. coli* had a significant OM permeability and growth defect when either Δ*tamB*, Δ*yfgH* or Δ*yceK* mutants additionally contained an *lptD* hypomorphic allele. While YfgH and YceK are both known lipoproteins that respectively map to the OM and IM (Juncker *et al.*, 2003), they remain functionally uncharacterised. Therefore, it is tempting to speculate

that - due to the similar phenotypes reported by Babu *et al.* (2011) - YfgH and/or YceK may (i) be novel interacting partners of the TAM, or indeed novel TAM components, or; (ii) represent components of additional assembly pathway mechanisms that also facilitate (at least) LptDE assembly.

Until now, it has been implied that the BAM complex and the TAM work independently toward the same goal, but this is not necessarily the case. Although no TAM-BAM native complexes have been reported in the literature thus far (see also Figure 4.4.5), this does not mean that the BAM complex and the TAM interact co-dependently within the same pathway. If the assembly mechanism of these Omp85 insertases indeed requires distortion of the protein-lipid interface (Noinaj *et al.*, 2015; Selkrig *et al.*, 2015; Shen *et al.*, 2014), it is possible that nearby BAM complexes and TAMs may generally facilitate protein-lipid interface distortion. One of the most important questions is therefore: what distinguishes TAM substrates from those of the BAM complex? Does the TAM have a greater affinity for substrates with complex domains? Or do periplasmic chaperones selectively deliver their cargo? To date, despite the number of known OMP periplasmic chaperones (Rollauer *et al.*, 2015), the only known interaction between periplasmic chaperones and Omp85 insertases is between SurA and BamA (Sklar *et al.*, 2007).

Regardless of whether the two β-barrel assembly pathways physically interact, it is clear that during fimbrial usher biogenesis the BAM complex is subsidiary to the TAM, and during LptDE biogenesis the TAM is subsidiary to the BAM complex. Although there is a general consensus in the field that the BAM complex assembles

all β-barrel proteins (Fleming, 2015; Rollauer *et al.*, 2015), this work expands that hypothesis by demonstrating that the TAM also plays various major and minor roles in β-barrel biogenesis.

Appendices

Appendix 1 - List of plasmid mutations located within the gene of interest

All vectors used in this study are as per Table 2.1.2. The gene of interest in each vector may contain a number of intentional mutations to ensure an optimal start or stop codon, or incorporation of particular cleavage site or fusion tag. Alternatively, silent mutations are listed that were introduced (presumably) during PCR amplification of the gene of interest.

			Mutations				
Plasmid	**Gene**	**Locus tag identifier[1]**	**Start codon[2]**	**in signal peptide due to *NcoI*[3]**	**Stop codon[4]**	**Other intentional mutations[5]**	**PCR[6]**
pKS02	*fimD*	b4317		c.3_4insGGC (p.M1_S2insG)			
pKS06	*tamA*	b4220	c.1G>A	c.3_4insGGC (p.M1_R2insG			
pKS07	*phoE*	b0241		c.3_4insGGC (p.M1_K2insG)			
pMB11	*tolC*	b3035		c.3_4insGGC (p.M1_K2insG)			
pWAK05	*fimD*	b4317				c.397_398delinsTG (p.D133C); c.1489_1508delinsATGA ATGGCTACAACATCGA (p.496R_504TdelinsENLYFQG); c.2401_2403delinsACG (p.T801C); c.2634_2635ins30 (p.R878_*879insASWSHPQFEK)	
pCJS29	*fimD*	b4317		c.3_4insGGC (p.M1_S2insG)		c.397_398delinsTG (p.D133C); c.1489_1508delins20 (p.496R_504TdelinsENLYFQG); c.2401_2403delinsACG (p.T801C); c.2634_2635ins30 (p.R878_*879insASWSHP QFEK)	
pCJS30	*eaeA*	E2348C_3939		c.4A>G (p.I2V)			
pCJS31	*yqiG*	c3792		c.4A>G (p.N2D)			

pCJS32	*papC*	c3590/c5186					c.165T>C (still G); c.291A>G (still P); c.300T>C (still C)
pCJS33	*yfcU*	c2883		c.4T>G (p.C2G)			
pCJS37	*ecpC*	E2348C_0248					
pCJS40	*ybgQ*	b0718	c.1G>A	c.4A>G (p.N2D)			
pCJS42	*lptD*	b0054		c.3_4insGGC (p.M1_K2insG)	c.2354G>A		
pCJS43	*lptE*	b0641	c.1G>A		c.581G>A		
pCJS44 (MCS1)	*lptD*	b0054		c.3_4insGGC (p.M1_K2insG)	c.2354G>A		
pCJS44 (MCS2)	*lptE*	b0641	c.1G>A		c.581G>A		
pCJS49	*htrE*	c0170	c.1G>A	c.4A>G (p.T2A)			
pCJS50	*fdeC*	c0415		c.4A>G (p.K2E)			
pCJS51	*fimD*	b4317				c.1_135delins99 (p.1_45delinsMKYLLPTAAAGLLLLAAQ PAMAMASWSHPQFEK)	
pCJS52	*fimD*	b4317		c.3_4insGGC (p.M1_S2insG)			
pCJS58 (MCS1)	*fimD*	b4317		c.3_4insGGC (p.M1_S2insG)			
pCJS58 (MCS2)	*tamA*	b4220	c.1G>A	c.3_4insGGC (p.M1_R2insG			
pCJS60 (MCS1)	*fimD*	b4317		c.3_4insGGC (p.M1_S2insG)			
pCJS60 (MCS2)	*bam B*	E2348C_2795					
pCJS64	*fimD*	b4317		c.3_4insGGC (p.M1_S2insG)		c.540_541insCATCATCACCATCACCACTCCGGA (p.P180_G181insHHHHHHSG)	
pCJS65	*fimD*	b4317		c.3_4insGGC (p.M1_S2insG)		c.768_769insCATCATCACCATCACCACTCCGGA (p.P256_L257insHHHHHHSG)	
pCJS66	*fimD*	b4317		c.3_4insGGC (p.M1_S2insG)		c.1824_1825insCATCATCACCATCACCACTCCGGA (p.S608_D609insHHHHHHSG)	

pCJS69	*tamA*	b4220	c.1G>A	c.3_4insGGC (p.M1_R2insG)			
pCJS72	*tamB*	b4221	c.3780G>A				
pCJS73	*bamB*	E2348C_2795					
pCJS74	*bamE*	b2617					

[1] Locus tag is used because GI numbers are being retired as of September 2016. [2] Mutations that alter the GTG to ATG start codon (silent mutation). [3] Mutations introduced from cloning into *Nco*I restriction site, which contains an out-of-frame start codon (usually resulting in the second amino acid being glycine). Because this mutation is within the cleavable signal peptide region, this mutation will not affect the mature protein's primary sequence. [4] Mutations that alter the stop codon used (silent mutation). [5] Other mutations that were introduced by primers, including fusion tags and protease cleavage sites. [6] Silent mutations that were likely introduced during PCR amplification

Appendix 2 - TAM mutant phenotypes indicated by colony size in the presence of various "stress" conditions

This data was originally reported by Nichols *et al.* (2011) as a normalised colony size distribution, where the authors defined significant mutant phenotypes as colony sizes located within the top fifth (more resistant) or bottom fifth percentile (more sensitive). Shown here are the calculated percentiles for the 324 "stress" conditions of both *E. coli* TAM mutants, according to their "ranking" within the 3979 *E. coli* strains originally tested. Significant phenotypes are highlighted in blue with white and bold typeface. The six antimicrobials where at least one concentration indicated a significant growth phenotype are listed at the top of the table and were reproduced for Table 3.1.1. The remaining 299 "stress" conditions that did not show significant growth phenotypes are listed below.

		Concentrations screened	Percentile	
Condition	**Target**		**Δ*tamA***	**Δ*tamB***
More resistant				
Actinomycin	RNA biosynthesis	2.5 µg/mL	79.8	78.5
		5.0 µg/mL	56.6	57.7
		10.0 µg/mL	94.4	67.6
		15.0 µg/mL	**95.3**	86.1
Bleomycin	DNA and RNA degradation	0.1 µg/mL	59.1	88.6
		0.5 µg/mL	**97.1**	90.2
		1.0 µg/mL	**99.0**	93.1
		2.0 µg/mL	**97.0**	**98.3**
Novobiocin	DNA gyrase	4 µg/mL	29.4	60.8
		6 µg/mL	37.0	73.0
		8 µg/mL	41.4	86.2
		10 µg/mL	66.4	76.8
		12 µg/mL	69.8	75.2
		30 µg/mL	**96.1**	53.6
More sensitive				
Cholate	Membrane	0.1 % (w/v)	68.9	66.5
		0.5 % (w/v)	**0.3**	44.7
		1.0 % (w/v)	14.7	47.4
		2.0 % (w/v)	35.4	55.0
Dibucaine	Proton motive force	0.4 mM	31.9	8.1
		0.8 mM	15.8	15.0
		1.2 mM	17.4	**2.4**

H_2O_2	Oxidative stress	0.1 mM	**1.3**	70.7
		0.5 mM	16.7	10.1
		1.0 mM	51.5	85.5
		2.0 mM	39.0	20.3
No growth phenotype				
A22	MreB inhibitor	0.5 µg/mL	68.7	48.2
		2.0 µg/mL	30.1	36.2
		5.0 µg/mL	34.5	42.0
		15.0 µg/mL	55.8	36.2
Acetate (in M9 media)	Alternative carbon source	0.6 % (w/v)	67.5	60.4
acidic pH (MES-HOMOPIPES)	-	pH 4.0 in 100 mM buffer	13.3	38.7
		pH 4.5 in 100 mM buffer	17.9	58.8
		pH 5.0 in 100 mM buffer	49.6	78.1
		pH 6.0 in 100 mM buffer	35.2	44.0
Acriflavine	Peptidoglycan dye	2.0 µg/mL	14.9	79.3
		10.0 µg/mL	46.4	20.8
Amikacin	protein synthesis: 30S ribosomal subunit	0.05 µg/mL	8.3	66.5
		0.10 µg/mL	26.6	65.0
		0.20 µg/mL	7.1	89.8
Amoxicillin	peptidoglycan biosynthesis: transpeptidases	0.25 µg/mL	90.2	87.3
		0.50 µg/mL	84.0	68.1
		1.00 µg/mL	16.9	39.0
		1.50 µg/mL	91.5	86.1
Ampicillin	peptidoglycan biosynthesis: transpeptidases	1.0 µg/mL	38.5	43.7
		2.0 µg/mL	22.1	34.1
		4.0 µg/mL	51.5	30.7
		8.0 µg/mL	23.5	46.1
Anaerobic	-	-	41.1	15.9
Azidothymidine	DNA damage (strand breakage)	0.5 ng/mL	69.8	58.9
		1.0 ng/mL	73.4	68.8
		2.5 ng/mL	77.1	92.2
Azithromycin	protein synthesis: 50S ribosomal subunit	0.02 µg/mL	42.8	57.0
		0.10 µg/mL	91.4	31.8
		1.00 µg/mL	82.6	67.6
Aztreonam	peptidoglycan biosynthesis: transpeptidases	20 ng/mL	30.8	56.4
		40 ng/mL	68.7	72.4
Bacitracin	peptidoglycan biosynthesis: bactoprenol	100 µg/mL	55.4	34.7
		200 µg/mL	52.2	69.2
		300 µg/mL	47.0	62.5
basic pH (TAPS)	-	pH 8.0 in 100 mM buffer	43.9	25.3
		pH 9.0 in 100 mM buffer	47.8	37.7
		pH 9.5 in 100 mM buffer	35.0	18.5
		pH 10.0 in 100 mM buffer	60.4	24.8
Benzalkonium	Membrane	1 µg/mL	32.5	42.8
		10 µg/mL	44.7	75.7
		25 µg/mL	17.0	49.9
Bicyclomycin		1 µg/mL	44.9	35.9

	Transcription termination factor, Rho	10 µg/mL	38.7	64.9
Bile salts	Membrane	0.1 % (w/v)	72.7	53.6
		0.5 % (w/v)	33.2	83.5
		1.0 % (w/v)	28.6	49.2
		2.0 % (w/v)	52.1	56.3
Calcofluor White M2R	Cellulose and chitin dye	50 µg/mL	76.5	73.3
Carbenicillin	peptidoglycan biosynthesis: transpeptidases	0.5 µg/mL	76.4	38.8
		1.0 µg/mL	55.5	24.6
		1.5 µg/mL	75.8	29.1
Carbonyl cyanide 3-chlorophenylhydrazone	disrupts proton motive force	0.1 µg/mL	64.1	75.8
		0.5 µg/mL	58.9	62.0
		2.0 µg/mL	57.6	50.7
Cecropin B	Membrane	0.1 µg/mL	71.9	73.8
		0.3 µg/mL	62.3	45.1
Cefaclor	peptidoglycan biosynthesis: transpeptidases	1 µg/mL	69.0	88.6
		2 µg/mL	83.5	85.2
		3 µg/mL	71.5	90.0
Cefoxitin	peptidoglycan biosynthesis: transpeptidases	0.25 µg/mL	16.5	26.5
		0.50 µg/mL	41.8	55.5
		0.75 µg/mL	5.8	55.9
		1.00 µg/mL	17.8	59.6
Cefsulodin	peptidoglycan biosynthesis: transpeptidases	6 µg/mL	47.1	39.0
		12 µg/mL	17.9	50.4
		18 µg/mL	63.7	37.8
		24 µg/mL	39.0	48.9
Cefsulodin (Cef) + Mecillinam (Mec)	as single compounds	6 µg/mL (Cef) + 0.03 µg/mL (Mec)	44.7	19.9
Ceftazidime	peptidoglycan biosynthesis: transpeptidases	25 ng/mL	33.7	59.5
		50 ng/mL	67.1	69.2
		75 ng/mL	74.0	55.8
Cerulenin	Fatty acid biosynthesis	1 µg/mL	15.4	51.8
		2 µg/mL	16.9	79.4
		4 µg/mL	8.0	57.3
		6 µg/mL	22.2	81.2
CHIR-090	LPS biosynthesis: inhibits LpxC	20 ng/mL	75.6	33.8
		25 ng/mL	80.7	60.8
		40 ng/mL	66.4	32.5
		50 ng/mL	79.9	79.6
		75 ng/mL	65.5	74.6
Chloramphenicol	protein synthesis: 50S ribosomal subunit	0.5 µg/mL	64.8	74.6
		1.0 µg/mL	78.9	73.3
		1.5 µg/mL	86.9	64.2
		2.0 µg/mL	51.9	55.8
Chlorpromazine	Membrane	3 µM	61.5	62.3
		6 µM	60.0	64.5
		12 µM	44.2	49.7
		24 µM	74.3	44.3
Ciprofloxacin	DNA gyrase	4 ng/mL	86.9	46.6
		6 ng/mL	81.1	64.6
		8 ng/mL	94.5	83.8
Cisplatin	DNA	20 µg/mL	11.5	14.7
		50 µg/mL	19.1	82.9
		100 µg/mL	44.0	62.2

Clarythromycin	protein synthesis: 50S ribosomal subunit	0.1 µg/mL	16.1	53.3
		1.0 µg/mL	30.9	81.6
		5 µg/mL	7.2	73.1
		10 µg/mL	33.8	74.1
Cobalt stress ($CoCl_2$) in LB	-	0.1 mM	53.1	76.5
		0.5 mM	73.6	49.7
Cold shock	-	16 °C	83.9	76.6
		18 °C	42.7	86.4
		20 °C	58.9	31.3
Copper stress ($CuCl_2$) in LB	-	1 mM	68.6	58.6
		2 mM	76.0	68.3
		4 mM	82.7	77.5
Cycloserine D	Peptidoglycan biosynthesis: racemase and ligase	16 µg/mL	42.7	22.3
Deoxycholate	Membrane	0.1 % (w/v)	62.6	58.7
		0.5 % (w/v)	39.6	26.6
		2.0 % (w/v)	17.4	68.9
Doxorubicin	ArcB (multidrug transporter)	1 µg/mL	12.9	68.3
		10 µg/mL	13.5	87.4
Doxycycline	protein synthesis: 16S ribosomal subunit	0.25 µg/mL	84.2	27.3
		0.50 µg/mL	88.0	45.9
		0.75 µg/mL	63.8	18.7
		1.00 µg/mL	40.3	21.1
EDTA	Outer membrane	0.1 mM	19.2	58.3
		0.5 mM	26.0	54.8
		1.0 mM	48.5	20.0
EGTA	Outer membrane	0.1 mM	53.8	76.7
		0.5 mM	79.1	50.5
		1.0 mM	74.7	55.7
		2.0 mM	68.4	73.5
Epigallocatechin gallate	Fatty acid biosynthesis	5 µM	68.2	67.0
		20 µM	75.5	65.4
		50 µM	50.9	43.5
Epinephrine	QseC (quorum sensor kinase)	50 µM	92.8	60.0
		250 µM	62.5	45.3
		1000 µM	68.3	29.8
Erythromycin	protein synthesis: 50S ribosomal subunit	0.1 µg/mL	37.5	52.8
		1.0 µg/mL	60.1	81.3
		5.0 µg/mL	69.4	75.6
		10.0 µg/mL	70.4	76.3
Ethidium Bromide	DNA	2 µg/mL	35.8	67.8
		10 µg/mL	61.5	44.4
		50 µg/mL	54.9	55.6
EtOH	protein folding	2 % (v/v)	41.8	39.4
		4 % (v/v)	67.2	23.1
		6 % (v/v)	59.4	37.1
Fosfomycin	Peptidoglycan biosynthesis: NAM synthesis	1 µg/mL	31.3	10.7
Fosfomycin + Glucose-6-Phosphate (G6P)	same as single compounds	0.05 µg/mL + 50 mg/mL (G6P)	41.1	27.4
		0.20 µg/mL + 50 mg/mL (G6P)	44.8	37.4
Fusidic acid	protein synthesis: Elongation factor G	1 µg/mL	10.6	62.6
		5 µg/mL	6.0	73.2
		20 µg/mL	11.2	94.3

		50 µg/mL	36.6	83.3
Gentamicin	protein synthesis: 30S ribosomal subunit	50 ng/mL	75.2	82.8
		100 ng/mL	85.2	90.7
Glucosamine (in M9 media)	alternative carbon source	0.2 % (w/v)	59.3	41.3
Glucose (in M9 media)	alternative carbon source	0.2 % (w/v)	59.2	87.8
Glycerol (in M9 media)	alternative carbon source	0.4 % (w/v)	83.3	61.3
Heat shock	-	40 °C	72.9	84.6
		42 °C	33.5	50.2
		43.5 °C	49.4	52.7
		45 °C	26.8	59.4
Hydroxyurea	DNA damage	1 mM	48.2	45.0
		5 mM	56.0	17.2
		10 mM	84.7	22.2
Indolicidin	Outer membrane and LPS	0.1 µg/mL	86.3	52.8
Iron excess ($FeSO_4$) in MOPS	-	1 mM (normal [$FeSO_4$] 100 µM)	60.8	44.3
Iron starvation ($FeSO_4$) in MOPS	-	2 µM (normal [$FeSO_4$] 100 µM)	66.6	42.5
Isoniazid	*Mycobacterium* antibiotic: mycolic acid biosynthesis	0.2 mM	63.1	20.7
		1.0 mM	20.3	79.9
		1.5 mM	30.4	26.5
Levofloxacin	DNA gyrase	2 ng/mL	31.4	17.2
Maltose (in M9 media)	alternative carbon source	0.1 % (w/v)	78.8	66.0
Mecillinam	peptidoglycan biosynthesis: transpeptidases	30 ng/mL	54.5	65.8
		60 ng/mL	39.6	65.2
		90 ng/mL	33.8	61.8
		120 ng/mL	60.7	44.5
Methotrexate	folic acid biosynthesis: dihydro folate reductase	1 µg/mL	26.2	55.1
		25 µg/mL	22.3	82.9
Minocycline	protein synthesis: 16S ribosomal subunit; also binds AcrB (multidrug transporter)	0.2 µg/mL	93.1	42.2
		0.5 µg/mL	49.2	32.9
		1.0 µg/mL	34.1	21.2
Mitomycin C	DNA and RNA synthesis	0.1 µg/mL	35.4	30.7
Methyl methanesulfonate	DNA damage (methylation)	0.05 % (w/v)	20.2	24.9
N-acetyl Glucosamine (in M9 media)	alternative carbon source	0.15 % (w/v)	52.4	49.7
NaCl	-	150 mM	35.0	25.1
		300 mM	63.4	28.6
		450 mM	77.5	67.6
		600 mM	68.9	66.9
Nalidixic acid	DNA gyrase	0.5 µg/mL	68.2	79.8
		1.0 µg/mL	50.0	85.6
		1.5 µg/mL	65.0	48.0
		2.0 µg/mL	42.9	58.6
NH_4Cl (in MOPS)	alternative nitrogen source	9.5 mM	36.8	53.4

Nickel stress ($NiCl_2$) in LB	-	0.1 mM	43.0	73.2
		1.0 mM	72.6	39.7
Nigericin	Proton motive force	0.1 µM	22.7	91.5
		1.0 µM	19.1	82.5
		5.0 µM	19.5	86.8
Nitrofurnatoin	DNA and Krebs cycle	0.1 µg/mL	53.5	93.4
		0.5 µg/mL	63.6	59.2
		1.0 µg/mL	85.1	53.9
		1.5 µg/mL	65.9	68.5
		2.0 µg/mL	73.4	86.9
Norepinephrine	QseC (quorum sensor kinase)	100 µM	31.2	37.3
		1000 µM	64.9	63.0
Norfloxacin	DNA gyrase	10 ng/mL	21.3	46.8
		20 ng/mL	72.4	71.6
		40 ng/mL	75.8	76.1
Oxacillin	peptidoglycan biosynthesis: transpeptidases	0.5 µg/mL	67.4	65.2
		5.0 µg/mL	45.0	77.6
		40.0 µg/mL	83.1	61.6
Paraquat dichloride	oxidative stress	0.2 µM	72.0	78.3
		1.0 µM	10.4	50.9
		5.0 µM	18.1	11.9
		10.0 µM	67.6	89.6
		18.0 µM	66.9	42.3
Phenazine methosulfate	Superoxide stress	20 nM	82.0	20.9
		50 nM	80.0	39.0
		100 nM	74.3	31.9
Phleomycin	DNA damage	0.2 µg/mL	18.6	31.7
		0.5 µg/mL	83.6	56.5
		1.0 µg/mL	84.9	57.5
Polymyxin B	LPS	1 µg/mL	9.8	18.2
		2 µg/mL	15.0	19.6
		4 µg/mL	54.4	11.9
		6 µg/mL	10.0	10.1
Procaine	Membrane: EnvZ/OmpR two-component system	1 mM	48.4	50.6
		5 mM	10.3	46.6
		10 mM	53.5	47.6
		30 mM	20.0	41.8
Propidium iodide	DNA/RNA	1 µg/mL	56.8	88.0
		20 µg/mL	90.6	84.7
		50 µg/mL	64.2	86.0
Puromycin	protein synthesis inhibitor	1 µg/mL	26.6	64.4
		5 µg/mL	57.5	37.2
		25 µg/mL	22.2	68.9
Pyocyanin	Superoxide stress	0.2 µg/mL	16.9	69.7
		1 µg/mL	33.9	48.2
		10 µg/mL	21.8	42.9
Radicicol	HtpG and PhoQ	1 µM	74.7	51.5
		5 µM	32.2	43.7
		10 µM	58.6	48.4
Rifampicin	RNA polymerase	1 µg/mL	12.7	44.1
		2 µg/mL	51.9	77.0
SDS	Membrane	0.5 % (w/v)	85.5	32.9
		1.0 % (w/v)	60.9	22.2
		2.0 % (w/v)	35.6	28.7
		3.0 % (w/v)	55.9	34.7
		4.0 % (w/v)	53.0	25.2

SDS+EDTA	Membrane	0.5 % (w/v) SDS + 0.1 mM EDTA	85.6	12.2
		0.5 % (w/v) SDS + 0.5 mM EDTA	92.6	21.0
		1.0% (w/v) SDS + 0.5 mM EDTA	52.7	9.5
Spectinomycin	protein synthesis: 30S ribosomal subunit	4 µg/mL	77.4	30.9
		6 µg/mL	87.4	66.9
Spiramycin	protein synthesis: 50S ribosomal subunit	1 µg/mL	22.3	73.8
		5 µg/mL	77.9	66.7
		20 µg/mL	65.1	79.9
Streptomycin	protein synthesis: 30S ribosomal subunit	0.05 µg/mL	77.8	43.2
Streptonigrin	DNA metabolism: respiration	0.1 µg/mL	81.4	75.6
		0.2 µg/mL	43.2	42.3
		0.5 µg/mL	72.1	67.9
Succinate (in M9 media)	alternative carbon source	0.3 % (w/v)	66.4	44.8
Sulfamethizole	folic acid biosynthesis	100 µg/mL	75.7	40.7
		200 µg/mL	34.8	56.5
		300 µg/mL	67.5	27.6
Sulfamonomethoxine	folic acid biosynthesis	50 µg/mL	48.5	66.5
		100 µg/mL	64.6	27.5
Taurocholate	Membrane	0.1 % (w/v)	18.0	67.4
		0.5 % (w/v)	5.0	65.7
		1.0 % (w/v)	6.0	51.1
Tetracycline	protein synthesis: 30S ribosomal subunit	0.25 µg/mL	85.8	51.8
		0.50 µg/mL	82.4	31.4
		0.75 µg/mL	66.3	61.3
		1.00 µg/mL	82.9	67.5
Theophylline	-	10 µg/mL	82.4	68.4
		100 µg/mL	83.6	79.7
Thiolactomycin	Fatty acid biosynthesis	1 µM	70.6	74.1
		5 µM	65.1	26.9
		50 µM	84.8	22.2
Tobramycin	protein synthesis: 30S ribosomal subunit	50 ng/mL	37.6	81.6
		100 ng/mL	39.5	61.4
		200 ng/mL	76.3	86.8
		400 ng/mL	75.5	76.5
Triclosan	Fatty acid biosynthesis	0.05 µg/mL	85.2	55.8
Trimethoprim	folic acid biosynthesis: dihydro folate reductase	100 ng/mL	68.2	66.6
		200 ng/mL	60.5	31.7
		300 ng/mL	26.6	51.6
		400 ng/mL	77.3	53.4
Trimethoprim (Tri) + Sulfamethizole (Sul)	same as single compounds	0.1 µg/mL (Tri) + 50 µg/mL (Sul)	26.2	31.5
Triton X-100	Membrane	0.01 % (v/v)	53.4	58.7
		0.03 % (v/v)	55.7	54.1
		0.20 % (v/v)	62.8	82.1
Tunicamycin	peptidogylcan biosynthesis: MraY	1.0 µg/mL	81.6	91.4
		3.0 µg/mL	56.6	64.5
		7.5 µg/mL	20.4	68.1
UV	DNA damage	6 seconds	88.0	63.9
		12 seconds	7.9	94.3
		18 seconds	53.2	94.2

		24 seconds	85.7	79.8
Vancomycin	peptidoglycan elongation	10 µg/mL	33.4	67.8
		20 µg/mL	31.6	87.3
		50 µg/mL	81.8	83.3
Verapamil	cell division and proton motive force	0.1 mM	57.3	61.9
		0.5 mM	34.0	16.3
		1.0 mM	20.2	39.3
		200 µg/mL	34.8	56.5
		300 µg/mL	67.5	27.6
Sulfamonomethoxine	folic acid biosynthesis	50 µg/mL	48.5	66.5
		100 µg/mL	64.6	27.5
Taurocholate	Membrane	0.1 % (w/v)	18.0	67.4
		0.5 % (w/v)	5.0	65.7
		1.0 % (w/v)	6.0	51.1
Tetracycline	protein synthesis: 30S ribosomal subunit	0.25 µg/mL	85.8	51.8
		0.50 µg/mL	82.4	31.4
		0.75 µg/mL	66.3	61.3
		1.00 µg/mL	82.9	67.5
Theophylline	-	10 µg/mL	82.4	68.4
		100 µg/mL	83.6	79.7
Thiolactomycin	Fatty acid biosynthesis	1 µM	70.6	74.1
		5 µM	65.1	26.9
		50 µM	84.8	22.2
Tobramycin	protein synthesis: 30S ribosomal subunit	50 ng/mL	37.6	81.6
		100 ng/mL	39.5	61.4
		200 ng/mL	76.3	86.8
		400 ng/mL	75.5	76.5
Triclosan	Fatty acid biosynthesis	0.05 µg/mL	85.2	55.8
Trimethoprim	folic acid biosynthesis: dihydro folate reductase	100 ng/mL	68.2	66.6
		200 ng/mL	60.5	31.7
		300 ng/mL	26.6	51.6
		400 ng/mL	77.3	53.4
Trimethoprim (Tri) + Sulfamethizole (Sul)	same as single compounds	0.1 µg/mL (Tri) + 50 µg/mL (Sul)	26.2	31.5
Triton X-100	Membrane	0.01 % (v/v)	53.4	58.7
		0.03 % (v/v)	55.7	54.1
		0.20 % (v/v)	62.8	82.1
Tunicamycin	peptidogylcan biosynthesis: MraY	1.0 µg/mL	81.6	91.4
		3.0 µg/mL	56.6	64.5
		7.5 µg/mL	20.4	68.1
UV	DNA damage	6 seconds	88.0	63.9
		12 seconds	7.9	94.3
		18 seconds	53.2	94.2
		24 seconds	85.7	79.8
Vancomycin	peptidoglycan elongation	10 µg/mL	33.4	67.8
		20 µg/mL	31.6	87.3
		50 µg/mL	81.8	83.3
Verapamil	cell division and proton motive force	0.1 mM	57.3	61.9
		0.5 mM	34.0	16.3
		1.0 mM	20.2	39.3

References

Aeckersberg, F., Lupp, C., Feliciano, B. & Ruby, E. G. (2001). *Vibrio fischeri* outer membrane protein OmpU plays a role in normal symbiotic colonization. *J Bacteriol* **183**, 6590-6597.

Amor, K., Heinrichs, D. E., Frirdich, E., Ziebell, K., Johnson, R. P. & Whitfield, C. (2000). Distribution of core oligosaccharide types in lipopolysaccharides from *Escherichia coli*. *Infect Immun* **68**, 1116-1124.

Anderson, T. F. (1949). On the mechanism of adsorption of bacteriophages on host cells. In *The Nature of the Bacterial Surface*, pp. 76-95. Edited by A. A. Miles & N. W. Pirie. Oxford, UK: Oxford University Press.

Antão, E. M., Ewers, C., Gürlebeck, D., Preisinger, R., Homeier, T., Li, G. & Wieler, L. H. (2009). Signature-tagged mutagenesis in a chicken infection model leads to the identification of a novel avian pathogenic *Escherichia coli* fimbrial adhesin. *PLoS One* **4**, e7796.

Appelmelk, B. J., An, Y. Q., Hekker, T. A., Thijs, L. G., MacLaren, D. M. & de Graaf, J. (1994). Frequencies of lipopolysaccharide core types in *Escherichia coli* strains from bacteraemic patients. *Microbiology* **140 (Pt 5)**, 1119-1124.

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

Azari, F., Nyland, L., Yu, C., Radermacher, M., Mintz, K. P. & Ruiz, T. (2013). Ultrastructural analysis of the rugose cell envelope of a member of the *Pasteurellaceae* family. *J Bacteriol* **195**, 1680-1688.

Babu, M., Diaz-Mejia, J. J., Vlasblom, J., Gagarinova, A., Phanse, S., Graham, C., Yousif, F., Ding, H., Xiong, X., Nazarians-Armavil, A., Alamgir, M., Ali, M., Pogoutse, O., Pe'er, A., Arnold, R., Michaut, M., Parkinson, J., Golshani, A., Whitfield, C., Wodak, S. J., Moreno-Hagelsieb, G., Greenblatt, J. F. & Emili, A. (2011). Genetic interaction maps in *Escherichia coli* reveal functional crosstalk among cell envelope biogenesis pathways. *PLoS Genet* **7**, e1002377.

Bachmann, B. J. (1996). Derivations and genotypes of some mutant derivatives of *Escherichia coli* K-12. In Escherichia coli *and* Salmonella typhimurium *Cellular and Molecular Biology*, 2nd edn, pp. 2460-2488. Edited by F. C. Neidhardt, R. Curtis III, J. L. Ingraham, E. C. C. Lin, K. B. Low Jr, B. Magasanik, W. S. Rezmikoff, M. Riley, M. Schaechter & H. E. Umbarger. Washington, DC: ASM Press.

Bader, M. W., Xie, T., Yu, C. A. & Bardwell, J. C. (2000). Disulfide bonds are generated by quinone reduction. *J Biol Chem* **275**, 26082-26088.

Båga, M., Norgren, M. & Normark, S. (1987). Biogenesis of *E. coli* Pap pili: papH, a minor pilin subunit involved in cell anchoring and length modulation. *Cell* **49**, 241-251.

Balashova, N. V., Park, D. H., Patel, J. K., Figurski, D. H. & Kachlany, S. C. (2007). Interaction between leukotoxin and Cu,Zn superoxide dismutase in *Aggregatibacter actinomycetemcomitans*. *Infect Immun* **75**, 4490-4497.

Baldi, D. L., Higginson, E. E., Hocking, D. M., Praszkier, J., Cavaliere, R., James, C. E., Bennett-Wood, V., Azzopardi, K. I., Turnbull, L., Lithgow, T., Robins-Browne, R. M., Whitchurch, C. B. & Tauschek, M. (2012). The type II secretion system and its ubiquitous lipoprotein substrate, SslE, are required for biofilm formation and virulence of enteropathogenic *Escherichia coli*. *Infect Immun* **80**, 2042-2052.

Barton, L. L. (2005). *Structural and Functional Relationships in Prokaryotes*. New York: Springer New York.

Basle, A., Rummel, G., Storici, P., Rosenbusch, J. P. & Schirmer, T. (2006). Crystal structure of osmoporin OmpC from *E. coli* at 2.0 A. *J Mol Biol* **362**, 933-942.

Bečárová, Z. (2015). Mechanism of FimI, the assembly termination subunit of the type 1 pili from uropathogenic *Escherichia coli*. Switzerland: ETH Zürich.

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

Bendtsen, J. D., Kiemer, L., Fausboll, A. & Brunak, S. (2005a). Non-classical protein secretion in bacteria. *BMC Microbiol* **5**, 58.

Bendtsen, J. D., Nielsen, H., Widdick, D., Palmer, T. & Brunak, S. (2005b). Prediction of twin-arginine signal peptides. *BMC Bioinformatics* **6**, 167.

Bennion, D., Charlson, E. S., Coon, E. & Misra, R. (2010). Dissection of β-barrel outer membrane protein assembly pathways through characterizing BamA POTRA 1 mutants of *Escherichia coli*. *Mol Microbiol* **77**, 1153-1171.

Berkmen, M. (2012). Production of disulfide-bonded proteins in *Escherichia coli*. *Protein Expr Purif* **82**, 240-251.

Berks, B. C. (1996). A common export pathway for proteins binding complex redox cofactors? *Mol Microbiol* **22**, 393-404.

Bernsel, A. & Daley, D. O. (2009). Exploring the inner membrane proteome of *Escherichia coli*: which proteins are eluding detection and why? *Trends Microbiol* **17**, 444-449.

Berry, J. L. & Pelicic, V. (2015). Exceptionally widespread nanomachines composed of type IV pilins: the prokaryotic Swiss Army knives. *FEMS Microbiol Rev* **39**, 134-154.

Bhinder, G., Sham, H. P., Chan, J. M., Morampudi, V., Jacobson, K. & Vallance, B. A. (2013). The Citrobacter rodentium mouse model: studying pathogen and host contributions to infectious colitis. *J Vis Exp*, e50222.

Bodelón, G., Marin, E. & Fernández, L. Á. (2009). Role of periplasmic chaperones and BamA (YaeT/Omp85) in folding and secretion of intimin from enteropathogenic *Escherichia coli* strains. *J Bacteriol* **191**, 5169-5179.

Bodelón, G., Palomino, C. & Fernández, L. Á. (2013). Immunoglobulin domains in *Escherichia coli* and other enterobacteria: from pathogenesis to applications in antibody technologies. *FEMS Microbiol Rev* **37**, 204-250.

Bonardi, F., Halza, E., Walko, M., Du Plessis, F., Nouwen, N., Feringa, B. L. & Driessen, A. J. (2011). Probing the SecYEG translocation pore size with preproteins conjugated with sizable rigid spherical molecules. *Proc Natl Acad Sci U S A* **108**, 7775-7780.

Bongrand, C., Koch, E. J., Moriano-Gutierrez, S., Cordero, O. X., McFall-Ngai, M., Polz, M. F. & Ruby, E. G. (2016). A genomic comparison of 13 symbiotic *Vibrio fischeri* isolates from the perspective of their host source and colonization behavior. *ISME J*.

Bork, P., Holm, L. & Sander, C. (1994). The immunoglobulin fold. Structural classification, sequence patterns and common core. *J Mol Biol* **242**, 309-320.

Braun, V., Bosch, V., Hantke, K. & Schaller, K. (1974). Structure and biosynthesis of functionally defined areas of the *Escherichia coli* outer membrane. *Ann N Y Acad Sci* **235**, 66-82.

Brinton, C. C., Jr. (1959). Non-flagellar appendages of bacteria. *Nature* **183**, 782-786.

Brooks, J. F., 2nd, Gyllborg, M. C., Cronin, D. C., Quillin, S. J., Mallama, C. A., Foxall, R., Whistler, C., Goodman, A. L. & Mandel, M. J. (2014). Global discovery of colonization determinants in the squid symbiont *Vibrio fischeri*. *Proc Natl Acad Sci U S A* **111**, 17284-17289.

Browning, D. F., Wells, T. J., Franca, F. L., Morris, F. C., Sevastsyanovich, Y. R., Bryant, J. A., Johnson, M. D., Lund, P. A., Cunningham, A. F., Hobman, J. L., May, R. C., Webber, M. A. & Henderson, I. R. (2013). Laboratory adapted *Escherichia coli* K-12 becomes a pathogen of *Caenorhabditis elegans* upon restoration of O antigen biosynthesis. *Mol Microbiol* **87**, 939-950.

Buddelmeijer, N. (2015). The molecular mechanism of bacterial lipoprotein modification--how, when and why? *FEMS Microbiol Rev* **39**, 246-261.

Burall, L. S., Harro, J. M., Li, X., Lockatell, C. V., Himpsl, S. D., Hebel, J. R., Johnson, D. E. & Mobley, H. L. (2004). *Proteus mirabilis* genes that contribute to pathogenesis of urinary tract infection: identification of 25 signature-tagged mutants attenuated at least 100-fold. *Infect Immun* **72**, 2922-2938.

Burkinshaw, B. J., Deng, W., Lameignere, E., Wasney, G. A., Zhu, H., Worrall, L. J., Finlay, B. B. & Strynadka, N. C. (2015). Structural analysis of a specialized type III secretion system peptidoglycan-cleaving enzyme. *J Biol Chem* **290**, 10406-10417.

Burkinshaw, B. J. & Strynadka, N. C. (2014). Assembly and structure of the T3SS. *Biochim Biophys Acta* **1843**, 1649-1663.

Burmann, B. M., Wang, C. & Hiller, S. (2013). Conformation and dynamics of the periplasmic membrane-protein-chaperone complexes OmpX-Skp and tOmpA-Skp. *Nat Struct Mol Biol* **20**, 1265-1272.

Busch, A., Phan, G. & Waksman, G. (2015). Molecular mechanism of bacterial type 1 and P pili assembly. *Philos Trans A Math Phys Eng Sci* **373**.

Busch, A. & Waksman, G. (2012). Chaperone-usher pathways: diversity and pilus assembly mechanism. *Philos Trans R Soc Lond B Biol Sci* **367**, 1112-1122.

Cabezón, E., Ripoll-Rozada, J., Pena, A., de la Cruz, F. & Arechaga, I. (2015). Towards an integrated model of bacterial conjugation. *FEMS Microbiol Rev* **39**, 81-95.

Cao, B., Zhao, Y., Kou, Y., Ni, D., Zhang, X. C. & Huang, Y. (2014). Structure of the nonameric bacterial amyloid secretion channel. *Proc Natl Acad Sci U S A* **111**, E5439-5444.

Casadei, M. A., Manas, P., Niven, G., Needs, E. & Mackey, B. M. (2002). Role of membrane fluidity in pressure resistance of *Escherichia coli* NCTC 8164. *Appl Environ Microbiol* **68**, 5965-5972.

Caspi, R., Altman, T., Billington, R., Dreher, K., Foerster, H., Fulcher, C. A., Holland, T. A., Keseler, I. M., Kothari, A., Kubo, A., Krummenacker, M., Latendresse, M., Mueller, L. A., Ong, Q., Paley, S., Subhraveti, P., Weaver, D. S., Weerasinghe, D., Zhang, P. & Karp, P. D. (2014). The MetaCyc database of metabolic pathways and enzymes and the BioCyc collection of Pathway/Genome Databases. *Nucleic Acids Res* **42**, D459-471.

Celik, N., Webb, C. T., Leyton, D. L., Holt, K. E., Heinz, E., Gorrell, R., Kwok, T., Naderer, T., Strugnell, R. A., Speed, T. P., Teasdale, R. D., Likic, V. A. & Lithgow, T. (2012). A bioinformatic strategy for the detection, classification and analysis of bacterial autotransporters. *PLoS One* **7**, e43245.

Cha, H. J., Muller, R. T. & Pos, K. M. (2014). Switch-loop flexibility affects transport of large drugs by the promiscuous AcrB multidrug efflux transporter. *Antimicrob Agents Chemother* **58**, 4767-4772.

Charlson, E. S., Werner, J. N. & Misra, R. (2006). Differential effects of *yfgL* mutation on *Escherichia coli* outer membrane proteins and lipopolysaccharide. *J Bacteriol* **188**, 7186-7194.

Chart, H., Smith, H. R., La Ragione, R. M. & Woodward, M. J. (2000). An investigation into the pathogenic properties of *Escherichia coli* strains BLR, BL21, DH5alpha and EQ1. *J Appl Microbiol* **89**, 1048-1058.

Chimalakonda, G., Ruiz, N., Chng, S. S., Garner, R. A., Kahne, D. & Silhavy, T. J. (2011). Lipoprotein LptE is required for the assembly of LptD by the β-barrel assembly machine in the outer membrane of *Escherichia coli*. *Proc Natl Acad Sci U S A* **108**, 2492-2497.

Chingcuanco, F., Yu, Y., Kus, J. V., Que, L., Lackraj, T., Levesque, C. M. & Barnett Foster, D. (2012). Identification of a novel adhesin involved in acid-induced adhesion of enterohaemorrhagic *Escherichia coli* O157 : H7. *Microbiology* **158**, 2399-2407.

Choi, D. S., Yamada, H., Mizuno, T. & Mizushima, S. (1986). Trimeric structure and localization of the major lipoprotein in the cell surface of *Escherichia coli*. *J Biol Chem* **261**, 8953-8957.

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

Chou, S., Bui, N. K., Russell, A. B., Lexa, K. W., Gardiner, T. E., LeRoux, M., Vollmer, W. & Mougous, J. D. (2012). Structure of a peptidoglycan amidase effector targeted to Gram-negative bacteria by the type VI secretion system. *Cell Rep* **1**, 656-664.

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

Clements, A., Bursac, D., Gatsos, X., Perry, A. J., Civciristov, S., Celik, N., Likic, V. A., Poggio, S., Jacobs-Wagner, C., Strugnell, R. A. & Lithgow, T. (2009). The reducible complexity of a mitochondrial molecular machine. *Proc Natl Acad Sci U S A* **106**, 15791-15795.

Cold Spring Harbor (CSH) Protocols (2006a). Low-salt LB medium. *Cold Spring Harbor Protocols* **2006**, pdb.rec10789.

Cold Spring Harbor (CSH) Protocols (2006b). M9 Salts. *Cold Spring Harbor Protocols* **2006**, pdb.rec614.

Cold Spring Harbor (CSH) Protocols (2010). M9 minimal medium (standard). *Cold Spring Harbor Protocols* **2010**, pdb.rec12295.

Cold Spring Harbor (CSH) Protocols (2012). SOC Medium for E. coli. *Cold Spring Harbor Protocols* **2012**, pdb.rec069732.

Cold Spring Harbor (CSH) Protocols (2013). TBST for Western Blotting. *Cold Spring Harbor Protocols* **2013**, pdb.rec074104.

Cold Spring Harbor (CSH) Protocols (2014). 2× YT Medium. *Cold Spring Harbor Protocols* **2014**, pdb.rec082222.

Cold Spring Harbor (CSH) Protocols (2016). Amino Acid Supplement Powder (Drop-Out Mixture). *Cold Spring Harbor Protocols* **2016**, pdb.rec089805.

Costa, T. R., Felisberto-Rodrigues, C., Meir, A., Prevost, M. S., Redzej, A., Trokter, M. & Waksman, G. (2015). Secretion systems in Gram-negative bacteria: structural and mechanistic insights. *Nat Rev Microbiol* **13**, 343-359.

Cowan, S. W., Schirmer, T., Rummel, G., Steiert, M., Ghosh, R., Pauptit, R. A., Jansonius, J. N. & Rosenbusch, J. P. (1992). Crystal structures explain functional properties of two *E. coli* porins. *Nature* **358**, 727-733.

Cowles, C. E., Li, Y., Semmelhack, M. F., Cristea, I. M. & Silhavy, T. J. (2011). The free and bound forms of Lpp occupy distinct subcellular locations in *Escherichia coli*. *Mol Microbiol* **79**, 1168-1181.

Craig, L. & Li, J. (2008). Type IV pili: paradoxes in form and function. *Curr Opin Struct Biol* **18**, 267-277.

Cristóbal, S., de Gier, J. W., Nielsen, H. & von Heijne, G. (1999). Competition between Sec- and TAT-dependent protein translocation in *Escherichia coli*. *EMBO J* **18**, 2982-2990.

Cronan, J. E. (2001). *Escherichia coli* as an experimental organism. In *eLS*: John Wiley & Sons, Ltd.

Crowlesmith, I., Gamon, K. & Henning, U. (1981). Precursor proteins are intermediates in vivo in the synthesis of two major outer membrane proteins, the OmpA and OmpF proteins, of *Escherichia coli* K12. *Eur J Biochem* **113**, 375-380.

Daegelen, P., Studier, F. W., Lenski, R. E., Cure, S. & Kim, J. F. (2009). Tracing ancestors and relatives of *Escherichia coli* B, and the derivation of B strains REL606 and BL21(DE3). *J Mol Biol* **394**, 634-643.

Dahl, J. U., Koldewey, P., Salmon, L., Horowitz, S., Bardwell, J. C. & Jakob, U. (2015). HdeB functions as an acid-protective chaperone in bacteria. *J Biol Chem* **290**, 65-75.

Dalbey, R. E., Kuhn, A., Zhu, L. & Kiefer, D. (2014). The membrane insertase YidC. *Biochim Biophys Acta* **1843**, 1489-1496.

Dalebroux, Z. D., Edrozo, M. B., Pfuetzner, R. A., Ressl, S., Kulasekara, B. R., Blanc, M. P. & Miller, S. I. (2015). Delivery of cardiolipins to the *Salmonella* outer membrane is necessary for survival within host tissues and virulence. *Cell Host Microbe* **17**, 441-451.

de Cock, H., Struyve, M., Kleerebezem, M., van der Krift, T. & Tommassen, J. (1997). Role of the carboxy-terminal phenylalanine in the biogenesis of outer membrane protein PhoE of *Escherichia coli* K-12. *J Mol Biol* **269**, 473-478.

de Jonge, B. L. (1990). Isogenic variants of *Escherichia coli* with altered morphology have peptidoglycan with identical muropeptide composition. *J Bacteriol* **172**, 4682-4684.

de Jonge, B. L., Wientjes, F. B., Jurida, I., Driehuis, F., Wouters, J. T. & Nanninga, N. (1989). Peptidoglycan synthesis during the cell cycle of *Escherichia coli*: composition and mode of insertion. *J Bacteriol* **171**, 5783-5794.

DeCanio, M. S., Landick, R. & Haft, R. J. (2013). The non-pathogenic *Escherichia coli* strain W secretes SslE via the virulence-associated type II secretion system β. *BMC Microbiol* **13**, 130.

Dekker, N. (2000). Outer-membrane phospholipase A: known structure, unknown biological function. *Mol Microbiol* **35**, 711-717.

Demchick, P. & Koch, A. L. (1996). The permeability of the wall fabric of *Escherichia coli* and *Bacillus subtilis*. *J Bacteriol* **178**, 768-773.

Denks, K., Vogt, A., Sachelaru, I., Petriman, N. A., Kudva, R. & Koch, H. G. (2014). The Sec translocon mediated protein transport in prokaryotes and eukaryotes. *Mol Membr Biol* **31**, 58-84.

Denoncin, K., Vertommen, D., Arts, I. S., Goemans, C. V., Rahuel-Clermont, S., Messens, J. & Collet, J. F. (2014). A new role for *Escherichia coli* DsbC protein in protection against oxidative stress. *J Biol Chem* **289**, 12356-12364.

Depuydt, M., Leonard, S. E., Vertommen, D., Denoncin, K., Morsomme, P., Wahni, K., Messens, J., Carroll, K. S. & Collet, J. F. (2009). A periplasmic reducing system protects single cysteine residues from oxidation. *Science* **326**, 1109-1111.

Desbois, A. P. & Smith, V. J. (2010). Antibacterial free fatty acids: activities, mechanisms of action and biotechnological potential. *Appl Microbiol Biotechnol* **85**, 1629-1642.

Desvaux, M., Scott-Tucker, A., Turner, S. M., Cooper, L. M., Huber, D., Nataro, J. P. & Henderson, I. R. (2007). A conserved extended signal peptide region directs posttranslational protein translocation via a novel mechanism. *Microbiology* **153**, 59-70.

Doerrler, W. T., Gibbons, H. S. & Raetz, C. R. (2004). MsbA-dependent translocation of lipids across the inner membrane of *Escherichia coli*. *J Biol Chem* **279**, 45102-45109.

Doerrler, W. T., Reedy, M. C. & Raetz, C. R. (2001). An *Escherichia coli* mutant defective in lipid export. *J Biol Chem* **276**, 11461-11464.

Dong, C., Beis, K., Nesper, J., Brunkan-Lamontagne, A. L., Clarke, B. R., Whitfield, C. & Naismith, J. H. (2006). Wza the translocon for *E. coli* capsular polysaccharides defines a new class of membrane protein. *Nature* **444**, 226-229.

Dong, H., Xiang, Q., Gu, Y., Wang, Z., Paterson, N. G., Stansfeld, P. J., He, C., Zhang, Y., Wang, W. & Dong, C. (2014). Structural basis for outer membrane lipopolysaccharide insertion. *Nature* **511**, 52-56.

Donohue-Rolfe, A. M. & Schaechter, M. (1980). Translocation of phospholipids from the inner to the outer membrane of *Escherichia coli*. *Proc Natl Acad Sci U S A* **77**, 1867-1871.

Dowhan, W. (1997). Molecular basis for membrane phospholipid diversity: why are there so many lipids? *Annu Rev Biochem* **66**, 199-232.

du Plessis, D. J., Berrelkamp, G., Nouwen, N. & Driessen, A. J. (2009). The lateral gate of SecYEG opens during protein translocation. *J Biol Chem* **284**, 15805-15814.

Duguid, J. P., Smith, I. W., Dempster, G. & Edmunds, P. N. (1955). Non-flagellar filamentous appendages (fimbriae) and haemagglutinating activity in *Bacterium coli*. *J Pathol Bacteriol* **70**, 335-348.

Dunstan, R. A., Hay, I. D., Wilksch, J. J., Schittenhelm, R. B., Purcell, A. W., Clark, J., Costin, A., Ramm, G., Strugnell, R. A. & Lithgow, T. (2015). Assembly of the secretion pores GspD, Wza and CsgG into bacterial outer membranes does not require the Omp85 proteins BamA or TamA. *Mol Microbiol* **97**, 616-629.

Dunstan, R. A., Heinz, E., Wijeyewickrema, L. C., Pike, R. N., Purcell, A. W., Evans, T. J., Praszkier, J., Robins-Browne, R. M., Strugnell, R. A., Korotkov, K. V. & Lithgow, T. (2013). Assembly of the type II secretion system such as found in *Vibrio cholerae* depends on the novel Pilotin AspS. *PLoS Pathog* **9**, e1003117.

Dutton, R. J., Boyd, D., Berkmen, M. & Beckwith, J. (2008). Bacterial species exhibit diversity in their mechanisms and capacity for protein disulfide bond formation. *Proc Natl Acad Sci U S A* **105**, 11933-11938.

Eckford, P. D. & Sharom, F. J. (2010). The reconstituted *Escherichia coli* MsbA protein displays lipid flippase activity. *Biochem J* **429**, 195-203.

Edgar, R. C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* **32**, 1792-1797.

Elofsson, A. & von Heijne, G. (2007). Membrane protein structure: prediction versus reality. *Annu Rev Biochem* **76**, 125-140.

Emmerling, M., Dauner, M., Ponti, A., Fiaux, J., Hochuli, M., Szyperski, T., Wuthrich, K., Bailey, J. E. & Sauer, U. (2002). Metabolic flux responses to pyruvate kinase knockout in *Escherichia coli*. *J Bacteriol* **184**, 152-164.

Endo, T., Kawano, S. & Yamano, K. (2011). BamE structure: the assembly of β-barrel proteins in the outer membranes of bacteria and mitochondria. *EMBO Rep* **12**, 94-95.

Estrada Mallarino, L., Fan, E., Odermatt, M., Muller, M., Lin, M., Liang, J., Heinzelmann, M., Fritsche, F., Apell, H. J. & Welte, W. (2015). TtOmp85, a β-barrel assembly protein, functions by barrel augmentation. *Biochemistry* **54**, 844-852.

Evans, M. L. & Chapman, M. R. (2014). Curli biogenesis: order out of disorder. *Biochim Biophys Acta* **1843**, 1551-1558.

Fischer, E. & Sauer, U. (2003). A novel metabolic cycle catalyzes glucose oxidation and anaplerosis in hungry *Escherichia coli*. *J Biol Chem* **278**, 46446-46451.

Fleming, K. G. (2015). A combined kinetic push and thermodynamic pull as driving forces for outer membrane protein sorting and folding in bacteria. *Philos Trans R Soc Lond B Biol Sci* **370**.

Fleming, P. J., Patel, D. S., Wu, E. L., Qi, Y., Yeom, M. S., Sousa, M. C., Fleming, K. G. & Im, W. (2016). BamA POTRA Domain Interacts with a Native Lipid Membrane Surface. *Biophys J* **110**, 2698-2709.

Frankel, G., Phillips, A. D., Novakova, M., Field, H., Candy, D. C., Schauer, D. B., Douce, G. & Dougan, G. (1996). Intimin from enteropathogenic *Escherichia coli* restores murine virulence to a *Citrobacter rodentium eaeA* mutant: induction of an immunoglobulin A response to intimin and EspB. *Infect Immun* **64**, 5315-5325.

Freinkman, E., Chng, S. S. & Kahne, D. (2011). The complex that inserts lipopolysaccharide into the bacterial outer membrane forms a two-protein plug-and-barrel. *Proc Natl Acad Sci U S A* **108**, 2486-2491.

Frirdich, E. & Whitfield, C. (2005). Lipopolysaccharide inner core oligosaccharide structure and outer membrane stability in human pathogens belonging to the *Enterobacteriaceae*. *J Endotoxin Res* **11**, 133-144.

Gallant, C. V., Sedic, M., Chicoine, E. A., Ruiz, T. & Mintz, K. P. (2008). Membrane morphology and leukotoxin secretion are associated with a novel membrane protein of *Aggregatibacter actinomycetemcomitans*. *J Bacteriol* **190**, 5972-5980.

Gally, D. L., Bogan, J. A., Eisenstein, B. I. & Blomfield, I. C. (1993). Environmental regulation of the *fim* switch controlling type 1 fimbrial phase variation in *Escherichia coli* K-12: effects of temperature and media. *J Bacteriol* **175**, 6186-6193.

Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S. e., Wilkins, M. R., Appel, R. D. & Bairoch, A. (2005). Protein identification and analysis tools on the ExPASy server. In *The Proteomics Protocols Handbook*, pp. 571-607. Edited by J. M. Walker. Totowa, NJ: Humana Press.

Ge, X., Lyu, Z. X., Liu, Y., Wang, R., Zhao, X. S., Fu, X. & Chang, Z. (2014a). Identification of FkpA as a key quality control factor for the biogenesis of outer membrane proteins under heat shock conditions. *J Bacteriol* **196**, 672-680.

Ge, X., Wang, R., Ma, J., Liu, Y., Ezemaduka, A. N., Chen, P. R., Fu, X. & Chang, Z. (2014b). DegP primarily functions as a protease for the biogenesis of β-barrel outer membrane proteins in the Gram-negative bacterium *Escherichia coli*. *FEBS J* **281**, 1226-1240.

Geibel, S., Procko, E., Hultgren, S. J., Baker, D. & Waksman, G. (2013). Structural and energetic basis of folded-protein transport by the FimD usher. *Nature* **496**, 243-246.

Geibel, S. & Waksman, G. (2014). The molecular dissection of the chaperone-usher pathway. *Biochim Biophys Acta* **1843**, 1559-1567.

Gennity, J. M. & Inouye, M. (1991). The protein sequence responsible for lipoprotein membrane localization in *Escherichia coli* exhibits remarkable specificity. *J Biol Chem* **266**, 16458-16464.

Gessmann, D., Chung, Y. H., Danoff, E. J., Plummer, A. M., Sandlin, C. W., Zaccai, N. R. & Fleming, K. G. (2014). 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.

Gibbs, R. J., Stewart, J. & Poxton, I. R. (2004). The distribution of, and antibody response to, the core lipopolysaccharide region of *Escherichia coli* isolated from the faeces of healthy humans and cattle. *J Med Microbiol* **53**, 959-964.

Giltner, C. L., Nguyen, Y. & Burrows, L. L. (2012). Type IV pilin proteins: versatile molecular modules. *Microbiol Mol Biol Rev* **76**, 740-772.

Glauert, A. M. & Thornley, M. J. (1969). The topography of the bacterial cell wall. *Annu Rev Microbiol* **23**, 159-198.

Glauner, B. (1988). Separation and quantification of muropeptides with high-performance liquid chromatography. *Anal Biochem* **172**, 451-464.

Glauner, B., Holtje, J. V. & Schwarz, U. (1988). The composition of the murein of *Escherichia coli*. *J Biol Chem* **263**, 10088-10095.

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

Gohlke, U., Pullan, L., McDevitt, C. A., Porcelli, I., de Leeuw, E., Palmer, T., Saibil, H. R. & Berks, B. C. (2005). The TatA component of the twin-arginine protein transport system forms channel complexes of variable diameter. *Proc Natl Acad Sci U S A* **102**, 10482-10486.

Goldberg, D. E., Rumley, M. K. & Kennedy, E. P. (1981). Biosynthesis of membrane-derived oligosaccharides: a periplasmic phosphoglyceroltransferase. *Proc Natl Acad Sci U S A* **78**, 5513-5517.

Goodman, A. L., Wu, M. & Gordon, J. I. (2011). Identifying microbial fitness determinants by insertion sequencing using genome-wide transposon mutant libraries. *Nat Protoc* **6**, 1969-1980.

Götzke, H., Muheim, C., Altelaar, A. F., Heck, A. J., Maddalo, G. & Daley, D. O. (2015). Identification of putative substrates for the periplasmic chaperone YfgM in *Escherichia coli* using quantitative proteomics. *Mol Cell Proteomics* **14**, 216-226.

Götzke, H., Palombo, I., Muheim, C., Perrody, E., Genevaux, P., Kudva, R., Müller, M. & Daley, D. O. (2014). YfgM is an ancillary subunit of the SecYEG translocon in *Escherichia coli*. *J Biol Chem* **289**, 19089-19097.

Goyal, P., Krasteva, P. V., Van Gerven, N., Gubellini, F., Van den Broeck, I., Troupiotis-Tsailaki, A., Jonckheere, W., Pehau-Arnaudet, G., Pinkner, J. S., Chapman, M. R., Hultgren, S. J.,

Howorka, S., Fronzes, R. & Remaut, H. (2014). Structural and mechanistic insights into the bacterial amyloid secretion channel CsgG. *Nature* **516**, 250-253.

Gruss, F., Zahringer, F., Jakob, R. P., Burmann, B. M., Hiller, S. & Maier, T. (2013). The structural basis of autotransporter translocation by TamA. *Nat Struct Mol Biol* **20**, 1318-1320.

Gu, Y., Li, H., Dong, H., Zeng, Y., Zhang, Z., Paterson, N. G., Stansfeld, P. J., Wang, Z., Zhang, Y., Wang, W. & Dong, C. (2016). Structural basis of outer membrane protein insertion by the BAM complex. *Nature* **531**, 64-69.

Gupta, S. D., Dowhan, W. & Wu, H. C. (1991). Phosphatidylethanolamine is not essential for the *N*-acylation of apolipoprotein in *Escherichia coli*. *J Biol Chem* **266**, 9983-9986.

Habenstein, B., Loquet, A., Hwang, S., Giller, K., Vasa, S. K., Becker, S., Habeck, M. & Lange, A. (2015). Hybrid structure of the type 1 pilus of uropathogenic *Escherichia coli*. *Angew Chem Int Ed Engl* **54**, 11691-11695.

Hagge, S. O., de Cock, H., Gutsmann, T., Beckers, F., Seydel, U. & Wiese, A. (2002). Pore formation and function of phosphoporin PhoE of *Escherichia coli* are determined by the core sugar moiety of lipopolysaccharide. *J Biol Chem* **277**, 34247-34253.

Hahn, E., Wild, P., Hermanns, U., Sebbel, P., Glockshuber, R., Haner, M., Taschner, N., Burkhard, P., Aebi, U. & Muller, S. A. (2002). Exploring the 3D molecular architecture of *Escherichia coli* type 1 pili. *J Mol Biol* **323**, 845-857.

Harvey, H. A., Swords, W. E. & Apicella, M. A. (2001). The mimicry of human glycolipids and glycosphingolipids by the lipooligosaccharides of pathogenic *Neisseria* and *Haemophilus*. *J Autoimmun* **16**, 257-262.

Heinrichs, D. E., Yethon, J. A., Amor, P. A. & Whitfield, C. (1998). The assembly system for the outer core portion of R1- and R4-type lipopolysaccharides of *Escherichia coli*. The R1 core-specific β-glucosyltransferase provides a novel attachment site for O-polysaccharides. *J Biol Chem* **273**, 29497-29505.

Heinz, E. & Lithgow, T. (2014). A comprehensive analysis of the Omp85/TpsB protein superfamily structural diversity, taxonomic occurrence, and evolution. *Front Microbiol* **5**, 370.

Heinz, E., Selkrig, J., Belousoff, M. J. & Lithgow, T. (2015). Evolution of the Translocation and Assembly Module (TAM). *Genome Biol Evol* **7**, 1628-1643.

Heinz, E., Stubenrauch, C. J., Grinter, R., Croft, N. P., Purcell, A. W., Strugnell, R. A., Dougan, G. & Lithgow, T. (2016). Conserved features in the structure, mechanism, and biogenesis of the inverse autotransporter protein family. *Genome Biol Evol* **8**, 1690-1705.

Henderson, I. R., Owen, P. & Nataro, J. P. (1999). Molecular switches--the ON and OFF of bacterial phase variation. *Mol Microbiol* **33**, 919-932.

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

Henry, R., Vithanage, N., Harrison, P., Seemann, T., Coutts, S., Moffatt, J. H., Nation, R. L., Li, J., Harper, M., Adler, B. & Boyce, J. D. (2012). Colistin-resistant, lipopolysaccharide-deficient *Acinetobacter baumannii* responds to lipopolysaccharide loss through increased expression of genes involved in the synthesis and transport of lipoproteins, phospholipids, and poly-β-1,6-*N*-acetylglucosamine. *Antimicrob Agents Chemother* **56**, 59-69.

Herlihey, F. A., Moynihan, P. J. & Clarke, A. J. (2014). The essential protein for bacterial flagella formation FlgJ functions as a β-*N*-acetylglucosaminidase. *J Biol Chem* **289**, 31029-31042.

Höhr, A. I., Straub, S. P., Warscheid, B., Becker, T. & Wiedemann, N. (2015). Assembly of β-barrel proteins in the mitochondrial outer membrane. *Biochim Biophys Acta* **1853**, 74-88.

Hong, W., Jiao, W., Hu, J., Zhang, J., Liu, C., Fu, X., Shen, D., Xia, B. & Chang, Z. (2005). Periplasmic protein HdeA exhibits chaperone-like activity exclusively within stomach pH range by transforming into disordered conformation. *J Biol Chem* **280**, 27029-27034.

Hospenthal, M. K., Redzej, A., Dodson, K., Ukleja, M., Frenz, B., Rodrigues, C., Hultgren, S. J., DiMaio, F., Egelman, E. H. & Waksman, G. (2016). Structure of a chaperone-usher pilus reveals the molecular basis of rod uncoiling. *Cell* **164**, 269-278.

Huijbregts, R. P., de Kroon, A. I. & de Kruijff, B. (1998). Rapid transmembrane movement of newly synthesized phosphatidylethanolamine across the inner membrane of *Escherichia coli*. *J Biol Chem* **273**, 18936-18942.

Humphries, A. D., Townsend, S. M., Kingsley, R. A., Nicholson, T. L., Tsolis, R. M. & Baumler, A. J. (2001). Role of fimbriae as antigens and intestinal colonization factors of *Salmonella* serovars. *FEMS Microbiol Lett* **201**, 121-125.

Ilangovan, A., Connery, S. & Waksman, G. (2015). Structural biology of the Gram-negative bacterial conjugation systems. *Trends Microbiol* **23**, 301-310.

Jabs, A., Weiss, M. S. & Hilgenfeld, R. (1999). Non-proline *cis* peptide bonds in proteins. *J Mol Biol* **286**, 291-304.

Jackups, R., Jr. & Liang, J. (2005). Interstrand pairing patterns in β-barrel membrane proteins: the positive-outside rule, aromatic rescue, and strand registration prediction. *J Mol Biol* **354**, 979-993.

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

Jansson, P. E., Lindberg, A. A., Lindberg, B. & Wollin, R. (1981). Structural studies on the hexose region of the core in lipopolysaccharides from *Enterobacteriaceae*. *Eur J Biochem* **115**, 571-577.

Jeong, H., Barbe, V., Lee, C. H., Vallenet, D., Yu, D. S., Choi, S. H., Couloux, A., Lee, S. W., Yoon, S. H., Cattolico, L., Hur, C. G., Park, H. S., Segurens, B., Kim, S. C., Oh, T. K., Lenski, R. E., Studier, F. W., Daegelen, P. & Kim, J. F. (2009). Genome sequences of *Escherichia coli* B strains REL606 and BL21(DE3). *J Mol Biol* **394**, 644-652.

Jepson, B. J., Mohan, S., Clarke, T. A., Gates, A. J., Cole, J. A., Butler, C. S., Butt, J. N., Hemmings, A. M. & Richardson, D. J. (2007). Spectropotentiometric and structural analysis of the periplasmic nitrate reductase from *Escherichia coli*. *J Biol Chem* **282**, 6425-6437.

Johansson, M. U., Alioth, S., Hu, K., Walser, R., Koebnik, R. & Pervushin, K. (2007). A minimal transmembrane β-barrel platform protein studied by nuclear magnetic resonance. *Biochemistry* **46**, 1128-1140.

Johnson, J. R., Stell, A. L., Scheutz, F., O'Bryan, T. T., Russo, T. A., Carlino, U. B., Fasching, C., Kavle, J., Van Dijk, L. & Gaastra, W. (2000). Analysis of the F antigen-specific papA alleles of extraintestinal pathogenic Escherichia coli using a novel multiplex PCR-based assay. *Infect Immun* **68**, 1587-1599.

Jong, W. S. & Luirink, J. (2008). The conserved extension of the Hbp autotransporter signal peptide does not determine targeting pathway specificity. *Biochem Biophys Res Commun* **368**, 522-527.

Jong, W. S., ten Hagen-Jongman, C. M., Ruijter, E., Orru, R. V., Genevaux, P. & Luirink, J. (2010). YidC is involved in the biogenesis of the secreted autotransporter hemoglobin protease. *J Biol Chem* **285**, 39682-39690.

Juncker, A. S., Willenbrock, H., Von Heijne, G., Brunak, S., Nielsen, H. & Krogh, A. (2003). Prediction of lipoprotein signal peptides in Gram-negative bacteria. *Protein Sci* **12**, 1652-1662.

Justice, S. S., Hunstad, D. A., Harper, J. R., Duguay, A. R., Pinkner, J. S., Bann, J., Frieden, C., Silhavy, T. J. & Hultgren, S. J. (2005). Periplasmic peptidyl prolyl *cis-trans* isomerases are not essential for viability, but SurA is required for pilus biogenesis in *Escherichia coli*. *J Bacteriol* **187**, 7680-7686.

Kadokura, H., Tian, H., Zander, T., Bardwell, J. C. & Beckwith, J. (2004). Snapshots of DsbA in action: detection of proteins in the process of oxidative folding. *Science* **303**, 534-537.

Kamp, F. & Hamilton, J. A. (2006). How fatty acids of different chain length enter and leave cells by free diffusion. *Prostaglandins Leukot Essent Fatty Acids* **75**, 149-159.

Kang'ethe, W. & Bernstein, H. D. (2013). Charge-dependent secretion of an intrinsically disordered protein via the autotransporter pathway. *Proc Natl Acad Sci U S A* **110**, E4246-4255.

Keenleyside, W. J., Perry, M., Maclean, L., Poppe, C. & Whitfield, C. (1994). A plasmid-encoded $rfb_{O:54}$ gene cluster is required for biosynthesis of the O:54 antigen in *Salmonella enterica* serovar Borreze. *Mol Microbiol* **11**, 437-448.

Keil, B. (1992). *Specificity of Proteolysis*. Berlin: Springer Berlin Heidelberg.

Kelley, L. A., Mezulis, S., Yates, C. M., Wass, M. N. & Sternberg, M. J. (2015). The Phyre2 web portal for protein modeling, prediction and analysis. *Nat Protoc* **10**, 845-858.

Kelly, M., Hart, E., Mundy, R., Marches, O., Wiles, S., Badea, L., Luck, S., Tauschek, M., Frankel, G., Robins-Browne, R. M. & Hartland, E. L. (2006). Essential role of the type III secretion system effector NleB in colonization of mice by *Citrobacter rodentium*. *Infect Immun* **74**, 2328-2337.

Keseler, I. M., Mackie, A., Peralta-Gil, M., Santos-Zavaleta, A., Gama-Castro, S., Bonavides-Martinez, C., Fulcher, C., Huerta, A. M., Kothari, A., Krummenacker, M., Latendresse, M., Muniz-Rascado, L., Ong, Q., Paley, S., Schroder, I., Shearer, A. G., Subhraveti, P., Travers, M., Weerasinghe, D., Weiss, V., Collado-Vides, J., Gunsalus, R. P., Paulsen, I. & Karp, P. D. (2013). EcoCyc: fusing model organism databases with systems biology. *Nucleic Acids Res* **41**, D605-612.

Kim, A. C., Oliver, D. C. & Paetzel, M. (2008). Crystal structure of a bacterial signal Peptide peptidase. *J Mol Biol* **376**, 352-366.

Kim, D. Y. (2015). Two stress sensor proteins for the expression of sigmaE regulon: DegS and RseB. *J Microbiol* **53**, 306-310.

Kleanthous, C., Rassam, P. & Baumann, C. G. (2015). Protein-protein interactions and the spatiotemporal dynamics of bacterial outer membrane proteins. *Curr Opin Struct Biol* **35**, 109-115.

Knowles, T. J., Browning, D. F., Jeeves, M., Maderbocus, R., Rajesh, S., Sridhar, P., Manoli, E., Emery, D., Sommer, U., Spencer, A., Leyton, D. L., Squire, D., Chaudhuri, R. R., Viant, M. R., Cunningham, A. F., Henderson, I. R. & Overduin, M. (2011). Structure and function of BamE within the outer membrane and the β-barrel assembly machine. *EMBO Rep* **12**, 123-128.

Koch, A. L. & Woeste, S. (1992). Elasticity of the sacculus of *Escherichia coli*. *J Bacteriol* **174**, 4811-4819.

Kol, M. A., de Kroon, A. I., Rijkers, D. T., Killian, J. A. & de Kruijff, B. (2001). Membrane-spanning peptides induce phospholipid flop: a model for phospholipid translocation across the inner membrane of *E. coli*. *Biochemistry* **40**, 10500-10506.

Kol, M. A., van Dalen, A., de Kroon, A. I. & de Kruijff, B. (2003a). Translocation of phospholipids is facilitated by a subset of membrane-spanning proteins of the bacterial cytoplasmic membrane. *J Biol Chem* **278**, 24586-24593.

Kol, M. A., van Laak, A. N., Rijkers, D. T., Killian, J. A., de Kroon, A. I. & de Kruijff, B. (2003b). Phospholipid flop induced by transmembrane peptides in model membranes is modulated by lipid composition. *Biochemistry* **42**, 231-237.

Konovalova, A., Perlman, D. H., Cowles, C. E. & Silhavy, T. J. (2014). Transmembrane domain of surface-exposed outer membrane lipoprotein RcsF is threaded through the lumen of β-barrel proteins. *Proc Natl Acad Sci U S A* **111**, E4350-4358.

Konovalova, A. & Silhavy, T. J. (2015). Outer membrane lipoprotein biogenesis: Lol is not the end. *Philos Trans R Soc Lond B Biol Sci* **370**.

Korat, B., Mottl, H. & Keck, W. (1991). Penicillin-binding protein 4 of *Escherichia coli*: molecular cloning of the *dacB* gene, controlled overexpression, and alterations in murein composition. *Mol Microbiol* **5**, 675-684.

Korea, C. G., Badouraly, R., Prevost, M. C., Ghigo, J. M. & Beloin, C. (2010). *Escherichia coli* K-12 possesses multiple cryptic but functional chaperone-usher fimbriae with distinct surface specificities. *Environ Microbiol* **12**, 1957-1977.

Korea, C. G., Ghigo, J. M. & Beloin, C. (2011). The sweet connection: Solving the riddle of multiple sugar-binding fimbrial adhesins in *Escherichia coli*: Multiple *E. coli* fimbriae form a versatile arsenal of sugar-binding lectins potentially involved in surface-colonisation and tissue tropism. *Bioessays* **33**, 300-311.

Kornberg, R. D. & McConnell, H. M. (1971). Inside-outside transitions of phospholipids in vesicle membranes. *Biochemistry* **10**, 1111-1120.

Koronakis, V., Eswaran, J. & Hughes, C. (2004). Structure and function of TolC: the bacterial exit duct for proteins and drugs. *Annu Rev Biochem* **73**, 467-489.

Koronakis, V., Sharff, A., Koronakis, E., Luisi, B. & Hughes, C. (2000). Crystal structure of the bacterial membrane protein TolC central to multidrug efflux and protein export. *Nature* **405**, 914-919.

Kovacs-Simon, A., Titball, R. W. & Michell, S. L. (2011). Lipoproteins of bacterial pathogens. *Infect Immun* **79**, 548-561.

Krogh, A., Larsson, B., von Heijne, G. & Sonnhammer, E. L. (2001). Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J Mol Biol* **305**, 567-580.

Kubelt, J., Menon, A. K., Muller, P. & Herrmann, A. (2002). Transbilayer movement of fluorescent phospholipid analogues in the cytoplasmic membrane of *Escherichia coli*. *Biochemistry* **41**, 5605-5612.

Kutter, E. & Goldman, E. (2015). Introduction to bacteriophages. In *Practical Handbook of Microbiology*, 3rd edn, pp. 855-882. Edited by E. Goldman & L. H. Green. London: CRC Press.

Lai, Y., Rosenshine, I., Leong, J. M. & Frankel, G. (2013). Intimate host attachment: enteropathogenic and enterohaemorrhagic *Escherichia coli*. *Cell Microbiol* **15**, 1796-1808.

Laird, M. W., Kloser, A. W. & Misra, R. (1994). Assembly of LamB and OmpF in deep rough lipopolysaccharide mutants of *Escherichia coli* K-12. *J Bacteriol* **176**, 2259-2264.

Lehr, U., Schutz, M., Oberhettinger, P., Ruiz-Perez, F., Donald, J. W., Palmer, T., Linke, D., Henderson, I. R. & Autenrieth, I. B. (2010). C-terminal amino acid residues of the trimeric autotransporter adhesin YadA of *Yersinia enterocolitica* are decisive for its recognition and assembly by BamA. *Mol Microbiol* **78**, 932-946.

Lehti, T. A., Bauchart, P., Heikkinen, J., Hacker, J., Korhonen, T. K., Dobrindt, U. & Westerlund-Wikstrom, B. (2010). Mat fimbriae promote biofilm formation by meningitis-associated *Escherichia coli*. *Microbiology* **156**, 2408-2417.

Leo, J. C., Grin, I. & Linke, D. (2012). Type V secretion: mechanism(s) of autotransport through the bacterial outer membrane. *Philos Trans R Soc Lond B Biol Sci* **367**, 1088-1101.

Leo, J. C., Oberhettinger, P., Schutz, M. & Linke, D. (2015). The inverse autotransporter family: intimin, invasin and related proteins. *Int J Med Microbiol* **305**, 276-282.

Leonardi, R. & Roach, P. L. (2004). Thiamine biosynthesis in *Escherichia coli*: in vitro reconstitution of the thiazole synthase activity. *J Biol Chem* **279**, 17054-17062.

Levine, M. M., Bergquist, E. J., Nalin, D. R., Waterman, D. H., Hornick, R. B., Young, C. R. & Sotman, S. (1978). *Escherichia coli* strains that cause diarrhoea but do not produce heat-labile or heat-stable enterotoxins and are non-invasive. *Lancet* **1**, 1119-1122.

Leyton, D. L., Johnson, M. D., Thapa, R., Huysmans, G. H., Dunstan, R. A., Celik, N., Shen, H. H., Loo, D., Belousoff, M. J., Purcell, A. W., Henderson, I. R., Beddoe, T., Rossjohn, J., Martin, L. L., Strugnell, R. A. & Lithgow, T. (2014). A mortise-tenon joint in the transmembrane domain modulates autotransporter assembly into bacterial outer membranes. *Nat Commun* **5**, 4239.

Leyton, D. L., Rossiter, A. E. & Henderson, I. R. (2012). From self sufficiency to dependence: mechanisms and factors important for autotransporter biogenesis. *Nat Rev Microbiol* **10**, 213-225.

Li, W., Cowley, A., Uludag, M., Gur, T., McWilliam, H., Squizzato, S., Park, Y. M., Buso, N. & Lopez, R. (2015a). The EMBL-EBI bioinformatics web and programmatic tools framework. *Nucleic Acids Res* **43**, W580-584.

Li, X., Gu, Y., Dong, H., Wang, W. & Dong, C. (2015b). Trapped lipopolysaccharide and LptD intermediates reveal lipopolysaccharide translocation steps across the *Escherichia coli* outer membrane. *Sci Rep* **5**, 11883.

Lillington, J., Geibel, S. & Waksman, G. (2015). Reprint of "Biogenesis and adhesion of type 1 and P pili". *Biochim Biophys Acta* **1850**, 554-564.

Lindberg, F., Lund, B., Johansson, L. & Normark, S. (1987). Localization of the receptor-binding protein adhesin at the tip of the bacterial pilus. *Nature* **328**, 84-87.

Liu, D. & Reeves, P. R. (1994). *Escherichia coli* K12 regains its O antigen. *Microbiology* **140 (Pt 1)**, 49-57.

Loman, N. J. & Pallen, M. J. (2015). Twenty years of bacterial genome sequencing. *Nat Rev Microbiol* **13**, 787-794.

Lou, H., Chen, M., Black, S. S., Bushell, S. R., Ceccarelli, M., Mach, T., Beis, K., Low, A. S., Bamford, V. A., Booth, I. R., Bayley, H. & Naismith, J. H. (2011). Altered antibiotic transport in OmpC mutants isolated from a series of clinical strains of multi-drug resistant *E. coli*. *PLoS One* **6**, e25825.

Lowe, M. A., Holt, S. C. & Eisenstein, B. I. (1987). Immunoelectron microscopic analysis of elongation of type 1 fimbriae in *Escherichia coli*. *J Bacteriol* **169**, 157-163.

Lugtenberg, E. J. & Peters, R. (1976). Distribution of lipids in cytoplasmic and outer membranes of *Escherichia coli* K12. *Biochim Biophys Acta* **441**, 38-47.

Luirink, J., Yu, Z., Wagner, S. & de Gier, J. W. (2012). Biogenesis of inner membrane proteins in *Escherichia coli*. *Biochim Biophys Acta* **1817**, 965-976.

Luo, C., Hu, G. Q. & Zhu, H. (2009). Genome reannotation of *Escherichia coli* CFT073 with new insights into virulence. *BMC Genomics* **10**, 552.

Maddalo, G., Stenberg-Bruzell, F., Götzke, H., Toddo, S., Björkholm, P., Eriksson, H., Chovanec, P., Genevaux, P., Lehtio, J., Ilag, L. L. & Daley, D. O. (2011). Systematic analysis of native membrane protein complexes in *Escherichia coli*. *J Proteome Res* **10**, 1848-1859.

Maderbocus, R. (2012). A study of outer membrane biogenesis in *Escherichia coli*. In *School of Cancer Sciences*. UK: University of Birmingham.

Magnet, S., Bellais, S., Dubost, L., Fourgeaud, M., Mainardi, J. L., Petit-Frere, S., Marie, A., Mengin-Lecreulx, D., Arthur, M. & Gutmann, L. (2007). Identification of the L,D-transpeptidases responsible for attachment of the Braun lipoprotein to *Escherichia coli* peptidoglycan. *J Bacteriol* **189**, 3927-3931.

Malinverni, J. C. & Silhavy, T. J. (2009). An ABC transport system that maintains lipid asymmetry in the gram-negative outer membrane. *Proc Natl Acad Sci U S A* **106**, 8009-8014.

Malinverni, J. C., Werner, J., Kim, S., Sklar, J. G., Kahne, D., Misra, R. & Silhavy, T. J. (2006). YfiO stabilizes the YaeT complex and is essential for outer membrane protein assembly in *Escherichia coli*. *Mol Microbiol* **61**, 151-164.

Malki, A., Le, H. T., Milles, S., Kern, R., Caldas, T., Abdallah, J. & Richarme, G. (2008). Solubilization of protein aggregates by the acid stress chaperones HdeA and HdeB. *J Biol Chem* **283**, 13679-13687.

Matern, Y., Barion, B. & Behrens-Kneip, S. (2010). PpiD is a player in the network of periplasmic chaperones in Escherichia coli. *BMC Microbiol* **10**, 251.

Maurer, L. & Orndorff, P. E. (1987). Identification and characterization of genes determining receptor binding and pilus length of *Escherichia coli* type 1 pili. *J Bacteriol* **169**, 640-645.

Michel, L. V., Shaw, J., MacPherson, V., Barnard, D., Bettinger, J., D'Arcy, B., Surendran, N., Hellman, J. & Pichichero, M. E. (2015). Dual orientation of the outer membrane lipoprotein Pal in *Escherichia coli*. *Microbiology* **161**, 1251-1259.

Misra, R., Peterson, A., Ferenci, T. & Silhavy, T. J. (1991). A genetic approach for analyzing the pathway of LamB assembly into the outer membrane of *Escherichia coli*. *J Biol Chem* **266**, 13592-13597.

Mobley, H. L., Green, D. M., Trifillis, A. L., Johnson, D. E., Chippendale, G. R., Lockatell, C. V., Jones, B. D. & Warren, J. W. (1990). Pyelonephritogenic *Escherichia coli* and killing of cultured human renal proximal tubular epithelial cells: role of hemolysin in some strains. *Infect Immun* **58**, 1281-1289.

Mohd Yusoff, M. Z., Hashiguchi, Y., Maeda, T. & Wood, T. K. (2013). Four products from *Escherichia coli* pseudogenes increase hydrogen production. *Biochem Biophys Res Commun* **439**, 576-579.

Morein, S., Andersson, A., Rilfors, L. & Lindblom, G. (1996). Wild-type *Escherichia coli* cells regulate the membrane lipid composition in a "window" between gel and non-lamellar structures. *J Biol Chem* **271**, 6801-6809.

Munera, D., Hultgren, S. & Fernández, L. Á. (2007). Recognition of the N-terminal lectin domain of FimH adhesin by the usher FimD is required for type 1 pilus biogenesis. *Mol Microbiol* **64**, 333-346.

Munera, D., Palomino, C. & Fernández, L. Á. (2008). Specific residues in the N-terminal domain of FimH stimulate type 1 fimbriae assembly in *Escherichia coli* following the initial binding of the adhesin to FimD usher. *Mol Microbiol* **69**, 911-925.

Murphy, C. N., Mortensen, M. S., Krogfelt, K. A. & Clegg, S. (2013). Role of *Klebsiella pneumoniae* type 1 and type 3 fimbriae in colonizing silicone tubes implanted into the bladders of mice as a model of catheter-associated urinary tract infections. *Infect Immun* **81**, 3009-3017.

Nakamura, K. & Mizushima, S. (1976). Effects of heating in dodecyl sulfate solution on the conformation and electrophoretic mobility of isolated major outer membrane proteins from *Escherichia coli* K-12. *J Biochem* **80**, 1411-1422.

Narita, S., Masui, C., Suzuki, T., Dohmae, N. & Akiyama, Y. (2013). Protease homolog BepA (YfgC) promotes assembly and degradation of β-barrel membrane proteins in *Escherichia coli*. *Proc Natl Acad Sci U S A* **110**, E3612-3621.

Natale, P., Bruser, T. & Driessen, A. J. (2008). Sec- and Tat-mediated protein secretion across the bacterial cytoplasmic membrane--distinct translocases and mechanisms. *Biochim Biophys Acta* **1778**, 1735-1756.

Needham, B. D. & Trent, M. S. (2013). Fortifying the barrier: the impact of lipid A remodelling on bacterial pathogenesis. *Nat Rev Microbiol* **11**, 467-481.

Nesta, B., Spraggon, G., Alteri, C., Moriel, D. G., Rosini, R., Veggi, D., Smith, S., Bertoldi, I., Pastorello, I., Ferlenghi, I., Fontana, M. R., Frankel, G., Mobley, H. L., Rappuoli, R.,

Pizza, M., Serino, L. & Soriani, M. (2012). FdeC, a novel broadly conserved *Escherichia coli* adhesin eliciting protection against urinary tract infections. *MBio* **3**.

Ng, T. W., Akman, L., Osisami, M. & Thanassi, D. G. (2004). The usher N terminus is the initial targeting site for chaperone-subunit complexes and participates in subsequent pilus biogenesis events. *J Bacteriol* **186**, 5321-5331.

Nguyen, L. T., Gumbart, J. C., Beeby, M. & Jensen, G. J. (2015). Coarse-grained simulations of bacterial cell wall growth reveal that local coordination alone can be sufficient to maintain rod shape. *Proc Natl Acad Sci U S A* **112**, E3689-3698.

Nichols, R. J., Sen, S., Choo, Y. J., Beltrao, P., Zietek, M., Chaba, R., Lee, S., Kazmierczak, K. M., Lee, K. J., Wong, A., Shales, M., Lovett, S., Winkler, M. E., Krogan, N. J., Typas, A. & Gross, C. A. (2011). Phenotypic landscape of a bacterial cell. *Cell* **144**, 143-156.

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

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

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

Noinaj, N., Rollauer, S. E. & Buchanan, S. K. (2015). The β-barrel membrane protein insertase machinery from Gram-negative bacteria. *Curr Opin Struct Biol* **31**, 35-42.

Novak, P. & Dev, I. K. (1988). Degradation of a signal peptide by protease IV and oligopeptidase A. *J Bacteriol* **170**, 5067-5075.

Novak, P., Ray, P. H. & Dev, I. K. (1986). Localization and purification of two enzymes from *Escherichia coli* capable of hydrolyzing a signal peptide. *J Biol Chem* **261**, 420-427.

Nuccio, S. P. & Bäumler, A. J. (2007). Evolution of the chaperone/usher assembly pathway: fimbrial classification goes Greek. *Microbiol Mol Biol Rev* **71**, 551-575.

Oberhettinger, P., Schutz, M., Leo, J. C., Heinz, N., Berger, J., Autenrieth, I. B. & Linke, D. (2012). Intimin and invasin export their C-terminus to the bacterial cell surface using an inverse mechanism compared to classical autotransport. *PLoS One* **7**, e47069.

Obermann, W. & Holtje, J. V. (1994). Alterations of murein structure and of penicillin-binding proteins in minicells from *Escherichia coli*. *Microbiology* **140 (Pt 1)**, 79-87.

Okuda, S., Sherman, D. J., Silhavy, T. J., Ruiz, N. & Kahne, D. (2016). Lipopolysaccharide transport and assembly at the outer membrane: the PEZ model. *Nat Rev Microbiol* **14**, 337-345.

Okuda, S. & Tokuda, H. (2011). Lipoprotein sorting in bacteria. *Annu Rev Microbiol* **65**, 239-259.

Oomen, C. J., van Ulsen, P., van Gelder, P., Feijen, M., Tommassen, J. & Gros, P. (2004). Structure of the translocator domain of a bacterial autotransporter. *EMBO J* **23**, 1257-1266.

Paetzel, M. (2014). Structure and mechanism of *Escherichia coli* type I signal peptidase. *Biochim Biophys Acta* **1843**, 1497-1508.

Pages, J. M. & Bolla, J. M. (1988). Assembly of the OmpF porin of *Escherichia coli* B. Immunological and kinetic studies of the integration pathway. *Eur J Biochem* **176**, 655-660.

Palmer, T. & Berks, B. C. (2012). The twin-arginine translocation (Tat) protein export pathway. *Nat Rev Microbiol* **10**, 483-496.

Palmer, T., Sargent, F. & Berks, B. C. (2005). Export of complex cofactor-containing proteins by the bacterial Tat pathway. *Trends Microbiol* **13**, 175-180.

Palomino, C., Marín, E. & Fernández, L. Á. (2011). The fimbrial usher FimD follows the SurA-BamB pathway for its assembly in the outer membrane of *Escherichia coli*. *J Bacteriol* **193**, 5222-5230.

Paramasivam, N., Habeck, M. & Linke, D. (2012). Is the C-terminal insertional signal in Gram-negative bacterial outer membrane proteins species-specific or not? *BMC Genomics* **13**, 510.

Park, J. S., Lee, W. C., Yeo, K. J., Ryu, K. S., Kumarasiri, M., Hesek, D., Lee, M., Mobashery, S., Song, J. H., Kim, S. I., Lee, J. C., Cheong, C., Jeon, Y. H. & Kim, H. Y. (2012). Mechanism of anchoring of OmpA protein to the cell wall peptidoglycan of the Gram-negative bacterial outer membrane. *FASEB J* **26**, 219-228.

Parsons, J. B. & Rock, C. O. (2013). Bacterial lipids: metabolism and membrane homeostasis. *Prog Lipid Res* **52**, 249-276.

Pasloske, B. L. & Paranchych, W. (1988). The expression of mutant pilins in *Pseudomonas aeruginosa*: fifth position glutamate affects pilin methylation. *Mol Microbiol* **2**, 489-495.

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

Petersen, T. N., Brunak, S., von Heijne, G. & Nielsen, H. (2011). SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nat Methods* **8**, 785-786.

Phan, G., Remaut, H., Wang, T., Allen, W. J., Pirker, K. F., Lebedev, A., Henderson, N. S., Geibel, S., Volkan, E., Yan, J., Kunze, M. B., Pinkner, J. S., Ford, B., Kay, C. W., Li, H., Hultgren, S. J., Thanassi, D. G. & Waksman, G. (2011). Crystal structure of the FimD usher bound to its cognate FimC-FimH substrate. *Nature* **474**, 49-53.

Phu, L., Izrael-Tomasevic, A., Matsumoto, M. L., Bustos, D., Dynek, J. N., Fedorova, A. V., Bakalarski, C. E., Arnott, D., Deshayes, K., Dixit, V. M., Kelley, R. F., Vucic, D. & Kirkpatrick, D. S. (2011). Improved quantitative mass spectrometry methods for characterizing complex ubiquitin signals. *Mol Cell Proteomics* **10**, M110 003756.

Pogliano, K. J. & Beckwith, J. (1994). Genetic and molecular characterization of the *Escherichia coli secD* operon and its products. *J Bacteriol* **176**, 804-814.

Poquet, I., Kornacker, M. G. & Pugsley, A. P. (1993). The role of the lipoprotein sorting signal (aspartate +2) in pullulanase secretion. *Mol Microbiol* **9**, 1061-1069.

Pouttu, R., Westerlund-Wikström, B., Lång, H., Alsti, K., Virkola, R., Saarela, U., Siitonen, A., Kalkkinen, N. & Korhonen, T. K. (2001). *matB*, a common fimbrillin gene of *Escherichia coli*, expressed in a genetically conserved, virulent clonal group. *J Bacteriol* **183**, 4727-4736.

Prehm, P., Stirm, S., Jann, B. & Jann, K. (1975). Cell-wall lipopolysaccharide from *Escherichia coli* B. *Eur J Biochem* **56**, 41-55.

Pucciarelli, M. G. & García-del Portillo, F. (2003). Protein-peptidoglycan interactions modulate the assembly of the needle complex in the *Salmonella* invasion-associated type III secretion system. *Mol Microbiol* **48**, 573-585.

Pugsley, A. P. (1993). The complete general secretory pathway in Gram-negative bacteria. *Microbiol Rev* **57**, 50-108.

Puorger, C., Eidam, O., Capitani, G., Erilov, D., Grütter, M. G. & Glockshuber, R. (2008). Infinite kinetic stability against dissociation of supramolecular protein complexes through donor strand complementation. *Structure* **16**, 631-642.

Qasim, M. A. (2014). Specificity of proteinase K at P2 to P3' sub-sites and its comparison to other serine proteases. *Protein Pept Lett* **21**, 164-170.

Qiao, S., Luo, Q., Zhao, Y., Zhang, X. C. & Huang, Y. (2014). Structural basis for lipopolysaccharide insertion in the bacterial outer membrane. *Nature* **511**, 108-111.

Raetz, C. R., Reynolds, C. M., Trent, M. S. & Bishop, R. E. (2007). Lipid A modification systems in Gram-negative bacteria. *Annu Rev Biochem* **76**, 295-329.

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

Raina, S., Missiakas, D., Baird, L., Kumar, S. & Georgopoulos, C. (1993). Identification and transcriptional analysis of the *Escherichia coli htrE* operon which is homologous to pap and related pilin operons. *J Bacteriol* **175**, 5009-5021.

Raivio, T. L. (2014). Everything old is new again: an update on current research on the Cpx envelope stress response. *Biochim Biophys Acta* **1843**, 1529-1541.

Rassam, P., Copeland, N. A., Birkholz, O., Toth, C., Chavent, M., Duncan, A. L., Cross, S. J., Housden, N. G., Kaminska, R., Seger, U., Quinn, D. M., Garrod, T. J., Sansom, M. S., Piehler, J., Baumann, C. G. & Kleanthous, C. (2015). Supramolecular assemblies underpin turnover of outer membrane proteins in bacteria. *Nature* **523**, 333-336.

Reddie, K. G. & Carroll, K. S. (2008). Expanding the functional diversity of proteins through cysteine oxidation. *Curr Opin Chem Biol* **12**, 746-754.

Reimer, U., Scherer, G., Drewello, M., Kruber, S., Schutkowski, M. & Fischer, G. (1998). Side-chain effects on peptidyl-prolyl *cis*/*trans* isomerisation. *J Mol Biol* **279**, 449-460.

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

Rendón, M. A., Saldaña, Z., Erdem, A. L., Monteiro-Neto, V., Vázquez, A., Kaper, J. B., Puente, J. L. & Girón, J. A. (2007). Commensal and pathogenic *Escherichia coli* use a common pilus

adherence factor for epithelial cell colonization. *Proc Natl Acad Sci U S A* **104**, 10637-10642.

Ried, G., Hindennach, I. & Henning, U. (1990). Role of lipopolysaccharide in assembly of *Escherichia coli* outer membrane proteins OmpA, OmpC, and OmpF. *J Bacteriol* **172**, 6048-6053.

Riley, J. G., Menggad, M., Montoya-Peleaz, P. J., Szarek, W. A., Marolda, C. L., Valvano, M. A., Schutzbach, J. S. & Brockhausen, I. (2005). The *wbbD* gene of *E. coli* strain VW187 (O7:K1) encodes a UDP-Gal: GlcNAcα-pyrophosphate-R β1,3-galactosyltransferase involved in the biosynthesis of O7-specific lipopolysaccharide. *Glycobiology* **15**, 605-613.

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

Robert, V., Volokhina, E. B., Senf, F., Bos, M. P., Van Gelder, P. & Tommassen, J. (2006). Assembly factor Omp85 recognizes its outer membrane protein substrates by a species-specific C-terminal motif. *PLoS Biol* **4**, e377.

Rock, C. O. & Cronan, J. E. (1996). *Escherichia coli* as a model for the regulation of dissociable (type II) fatty acid biosynthesis. *Biochim Biophys Acta* **1302**, 1-16.

Rollauer, S. E., Sooreshjani, M. A., Noinaj, N. & Buchanan, S. K. (2015). Outer membrane protein biogenesis in Gram-negative bacteria. *Philos Trans R Soc Lond B Biol Sci* **370**.

Ronald, A. (2002). The etiology of urinary tract infection: traditional and emerging pathogens. *Am J Med* **113 Suppl 1A**, 14S-19S.

Rondelet, A. & Condemine, G. (2013). Type II secretion: the substrates that won't go away. *Res Microbiol* **164**, 556-561.

Roujeinikova, A., Simon, W. J., Gilroy, J., Rice, D. W., Rafferty, J. B. & Slabas, A. R. (2007). Structural studies of fatty acyl-(acyl carrier protein) thioesters reveal a hydrophobic binding cavity that can expand to fit longer substrates. *J Mol Biol* **365**, 135-145.

Roussel-Jazédé, V., Grijpstra, J., van Dam, V., Tommassen, J. & van Ulsen, P. (2013). Lipidation of the autotransporter NalP of *Neisseria meningitidis* is required for its function in the release of cell-surface-exposed proteins. *Microbiology* **159**, 286-295.

Ruby, E. G., Urbanowski, M., Campbell, J., Dunn, A., Faini, M., Gunsalus, R., Lostroh, P., Lupp, C., McCann, J., Millikan, D., Schaefer, A., Stabb, E., Stevens, A., Visick, K., Whistler, C. & Greenberg, E. P. (2005). Complete genome sequence of *Vibrio fischeri*: a symbiotic bacterium with pathogenic congeners. *Proc Natl Acad Sci U S A* **102**, 3004-3009.

Sachelaru, I., Petriman, N. A., Kudva, R. & Koch, H. G. (2014). Dynamic interaction of the Sec translocon with the chaperone PpiD. *J Biol Chem* **289**, 21706-21715.

Sachelaru, I., Petriman, N. A., Kudva, R., Kuhn, P., Welte, T., Knapp, B., Drepper, F., Warscheid, B. & Koch, H. G. (2013). YidC occupies the lateral gate of the SecYEG translocon and is sequentially displaced by a nascent membrane protein. *J Biol Chem* **288**, 16295-16307.

Saenz, H. L. & Dehio, C. (2005). Signature-tagged mutagenesis: technical advances in a negative selection method for virulence gene identification. *Curr Opin Microbiol* **8**, 612-619.

Saldaña, Z., de la Cruz, M. A., Carrillo-Casas, E. M., Durán, L., Zhang, Y., Hernández-Castro, R., Puente, J. L., Daaka, Y. & Girón, J. A. (2014). Production of the *Escherichia coli* common pilus by uropathogenic *E. coli* is associated with adherence to HeLa and HTB-4 cells and invasion of mouse bladder urothelium. *PLoS One* **9**, e101200.

Sambrook, J. & Russell, D. W. (2006). Agarose Gel Electrophoresis. *Cold Spring Harbor Protocols* **2006**, pdb.prot4020.

Sánchez-Romero, M. A., Cota, I. & Casadesús, J. (2015). DNA methylation in bacteria: from the methyl group to the methylome. *Curr Opin Microbiol* **25**, 9-16.

Sauer, M. M., Jakob, R. P., Eras, J., Baday, S., Eris, D., Navarra, G., Bernèche, S., Ernst, B., Maier, T. & Glockshuber, R. (2016). Catch-bond mechanism of the bacterial adhesin FimH. *Nat Commun* **7**, 10738.

Saulino, E. T., Thanassi, D. G., Pinkner, J. S. & Hultgren, S. J. (1998). Ramifications of kinetic partitioning on usher-mediated pilus biogenesis. *EMBO J* **17**, 2177-2185.

Sauri, A., Soprova, Z., Wickstrom, D., de Gier, J. W., Van der Schors, R. C., Smit, A. B., Jong, W. S. & Luirink, J. (2009). The Bam (Omp85) complex is involved in secretion of the autotransporter haemoglobin protease. *Microbiology* **155**, 3982-3991.

Scheurwater, E. M. & Burrows, L. L. (2011). Maintaining network security: how macromolecular structures cross the peptidoglycan layer. *FEMS Microbiol Lett* **318**, 1-9.

Schmidpeter, P. A. & Schmid, F. X. (2015). Prolyl isomerization and its catalysis in protein folding and protein function. *J Mol Biol* **427**, 1609-1631.

Schneider, D., Duperchy, E., Coursange, E., Lenski, R. E. & Blot, M. (2000). Long-term experimental evolution in *Escherichia coli*. IX. Characterization of insertion sequence-mediated mutations and rearrangements. *Genetics* **156**, 477-488.

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

Schwalm, J., Mahoney, T. F., Soltes, G. R. & Silhavy, T. J. (2013). Role for Skp in LptD assembly in *Escherichia coli*. *J Bacteriol* **195**, 3734-3742.

Sekowska, A., Kung, H. F. & Danchin, A. (2000). Sulfur metabolism in *Escherichia coli* and related bacteria: facts and fiction. *J Mol Microbiol Biotechnol* **2**, 145-177.

Selkrig, J., Belousoff, M. J., Headey, S. J., Heinz, E., Shiota, T., Shen, H. H., Beckham, S. A., Bamert, R. S., Phan, M. D., Schembri, M. A., Wilce, M. C., Scanlon, M. J., Strugnell, R. A. & Lithgow, T. (2015). Conserved features in TamA enable interaction with TamB to drive the activity of the translocation and assembly module. *Sci Rep* **5**, 12905.

Selkrig, J., Leyton, D. L., Webb, C. T. & Lithgow, T. (2014). Assembly of β-barrel proteins into bacterial outer membranes. *Biochim Biophys Acta* **1843**, 1542-1550.

Selkrig, J., Mosbahi, K., Webb, C. T., Belousoff, M. J., Perry, A. J., Wells, T. J., Morris, F., Leyton, D. L., Totsika, M., Phan, M. D., Celik, N., Kelly, M., Oates, C., Hartland, E. L., Robins-Browne, R. M., Ramarathinam, S. H., Purcell, A. W., Schembri, M. A., Strugnell, R. A., Henderson, I. R., Walker, D. & Lithgow, T. (2012). Discovery of an archetypal protein transport system in bacterial outer membranes. *Nat Struct Mol Biol* **19**, 506-510, S501.

Sen, K. & Nikaido, H. (1991). Lipopolysaccharide structure required for *in vitro* trimerization of *Escherichia coli* OmpF porin. *J Bacteriol* **173**, 926-928.

Seydel, A., Gounon, P. & Pugsley, A. P. (1999). Testing the '+2 rule' for lipoprotein sorting in the *Escherichia coli* cell envelope with a new genetic selection. *Mol Microbiol* **34**, 810-821.

Shahid, S. A., Bardiaux, B., Franks, W. T., Krabben, L., Habeck, M., van Rossum, B. J. & Linke, D. (2012). Membrane-protein structure determination by solid-state NMR spectroscopy of microcrystals. *Nat Methods* **9**, 1212-1217.

Sheikh, A., Luo, Q., Roy, K., Shabaan, S., Kumar, P., Qadri, F. & Fleckenstein, J. M. (2014). Contribution of the highly conserved EaeH surface protein to enterotoxigenic Escherichia coli pathogenesis. *Infect Immun* **82**, 3657-3666.

Shen, H. H., Leyton, D. L., Shiota, T., Belousoff, M. J., Noinaj, N., Lu, J., Holt, S. A., Tan, K., Selkrig, J., Webb, C. T., Buchanan, S. K., Martin, L. L. & Lithgow, T. (2014). Reconstitution of a nanomachine driving the assembly of proteins into bacterial outer membranes. *Nat Commun* **5**, 5078.

Shiota, T., Imai, K., Qiu, J., Hewitt, V. L., Tan, K., Shen, H. H., Sakiyama, N., Fukasawa, Y., Hayat, S., Kamiya, M., Elofsson, A., Tomii, K., Horton, P., Wiedemann, N., Pfanner, N., Lithgow, T. & Endo, T. (2015). Molecular architecture of the active mitochondrial protein gate. *Science* **349**, 1544-1548.

Shu, W., Liu, J., Ji, H. & Lu, M. (2000). Core structure of the outer membrane lipoprotein from *Escherichia coli* at 1.9 Å resolution. *J Mol Biol* **299**, 1101-1112.

Signoretto, C., Di Stefano, F. & Canepari, P. (1996). Modified peptidoglycan chemical composition in shape-altered *Escherichia coli*. *Microbiology* **142 (Pt 8)**, 1919-1926.

Sijbrandi, R., Urbanus, M. L., ten Hagen-Jongman, C. M., Bernstein, H. D., Oudega, B., Otto, B. R. & Luirink, J. (2003). Signal recognition particle (SRP)-mediated targeting and Sec-dependent translocation of an extracellular *Escherichia coli* protein. *J Biol Chem* **278**, 4654-4659.

Simmerman, R. F., Dave, A. M. & Bruce, B. D. (2014). Structure and function of POTRA domains of Omp85/TPS superfamily. *Int Rev Cell Mol Biol* **308**, 1-34.

Simpson, R. J. (2006). SDS-PAGE of Proteins. *Cold Spring Harbor Protocols* **2006**, pdb.prot4313.

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

Skorko-Glonek, J., Laskowska, E., Sobiecka-Szkatula, A. & Lipinska, B. (2007). Characterization of the chaperone-like activity of HtrA (DegP) protein from *Escherichia coli* under the conditions of heat shock. *Arch Biochem Biophys* **464**, 80-89.

Smith, K. P., Fields, J. G., Voogt, R. D., Deng, B., Lam, Y. W. & Mintz, K. P. (2015). The cell envelope proteome of *Aggregatibacter actinomycetemcomitans*. *Mol Oral Microbiol* **30**, 97-110.

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

Spiess, C., Beil, A. & Ehrmann, M. (1999). A temperature-dependent switch from chaperone to protease in a widely conserved heat shock protein. *Cell* **97**, 339-347.

Spurbeck, R. R., Stapleton, A. E., Johnson, J. R., Walk, S. T., Hooton, T. M. & Mobley, H. L. (2011). Fimbrial profiles predict virulence of uropathogenic *Escherichia coli* strains: contribution of Ygi and Yad fimbriae. *Infect Immun* **79**, 4753-4763.

Stabb, E. V. & Ruby, E. G. (2003). Contribution of *pilA* to competitive colonization of the squid *Euprymna scolopes* by *Vibrio fischeri*. *Appl Environ Microbiol* **69**, 820-826.

Stegmeier, J. F., Gluck, A., Sukumaran, S., Mantele, W. & Andersen, C. (2007). Characterisation of YtfM, a second member of the Omp85 family in *Escherichia coli*. *Biol Chem* **388**, 37-46.

Stenberg, F., von Heijne, G. & Daley, D. O. (2007). Assembly of the cytochrome bo3 complex. *J Mol Biol* **371**, 765-773.

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

Strom, M. S., Nunn, D. N. & Lory, S. (1993). A single bifunctional enzyme, PilD, catalyzes cleavage and *N*-methylation of proteins belonging to the type IV pilin family. *Proc Natl Acad Sci U S A* **90**, 2404-2408.

Strömberg, N., Marklund, B. I., Lund, B., Ilver, D., Hamers, A., Gaastra, W., Karlsson, K. A. & Normark, S. (1990). Host-specificity of uropathogenic *Escherichia coli* depends on differences in binding specificity to Gal α 1-4Gal-containing isoreceptors. *EMBO J* **9**, 2001-2010.

Struve, C., Forestier, C. & Krogfelt, K. A. (2003). Application of a novel multi-screening signature-tagged mutagenesis assay for identification of *Klebsiella pneumoniae* genes essential in colonization and infection. *Microbiology* **149**, 167-176.

Struyve, M., Moons, M. & Tommassen, J. (1991). Carboxy-terminal phenylalanine is essential for the correct assembly of a bacterial outer membrane protein. *J Mol Biol* **218**, 141-148.

Stubenrauch, C., Belousoff, M. J., Hay, I. D., Shen, H.-H., Lillington, J., Tuck, K. L., Peters, K. M., Phan, M.-D., Lo, A. W., Schembri, M. A., Strugnell, R. A., Waksman, G. & Lithgow, T. (2016). Effective assembly of fimbriae in *Escherichia coli* depends on the translocation assembly module nanomachine. *Nat Microbiol* **1**, 16064.

Studier, F. W., Rosenberg, A. H., Dunn, J. J. & Dubendorff, J. W. (1990). Use of T7 RNA polymerase to direct expression of cloned genes. *Methods Enzymol* **185**, 60-89.

Stull, F., Koldewey, P., Humes, J. R., Radford, S. E. & Bardwell, J. C. (2016). Substrate protein folds while it is bound to the ATP-independent chaperone Spy. *Nat Struct Mol Biol* **23**, 53-58.

Szabady, R. L., Peterson, J. H., Skillman, K. M. & Bernstein, H. D. (2005). An unusual signal peptide facilitates late steps in the biogenesis of a bacterial autotransporter. *Proc Natl Acad Sci U S A* **102**, 221-226.

Tabor, S. & Richardson, C. C. (1985). A bacteriophage T7 RNA polymerase/promoter system for controlled exclusive expression of specific genes. *Proc Natl Acad Sci U S A* **82**, 1074-1078.

Taniguchi, Y., Choi, P. J., Li, G. W., Chen, H., Babu, M., Hearn, J., Emili, A. & Xie, X. S. (2010). Quantifying *E. coli* proteome and transcriptome with single-molecule sensitivity in single cells. *Science* **329**, 533-538.

Tarry, M. J., Schafer, E., Chen, S., Buchanan, G., Greene, N. P., Lea, S. M., Palmer, T., Saibil, H. R. & Berks, B. C. (2009). Structural analysis of substrate binding by the TatBC component of the twin-arginine protein transport system. *Proc Natl Acad Sci U S A* **106**, 13284-13289.

Taubert, J., Hou, B., Risselada, H. J., Mehner, D., Lunsdorf, H., Grubmuller, H. & Bruser, T. (2015). TatBC-independent TatA/Tat substrate interactions contribute to transport efficiency. *PLoS One* **10**, e0119761.

Taymaz-Nikerel, H., Borujeni, A. E., Verheijen, P. J., Heijnen, J. J. & van Gulik, W. M. (2010). Genome-derived minimal metabolic models for *Escherichia coli* MG1655 with estimated *in vivo* respiratory ATP stoichiometry. *Biotechnol Bioeng* **107**, 369-381.

Tefsen, B., Bos, M. P., Beckers, F., Tommassen, J. & de Cock, H. (2005). MsbA is not required for phospholipid transport in *Neisseria meningitidis*. *J Biol Chem* **280**, 35961-35966.

Thanassi, D. G., Bliska, J. B. & Christie, P. J. (2012). Surface organelles assembled by secretion systems of Gram-negative bacteria: diversity in structure and function. *FEMS Microbiol Rev* **36**, 1046-1082.

Thoma, J., Burmann, B. M., Hiller, S. & Muller, D. J. (2015). Impact of holdase chaperones Skp and SurA on the folding of β-barrel outer-membrane proteins. *Nat Struct Mol Biol* **22**, 795-802.

Tokuda, H. & Matsuyama, S. (2004). Sorting of lipoproteins to the outer membrane in *E. coli*. *Biochim Biophys Acta* **1694**, IN1-9.

Toporowski, M. C., Nomellini, J. F., Awram, P. & Smit, J. (2004). Two outer membrane proteins are required for maximal type I secretion of the *Caulobacter crescentus* S-layer protein. *J Bacteriol* **186**, 8000-8009.

Touzé, T., Hayward, R. D., Eswaran, J., Leong, J. M. & Koronakis, V. (2004). Self-association of EPEC intimin mediated by the β-barrel-containing anchor domain: a role in clustering of the Tir receptor. *Mol Microbiol* **51**, 73-87.

Tuomanen, E. & Cozens, R. (1987). Changes in peptidoglycan composition and penicillin-binding proteins in slowly growing *Escherichia coli*. *J Bacteriol* **169**, 5308-5310.

Turner, R. D., Hurd, A. F., Cadby, A., Hobbs, J. K. & Foster, S. J. (2013). Cell wall elongation mode in Gram-negative bacteria is determined by peptidoglycan architecture. *Nat Commun* **4**, 1496.

Tuteja, R. (2005). Type I signal peptidase: an overview. *Arch Biochem Biophys* **441**, 107-111.

Typas, A., Banzhaf, M., CGross, C. A. & Vollmer, W. (2012). From the regulation of peptidoglycan synthesis to bacterial growth and morphology. *Nat Rev Microbiol* **10**, 123-136.

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

van der Woude, M. W. & Bäumler, A. J. (2004). Phase and antigenic variation in bacteria. *Clin Microbiol Rev* **17**, 581-611, table of contents.

Vázquez-Laslop, N., Lee, H., Hu, R. & Neyfakh, A. A. (2001). Molecular sieve mechanism of selective release of cytoplasmic proteins by osmotically shocked *Escherichia coli*. *J Bacteriol* **183**, 2399-2404.

Verma, R., Rojas, T. C., Maluta, R. P., Leite, J. L., da Silva, L. P., Nakazato, G. & Dias da Silveira, W. (2016). Fimbria-encoding gene *yadC* has a pleiotropic effect on several biological characteristics and plays a role in avian pathogenic *Escherichia coli* pathogenicity. *Infect Immun* **84**, 187-193.

Vila-Farrés, X., Ferrer-Navarro, M., Callarisa, A. E., Martí, S., Espinal, P., Gupta, S., Rolain, J. M., Giralt, E. & Vila, J. (2015). Loss of LPS is involved in the virulence and resistance to colistin of colistin-resistant *Acinetobacter nosocomialis* mutants selected in vitro. *J Antimicrob Chemother* **70**, 2981-2986.

Vincent, S., Glauner, B. & Gutmann, L. (1991). Lytic effect of two fluoroquinolones, ofloxacin and pefloxacin, on *Escherichia coli* W7 and its consequences on peptidoglycan composition. *Antimicrob Agents Chemother* **35**, 1381-1385.

Volkan, E., Kalas, V., Pinkner, J. S., Dodson, K. W., Henderson, N. S., Pham, T., Waksman, G., Delcour, A. H., Thanassi, D. G. & Hultgren, S. J. (2013). Molecular basis of usher pore gating in *Escherichia coli* pilus biogenesis. *Proc Natl Acad Sci U S A* **110**, 20741-20746.

Vollmer, W. & Bertsche, U. (2008). Murein (peptidoglycan) structure, architecture and biosynthesis in *Escherichia coli*. *Biochim Biophys Acta* **1778**, 1714-1734.

Vollmer, W. & Seligman, S. J. (2010). Architecture of peptidoglycan: more data and more models. *Trends Microbiol* **18**, 59-66.

Walther, D. M., Papic, D., Bos, M. P., Tommassen, J. & Rapaport, D. (2009). Signals in bacterial β-barrel proteins are functional in eukaryotic cells for targeting to and assembly in mitochondria. *Proc Natl Acad Sci U S A* **106**, 2531-2536.

Wang, P., Shim, E., Cravatt, B., Jacobsen, R., Schoeniger, J., Kim, A. C., Paetzel, M. & Dalbey, R. E. (2008). *Escherichia coli* signal peptide peptidase A is a serine-lysine protease with a lysine recruited to the nonconserved amino-terminal domain in the S49 protease family. *Biochemistry* **47**, 6361-6369.

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

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

Weidel, W., Frank, H. & Martin, H. H. (1960). The rigid layer of the cell wall of *Escherichia coli* strain B. *J Gen Microbiol* **22**, 158-166.

Welch, R. A., Burland, V., Plunkett, G., 3rd, Redford, P., Roesch, P., Rasko, D., Buckles, E. L., Liou, S. R., Boutin, A., Hackett, J., Stroud, D., Mayhew, G. F., Rose, D. J., Zhou, S., Schwartz, D. C., Perna, N. T., Mobley, H. L., Donnenberg, M. S. & Blattner, F. R. (2002). Extensive mosaic structure revealed by the complete genome sequence of uropathogenic Escherichia coli. *Proc Natl Acad Sci U S A* **99**, 17020-17024.

Werneburg, G. T., Henderson, N. S., Portnoy, E. B., Sarowar, S., Hultgren, S. J., Li, H. & Thanassi, D. G. (2015). The pilus usher controls protein interactions via domain masking and is functional as an oligomer. *Nat Struct Mol Biol* **22**, 540-546.

Werner, J., Augustus, A. M. & Misra, R. (2003). Assembly of TolC, a structurally unique and multifunctional outer membrane protein of *Escherichia coli* K-12. *J Bacteriol* **185**, 6540-6547.

Werner, J. & Misra, R. (2005). YaeT (Omp85) affects the assembly of lipid-dependent and lipid-independent outer membrane proteins of *Escherichia coli*. *Mol Microbiol* **57**, 1450-1459.

Whitfield, C. & Trent, M. S. (2014). Biosynthesis and export of bacterial lipopolysaccharides. *Annu Rev Biochem* **83**, 99-128.

Wilson, M. M. & Bernstein, H. D. (2016). Surface-Exposed Lipoproteins: An Emerging Secretion Phenomenon in Gram-Negative Bacteria. *Trends Microbiol* **24**, 198-208.

Wiśniewski, J. R. & Rakus, D. (2014). Multi-enzyme digestion FASP and the 'Total Protein Approach'-based absolute quantification of the *Escherichia coli* proteome. *J Proteomics* **109**, 322-331.

Wittig, I., Braun, H. P. & Schagger, H. (2006). Blue native PAGE. *Nat Protoc* **1**, 418-428.

Workman, P., Heide, K., Giuliano, N., Lee, N., Mar, J., Vuong, P., Bennion, D. & Misra, R. (2012). Genetic, biochemical, and molecular characterization of the polypeptide transport-associated domain of *Escherichia coli* BamA. *J Bacteriol* **194**, 3512-3521.

Wu, E. L., Engstrom, O., Jo, S., Stuhlsatz, D., Yeom, M. S., Klauda, J. B., Widmalm, G. & Im, W. (2013). Molecular dynamics and NMR spectroscopy studies of *E. coli* lipopolysaccharide structure and dynamics. *Biophys J* **105**, 1444-1455.

Wu, T., Malinverni, J., Ruiz, N., Kim, S., Silhavy, T. J. & Kahne, D. (2005). Identification of a multicomponent complex required for outer membrane biogenesis in *Escherichia coli*. *Cell* **121**, 235-245.

Wu, T., McCandlish, A. C., Gronenberg, L. S., Chng, S. S., Silhavy, T. J. & Kahne, D. (2006). Identification of a protein complex that assembles lipopolysaccharide in the outer membrane of *Escherichia coli*. *Proc Natl Acad Sci U S A* **103**, 11754-11759.

Wu, X. R., Sun, T. T. & Medina, J. J. (1996). In vitro binding of type 1-fimbriated Escherichia coli to uroplakins Ia and Ib: relation to urinary tract infections. *Proc Natl Acad Sci U S A* **93**, 9630-9635.

Wurpel, D. J., Beatson, S. A., Totsika, M., Petty, N. K. & Schembri, M. A. (2013). Chaperone-usher fimbriae of *Escherichia coli*. *PLoS One* **8**, e52835.

Xie, K. & Dalbey, R. E. (2008). Inserting proteins into the bacterial cytoplasmic membrane using the Sec and YidC translocases. *Nat Rev Microbiol* **6**, 234-244.

Xie, K., Hessa, T., Seppala, S., Rapp, M., von Heijne, G. & Dalbey, R. E. (2007). Features of transmembrane segments that promote the lateral release from the translocase into the lipid phase. *Biochemistry* **46**, 15153-15161.

Xu, Z., Chen, H. & Zhou, R. (2011). Genome-wide evidence for positive selection and recombination in *Actinobacillus pleuropneumoniae*. *BMC Evol Biol* **11**, 203.

Yamaguchi, K., Yu, F. & Inouye, M. (1988). A single amino acid determinant of the membrane localization of lipoproteins in *E. coli*. *Cell* **53**, 423-432.

Yao, J. & Rock, C. O. (2013). Phosphatidic acid synthesis in bacteria. *Biochim Biophys Acta* **1831**, 495-502.

Zahrl, D., Wagner, M., Bischof, K., Bayer, M., Zavecz, B., Beranek, A., Ruckenstuhl, C., Zarfel, G. E. & Koraimann, G. (2005). Peptidoglycan degradation by specialized lytic transglycosylases associated with type III and type IV secretion systems. *Microbiology* **151**, 3455-3467.

Zavialov, A. V., Tischenko, V. M., Fooks, L. J., Brandsdal, B. O., Aqvist, J., Zav'yalov, V. P., Macintyre, S. & Knight, S. D. (2005). Resolving the energy paradox of chaperone/usher-mediated fibre assembly. *Biochem J* **389**, 685-694.

Zhang, G., Meredith, T. C. & Kahne, D. (2013). On the essentiality of lipopolysaccharide to Gram-negative bacteria. *Curr Opin Microbiol* **16**, 779-785.

Zhang, L. & Foxman, B. (2003). Molecular epidemiology of *Escherichia coli* mediated urinary tract infections. *Front Biosci* **8**, e235-244.

Zhang, M., Lin, S., Song, X., Liu, J., Fu, Y., Ge, X., Fu, X., Chang, Z. & Chen, P. R. (2011). A genetically incorporated crosslinker reveals chaperone cooperation in acid resistance. *Nat Chem Biol* **7**, 671-677.

Zhang, Y. M. & Rock, C. O. (2008). Membrane lipid homeostasis in bacteria. *Nat Rev Microbiol* **6**, 222-233.

Zhou, G., Mo, W. J., Sebbel, P., Min, G., Neubert, T. A., Glockshuber, R., Wu, X. R., Sun, T. T. & Kong, X. P. (2001). Uroplakin Ia is the urothelial receptor for uropathogenic *Escherichia coli*: evidence from *in vitro* FimH binding. *J Cell Sci* **114**, 4095-4103.

Zückert, W. R. (2014). Secretion of bacterial lipoproteins: through the cytoplasmic membrane, the periplasm and beyond. *Biochim Biophys Acta* **1843**, 1509-1516.

Zunino, P., Sosa, V., Allen, A. G., Preston, A., Schlapp, G. & Maskell, D. J. (2003). *Proteus mirabilis* fimbriae (PMF) are important for both bladder and kidney colonization in mice. *Microbiology* **149**, 3231-3237.