



MONASH University

Assembly of multimeric outer membrane proteins  
in *Escherichia coli*

Rhys Alexander Dunstan

BSc (Hons)

A thesis submitted for the degree of Doctor of Philosophy at Monash University in 2016

Department of Biochemistry and Molecular Biology

Department of Microbiology

## **Copyright notice**

© The author 2016. Except as provided in the Copyright Act 1968, this thesis may not be reproduced in any form without the written permission of the author.

## Table of Contents

Abstract.....	i
General Declaration .....	iii
Preface.....	iv
Acknowledgments.....	v
List of Figures .....	vii
List of Tables .....	x
Abbreviations.....	xi
CHAPTER 1: Introduction .....	1
1.1 Composition of Gram-negative bacterial membranes .....	1
1.2 Targeting proteins to the periplasm .....	3
1.3 Lipoprotein biogenesis and targeting to the outer membrane.....	4
1.4 Targeting of $\beta$ -barrel proteins to the outer membrane .....	6
1.5 Assembly of $\beta$ -barrel proteins by the BAM complex.....	9
1.6 The TAM .....	11
1.7 Multimeric bacterial outer membrane proteins.....	15
1.8 Secretins.....	17
1.9 Secretin architecture.....	17
1.10 Secretin domains.....	18

1.11 Piloting and assembly of secretins to the outer membrane.....	20
1.12 Role of accessory factors in the assembly of secretins .....	23
1.13 Does the BAM complex play a role in secretin assembly? .....	25
1.14 Aims of this study .....	26
CHAPTER 2: Identification of a novel pilotin AspS.....	27
2.1 INTRODUCTION .....	27
2.2 METHODS and MATERIALS.....	29
2.2.1 Bioinformatics.....	29
2.2.2 Strains, plasmids and growth conditions .....	30
2.2.3 Generation of isogenic gene deletions .....	32
2.2.4 Cloning of gspD, yacC and yghG (aspS).....	33
2.2.5 Sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE) .....	34
2.2.6 Immunoblotting.....	36
2.2.7 Assays to measure GspD assembly.....	36
2.2.8 Isolation of total membranes.....	37
2.2.9 Sucrose density fractionation.....	38
2.2.10 Secretome analysis.....	38
2.2.11 In-gel digestion of proteins and LC-MS/MS .....	39
2.3 RESULTS .....	40
2.3.1 PulS-OutS family HMM analysis detects YacC in the genome of EPEC E2348/69 .....	40
2.3.2 AspS is required for T2SS function .....	44

2.3.3 Assays to monitor the kinetics of assembly of secretin GspD in EPEC .....	44
2.3.4 Overexpression of AspS pilots GspD to the OM.....	47
2.3.5 The structure of AspS distinguishes it from the PulS-OutS family of proteins and interacts with the S-domain of its cognate secretin .....	49
2.4 DISCUSSION .....	54
2.4.1 Targeting secretins to the outer membrane .....	54
2.4.2 Secretin/Pilotin interactions .....	55
2.4.3 What is the role of YacC? .....	56
CHAPTER 3: Depletion of BamA.....	59
3.1 INTRODUCTION .....	59
3.2 METHODS and MATERIALS.....	60
3.2.1 Strains and Growth Conditions .....	60
3.2.2 Measuring BamA depletion .....	60
3.2.3 Fluorescence Microscopy .....	60
3.2.4 Isolation of total membranes after BamA expression or depletion.....	61
3.2.5 Sucrose density fractionation.....	61
3.2.6 Transmission Electron Microscopy .....	61
3.2.5 SDS-PAGE .....	62
3.2.8 Blue Native-PAGE (BN-PAGE).....	62
3.2.9 Immunoblotting.....	63
3.2.10 BamC Co-immunoprecipitation.....	64
3.2.11 In-gel digestion of proteins and LC-MS/MS .....	65

3.3 RESULTS .....	66
3.3.1 BamA depletion leads to growth defects and membrane damage .....	66
3.3.2 Depletion of BamA inhibits the assembly of $\beta$ -barrel proteins .....	69
3.3.3 BamC and BamD remain modular after BamA depletion.....	73
3.4 DISCUSSION .....	79
3.4.1 Depletion of BamA for outer membrane assembly assays .....	79
3.4.2 Hierarchal assembly of outer membrane proteins .....	80
3.4.3 Modular nature of the BAM complex.....	81
CHAPTER 4: Multimeric outer membrane proteins GspD, Wza and CsgG do not require BamA or TamA for assembly. ....	83
4.1 INTRODUCTION .....	83
4.2 METHODS and MATERIALS.....	85
4.2.1 Strains, Plasmids and Growth Conditions .....	85
4.2.2 Cloning of wza and csgG.....	86
4.2.3 SDS-PAGE .....	87
4.2.4 Immunoblotting.....	87
4.2.5 GspD assembly assays in MC4100A BamA depletion strain.....	87
4.2.6 Wza assembly assays in MC4100A BamA depletion strain.....	88
4.2.7 CsgG assembly assays in MC4100A BamA depletion strain.....	88
4.2.8 Isolation of total membranes.....	89
4.2.9 Sucrose density fractionation.....	89
4.2.10 Urea Extraction of peripherally associated membrane proteins .....	89

4.2.11 Wza complementation assay .....	90
4.2.12 CsgG complementation assay .....	91
4.2.13 AraB structural modelling.....	91
4.2.14 In-gel digestion of proteins and LC-MS/MS .....	91
4.3 RESULTS .....	92
4.3.1 GspD does not require BamA for assembly into bacterial outer membranes.....	92
4.3.2 Wza does not require BamA for assembly into bacterial outer membranes.....	98
4.3.3 CsgG does not require BamA for assembly into bacterial outer membranes .....	105
4.3.4 CsgE and CsgF increase the assembly efficiency of CsgG .....	105
4.4 DISCUSSION .....	112
4.4.1 New architectures for outer membrane proteins .....	112
4.4.2 Targeting of monomers to the outer membrane.....	113
4.4.3 Integration of the secretion pore into the outer membrane .....	114
CHAPTER 5: Conclusion .....	118
REFERENCES: .....	120
<i>APPENDIX A1 – Protein sequences used to build the PulS_OutS like pilotin HMM.....</i>	<i>133</i>
<i>APPENDIX A2 – Protein sequence scores from the Pilotin HMM in EPEC .....</i>	<i>136</i>
<i>APPENDIX A3 – Representative Coomassie stained SDS-PAGE of membrane fractions after BamA depletion and expression of GspD, Wza and CsgG. ....</i>	<i>138</i>

## Abstract

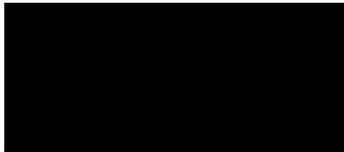
Gram-negative bacterial outer membranes incorporate proteins of at least three well characterized architectures:  $\beta$ -barrel proteins, lipoproteins and multimeric secretion pores. Assembly of  $\beta$ -barrel proteins into the outer membrane is mediated by the  $\beta$ -barrel assembly machinery (BAM), consisting of an essential core  $\beta$ -barrel BamA and four accessory lipoproteins (BamBCDE). Lipoproteins are targeted and subsequently anchored to the outer membrane by covalently attached lipid modifications by the localization of lipoproteins (Lol) machinery. Secretins are amongst the most characterized multimeric outer membrane proteins, and often rely on a small lipoprotein, termed pilotin, to catalyze the efficient translocation of secretin monomers to the outer membrane for assembly into a functional secretion pore. Through the use of cryo-EM and X-ray crystallography it has become clear that oligomeric outer membrane proteins can adopt either  $\alpha$ -helical or  $\beta$ -stranded transmembrane conformations, yet their mechanism of assembly remains unclear.

Chapter two describes the identification and characterisation of a novel pilotin AspS for the secretin GspD from the type 2 secretion system of enteropathogenic *Escherichia coli*. Sucrose density fractionation and novel time course assays show that AspS is required for the targeting of GspD to the outer membrane and for the efficient assembly of GspD multimers. The structure of AspS was solved, demonstrating convergent evolution wherein AspS is functionally equivalent and yet structurally unrelated to pilotins from other secretion systems.

Chapter three describes the optimization of a *bamA* depletion strain of *E. coli* to monitor outer membrane protein assembly. Assays were performed to analyse bacterial viability, morphology and functionality of the BAM after *bamA* depletion. Chapter four focusses on analysing the role of BamA in assembling secretion channels GspD, Wza and CsgG, using the depletion regime described in Chapter three. Assembly assays demonstrate that GspD, Wza and CsgG can assemble into the outer membrane in the absence of BamA. These assays also showed that periplasmic proteins CsgE and CsgF could enhance the assembly efficiency of both the mature CsgG multimer and a possible CsgG assembly intermediate.

## **General 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.



.....

**Rhys Alexander Dunstan**

**19 February 2016**

.....

**Date**

## Preface

- Prof. Roy Robbins-Browne and Dr. Judyta Praszkiar provided wild-type, *ΔgspD* and *ΔsslE* strains of enteropathogenic *Escherichia coli* and pJP117, pJP181 and pJP133 plasmids.
- A/Prof. Konstantin Korotkov and Tim Evans expressed, purified and determined the crystal structure of AspS from *Vibrio cholerae*, created the structural alignment between AspS and PA3611, and performed the AspS MBP-S binding experiment.
- Prof. Tony Purcell, Dr. Ralf Schittenhelm and Sri Ramarathinam analysed samples by mass spectrometry.
- Dr. Georg Ramm, Adam Costin and Joan Clark assisted with the imaging of samples by electron microscopy.
- Dr. Eva Heinz in the Lithgow laboratory performed the PulS\_OutS HMM in *Vibrio cholerae* and CLANS analysis of PulS\_OutS and AspS pilotins.
- Dr. Iain Hay in the Lithgow laboratory performed the colanic acid assay and constructed the CsgG, CsgE and CsgF expressing plasmids.

All other experimentation comprises my original work.

Some of the work undertaken towards this degree has been published in the papers listed below.

**Dunstan RA**, Heinz E, Wijeyewickrema LC, Pike RN, Purcell AW, Evans TJ, Praszkiar J, Robins-Browne RM, Strugnell RA, Korotkov KV and Lithgow T (2013), Assembly of the type II secretion system such as found in *Vibrio cholerae* depends on the novel pilotin AspS. *PLoS Pathogens* 9(1): e1003117. doi: 10.1371/journal.ppat.1003117

**Dunstan RA**, Hay ID, Wilksch JJ, Schittenhelm RB, Purcell AW, Clark J, Costin A, Ramm G, Strugnell RA and 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. *Molecular Microbiology*, 97: 616–629. doi:10.1111/mmi.13055

## Acknowledgments

First I would like to thank my fantastic supervisor Prof. Trevor Lithgow. I will be forever grateful for the opportunities that you have given me during my PhD period. I would not have been able to achieve this without your constant positivity and support. You gave me the freedom to expand on my own ideas and guidance when I was lost or unsure. I will strive to emulate your positive, calm and creative nature throughout the rest of my life.

To my great collaborators, thank you for all of your help during my PhD. I have learnt so much from all of you. To Prof. Roy Robins-Browne and Dr. Judy Praszker, thank you for allowing me to work in your lab and teaching me bacterial genetics. To A/Prof. Konstantin Korotkov thank you for your input and insight into the AspS work. To Prof. Tony Purcell and your research team, thank you for performing the mass spectrometry analysis of our samples. To Dr. Georg Ramm and your research team, thank you for helping me with the electron microscopy.

To all of the members of Lithgow lab, thank you for making my lab experience a fantastic one! I wish you all the very best in your future endeavours. Special thanks to Eva Heinz and Iain Hay. You have both played integral roles in different aspects of my PhD studies. To Denisse Leyton, Tara Wilson and Sam Palframan thank you for being my rock in the lab, you have always been there if I've ever needed someone to talk to. To Chaille Webb, Matt Belousoff and Chris Stubenrauch, thank you for all of the advice and support over the years. To Mary Speir, Victoria Hewitt, Takuya Shiota, Kher Shing Tan, Seong Hoong Chow,

Pankaj Deo and Hsin-Hui Shen, I'll never forget the squash and badminton sessions. To Julie Nguyen, Von Torres, Jonny Kuo and Jess Wisniewski, thanks for the many memorable games nights and I promise to not fall asleep at the next event. Maybe, just maybe, I'll join you at karaoke but you'll have to sweeten the deal.

To all of my wonderful friends outside of the lab, Matt Lee, Shannon Lee, Chris Lee, Melissa Lee, Tom Hyde, Teigen Bywater, Shane Landry, Daniel Guthrie, Katie Auchettl, Andy Laing, Kate Nowak and Jenny Flores. I couldn't ask for a better group of friends. You have all been there for me in one way or another. Thanks for always checking up on me if I've been absent or looked stressed. I hope there will be many more board game sessions, camping trips and overall good times in the future!

Finally to my family, words cannot describe how thankful I am to you. Mum, Dad, Erin and Bridget thank you for your everlasting love and support. You have supported me in everything I have ever wanted to do. If there was ever a time where I needed help, however big or small, you've been there and found a way to fix it. You've provided a loving environment for me to come home to and for that I am eternally grateful. To my extended family, especially my grandparents thank you for everything that you have done for me in my life. You have been a never ending stream of love, support and advice. I love you all!

## List of Figures

### Chapter 1:

Figure 1.1: The Gram-negative cell envelope

Figure 1.2: Translocation of lipoproteins to the outer membrane by the Lol pathway

Figure 1.3: Assembly of  $\beta$ -barrel proteins by the BAM complex

Figure 1.4: Assembly of an autotransporter by the BAM and TAM complexes

Figure 1.5: Structure of Omp85 proteins BamA and TamA

Figure 1.6: Structures of multimeric secretion pores Wza, CsgG and GspD

Figure 1.7: Domain organisation of secretins

Figure 1.8: Structures of the different pilotin classes

Figure 1.9: “Piggyback” model of secretin translocation to the outer membrane

### Chapter 2:

Figure 2.1: YacC is a novel member of the PulS-OutS family of proteins

Figure 2.2: Sequence characteristics of AspS and gene synteny of the T2SS in EPEC, *Shigella boydii* and *Vibrio cholerae*

Figure 2.3: AspS belongs to a novel group of pilotins and is required for T2SS function

Figure 2.4: GspD-C<sub>4</sub> is functional and is able to secrete SsIE

Figure 2.5: Assay to measure the *in vivo* assembly of GspD in EPEC

Figure 2.6: GspD requires AspS for targeting to the outer membrane

Figure 2.7: AspS provides a novel structure for pilotin like proteins

Figure 2.8: The S-domain is required for AspS dependant targeting of GspD to the outer membrane

Figure 2.9: Structural homology between AspS and PA3611

### Chapter 3:

Figure 3.1: Depletion of BamA

Figure 3.2: Depletion of BamA leads to a loss in cell viability

Figure 3.3: Cell morphology of BamA depleted *E. coli*

Figure 3.4: Membrane proteome of BamA depleted *E. coli*

Figure 3.5: Model of TamA assembly into the outer membrane

Figure 3.6: Bam lipoproteins are still targeted to the outer membrane after BamA depletion

Figure 3.7: BamC and BamD remain modular after BamA depletion

Figure 3.8: Modular nature of the BAM complex

Figure 3.9: Depletion of BamA does not inhibit the interaction between BamC and BamD

### Chapter 4:

Figure 4.1: GspD can assemble into multimers independent of BamA

Figure 4.2: Structural model of AraB showing Cysteines that may interact with the Lumio Green dye

Figure 4.3: GspD is targeted to the outer membrane after BamA depletion

Figure 4.4: GspD is integrated to the outer membrane after BamA depletion

Figure 4.5: Wza-C<sub>4</sub> is functional and able to secrete EPS

Figure 4.6: Wza does not require BamA for assembly into multimers

Figure 4.7: Wza is targeted to the outer membrane after BamA depletion

Figure 4.8: Wza is integrated to the outer membrane after BamA depletion

Figure 4.9: Assembly of a non-acylated chimera of Wza does not require BamA

Figure 4.10: CsgG-C4 is functional allowing the secretion of curli fibers

Figure 4.11: CsgG can assemble into multimers after BamA depletion

Figure 4.12: CsgG is targeted to the outer membrane after BamA depletion

Figure 4.13: CsgG is integrated to the outer membrane after BamA depletion

Figure 4.14: CsgE and CsgF both increase the assembly efficiency of CsgG into multimers

Figure 4.15: A model for the assembly of homo-oligomeric secretion pores

## List of Tables

Table 2.1: Strains and Plasmids used in Chapter 2

Table 2.2: Primers used to generate *gspD*, *yacC* and *yghG* isogenic gene deletions in EPEC E2348/69 and BL21 (DE3)

Table 2.3: Primers used to construct the GspD, YacC and YghG (AspS) expressing plasmids

Table 2.4: Solutions required to make 1 x 10% gradient SDS-PAGE gel

Table 2.5: Solutions required to make 1 x 3-14% gradient SDS-PAGE gel

Table 3.1: Solutions required to make 1 x 5-16% gradient BN-PAGE gel

Table 3.2: Antibodies used in Chapter 3

Table 4.1: Strains and Plasmids used in Chapter 4

Table 4.2: Primers used to construct the Wza expressing plasmids

## Abbreviations

AhT	anhydrotetracycline
APS	ammonium persulfate
AspS	alternate general secretion protein subunit S
BAM	$\beta$ -barrel assembly machinery
BamA+	<i>bamA</i> expressing conditions
BamA-	<i>bamA</i> depletion conditions
BN	blue native
CAYE	casamino acid yeast extract salts
CLANS	cluster analysis of sequences
DDM	<i>N</i> -dodecyl- $\beta$ -maltoside
ECL	enhanced chemiluminescence
EDTA	ethylenediaminetetraacetic acid
EPEC	enteropathogenic <i>Escherichia coli</i>
HMM	hidden Markov model
IM	inner membrane
IP	immunoprecipitation
IPTG	isopropyl-beta-D-thiogalactopyranoside
kDA	kilo Dalton
LB	Luria broth
Lol	localization of lipoproteins
LPS	lipopolysaccharide
MCS	multiple cloning site
min	minute
OD	optical density
OM	outer membrane
OMPs	integral $\beta$ -barrel outer membrane proteins
P	pellet

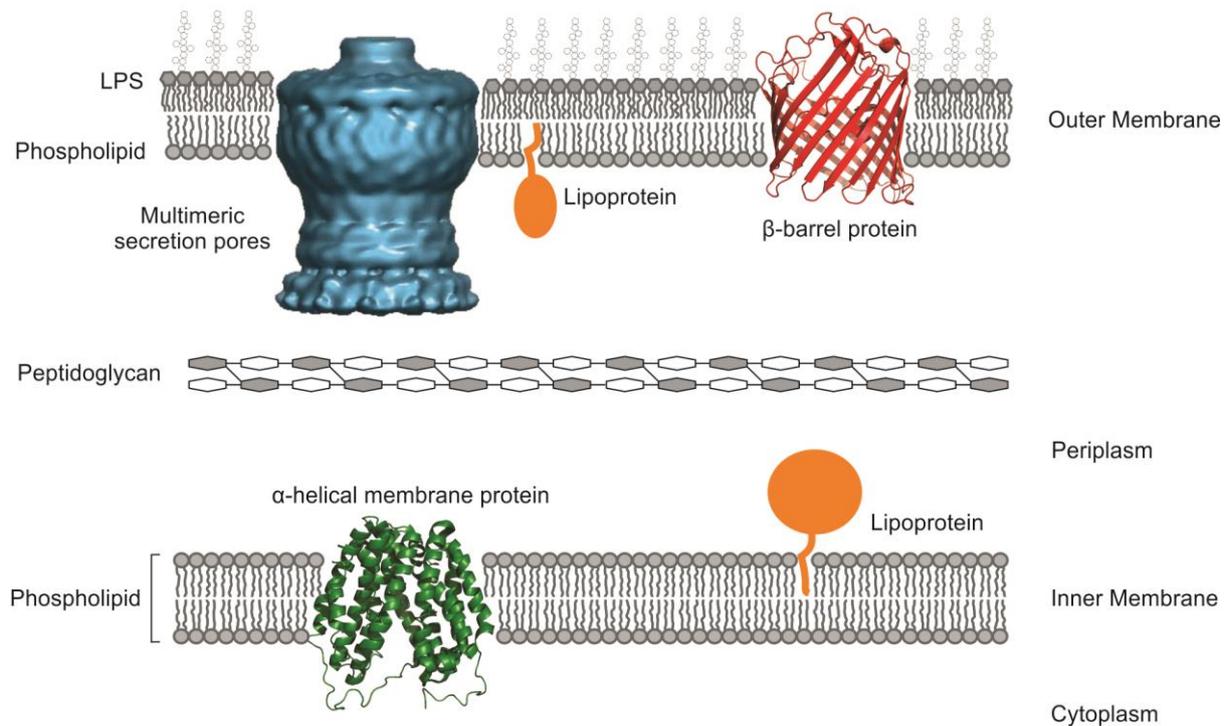
PAGE	polyacrylamide gel electrophoresis
PBS	phosphate buffered saline
PCR	polymerase chain reaction
PMSF	phenylmethanesulfonyl fluoride
POTRA	polypeptide transport-associated
rpm	revolutions per minute
SDS	sodium dodecyl sulphate
S	supernatant
SS	signal sequence
T	total
TAM	translocation and assembly module
TCA	trichloroacetic acid
TEMED	tetramethylethylenediamine
TM	total membranes
T2SS	type two secretion system
T3SS	type three secretion system
T4P	type four pili
° C	degrees Celsius
x g	times gravity

# CHAPTER 1: Introduction

## *1.1 Composition of Gram-negative bacterial membranes*

The bacterial outer membrane provides a permeability barrier to guard against environmental stress ranging from physical factors to antibiotics. Because Gram-negative bacteria are surrounded by two lipid bilayers, they have two aqueous compartments: the cytoplasm and periplasm [1]. The cytoplasm is encapsulated by the inner membrane and the periplasm containing the peptidoglycan layer separates the inner and outer membranes. Both membranes are rich in phosphatidylethanolamine, phosphatidylglycerol and cardiolipin [2] but differ in their symmetry and protein composition [1,3].

The inner membrane adopts a symmetrical phospholipid bilayer which houses integral membrane proteins of  $\alpha$ -helical topology, as well as lipoproteins attached to the periplasmic interface of the membrane [3,4]. In contrast to the inner membrane, the outer membrane is asymmetrical comprising of a phospholipid inner leaflet and a lipopolysaccharide (LPS) outer leaflet, and contains proteins of  $\beta$ -barrel topology, lipoproteins and large multimeric pores (Figure 1.1) [5,6]. The LPS layer assists in the stabilization and barrier function of the outer membrane by reducing the diffusion of hydrophobic molecules into the bacterial cell [7]. The outer membrane also acts as a molecular sieve allowing the transport of small hydrophilic molecules via porins embedded in the outer membrane [7]. The precise targeting and assembly of each of the membrane components is vital for bacterial cell viability, maintaining membrane potential, ensuring the uptake of nutrients into the bacterial cell and allowing for the export of proteins and carbohydrates into the extracellular milieu.



**Figure 1.1: The Gram-negative cell envelope.** The cell envelope of Gram-negative bacteria consists of two membranes of differing lipid and protein composition. The inner membrane adopts a symmetrical phospholipid bilayer which houses integral membrane proteins of  $\alpha$ -helical topology, as well as lipoproteins attached to the periplasmic interface of the membrane. The outer membrane is asymmetrical comprising of phospholipid and lipopolysaccharide (LPS), and contains proteins of  $\beta$ -barrel topology, lipoproteins and large multimeric pores.

## ***1.2 Targeting proteins to the periplasm***

Periplasmic and outer membrane proteins are synthesized in the cytoplasm as a preprotein with a cleavable N-terminal signal peptide that targets them to specific translocation machineries of the inner membrane. There are two translocation machines at the inner membrane, the Sec translocon and the TatABC complex. The SecYEG channel of the translocon facilitates the transport of unfolded proteins into and across the inner membrane in an ATP dependent manner [8,9]. The Sec translocon consists mainly of 7 proteins, including ATPase (SecA), chaperone (SecB), the membrane translocation channel (SecYEG) and a second membrane complex that aids in the efficiency of protein transport (SecDF). YidC and YajC are two other membrane proteins that can also interact with the Sec machinery and play a role protein insertion and translocation of target substrates [8,9].

In *E. coli*, the twin arginine translocase (Tat) is comprised of three key subunits: TatA, TatB and TatC, and facilitates the transport of folded substrate into the periplasm [10]. Tat substrates are targeted to the TatABC complex by virtue of an N-terminal signal sequence that contains the S-R-R-x-F-L-K sequence (with x being a polar amino acid), a hydrophobic central domain and positively charged C-terminal domain. Studies have shown that TatBC acts as the receptor binding complex binding the signal peptide of target substrates and TatA oligomerizes to form the translocation pore. The translocase is unusual in that it forms complexes of differing protomeric ratios to accommodate the transport of folded proteins of varying sizes (~25-70Å) across the inner membrane [10]. Assembly of the TatABC complex is dynamic and occurs on demand in the presence of the target substrates.

### ***1.3 Lipoprotein biogenesis and targeting to the outer membrane***

Lipoproteins are a diverse group of proteins that perform important cellular functions, some of which include the biogenesis and maintenance of the cell envelope, transport of macromolecules, drug efflux, protein assembly and bacterial pathogenesis [11-15]. The function of the majority of lipoproteins from *E. coli* are unknown, although three outer membrane lipoproteins (BamD, LptE and LolB) have been reported to be essential and all play a role in different aspects of outer membrane biogenesis [11-13].

Current estimates are that *E. coli* expresses at least 90 different lipoproteins, the majority of which are predicted to reside in the inner leaflet of the outer membrane and others on the periplasmic face of the inner membrane [16]. However it has been reported that some lipoproteins can adopt altered topologies in the bacterial outer membrane. This was first observed for the lipoprotein PulA from *Klebsiella oxytoca* [17,18]. It has become clear that PulA is secreted onto the cell surface by the type 2 secretion system (T2SS) but it is still unknown how PulA and similar secreted lipoproteins are anchored onto the surface of the outer membrane. Additionally, some lipoproteins, Wza and CsgG for example, adopt a more complicated topology in which the N-terminus is attached to the inner leaflet of the outer membrane and the C-terminal segment of the protein forms a translocation channel for the secretion of polysaccharide and curli fibres respectively to the cell surface [19-21].

Lipoproteins are synthesised in the cytosol as precursors and are targeted to the Sec translocon via a characteristic signal peptide referred to as a lipobox, L-A/S-G/A-C, with the N-terminal cysteine being critical for the biogenesis of lipoproteins. The maturation of lipoproteins takes place on the periplasmic side of the inner membrane in three critical steps

[22]. In the first step, phosphatidylglycerol/prolipoprotein diacylglyceryl transferase (Lgt) adds diacylglycerol to the N-terminal cysteine residue of the lipobox motif via a thioether linkage. The second step involves the cleavage of the signal peptide by signal peptidase II (LspA). Finally, the N-terminal Cys residue is acylated by phospholipid/apolipoprotein transacylase (Lnt) which allows the protein to be anchored into the inner or outer membrane.

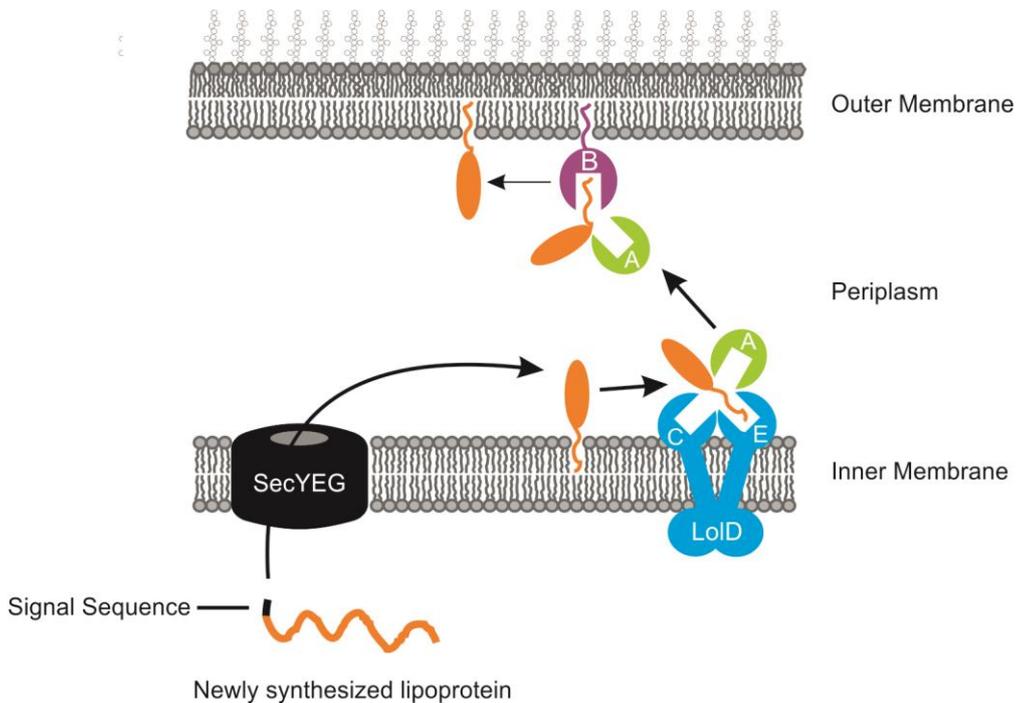
Once processed, lipoproteins are sorted to the outer membrane by the localization of lipoprotein (Lol) system, or else they will remain at the inner membrane after translocation into the periplasm [4]. This dichotomy in the destination of lipoproteins within the cell envelope is determined based on the sequence of amino acids directly adjacent to the acylated cysteine residue (position +1). Yamaguchi et al., (1988) first reported the importance of the amino acid Ser at position +2 for the targeting of lipoproteins to the outer membrane [23]. The substitution of Ser to Asp at position +2 served as a Lol avoidance signal resulting in the lipoprotein remaining at the inner membrane [23]. It is suggested that the retention of lipoproteins with Asp at position +2 is due to the increased interaction of the Asp with phospholipid blocking the recognition of the N-terminal cysteine and extraction of the lipoprotein by the LolCDE complex [24,25]. Moreover, the presence of Gly, Pro, Phe, Trp or Tyr at this position and Asn at the +3 position resulted in the retention of a lipidated maltose-binding protein (MalE) [26].

The Lol machinery comprises of an ATP-binding cassette (ABC) transporter LolCDE at the inner membrane, a periplasmic chaperone LolA and an outer membrane receptor LolB [4]. The first step in the translocation process by the Lol pathway is the interaction of the target substrate with the LolCDE complex at the periplasmic face of the inner membrane. The

binding of lipoprotein substrates to the LolCDE complex increases the specific interaction of LolA to LolC [27]. The hydrolysis of ATP by LolD weakens the interactions of the LolCDE machinery and facilitates the release of the N-terminal lipid moiety of the lipoprotein from the outer leaflet of the inner membrane [28]. The lipoprotein is then transported across the periplasm via its lipid moiety by the carrier protein LolA [29]. Once shuttled to the outer membrane, the LolA-lipoprotein complex interacts with its lipoprotein receptor LolB [13], whereby the lipoprotein is transferred by its lipid moiety from LolA to LolB which displays a higher affinity for the lipid [27]. Once transferred to LolB, the lipoprotein is localized to the outer membrane via the insertion of the N-terminal lipid moiety in the outer membrane [30] (Figure 1.2).

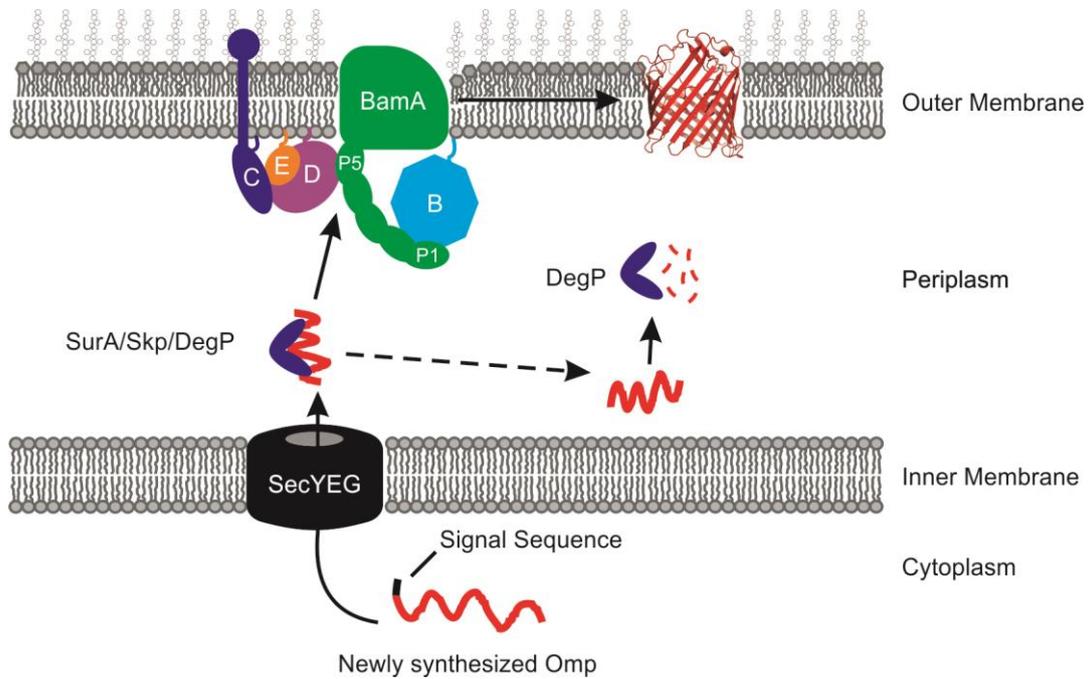
#### ***1.4 Targeting of $\beta$ -barrel proteins to the outer membrane***

Like lipoproteins, integral  $\beta$ -barrel outer membrane proteins (OMPs) are also synthesized as unfolded proteins in the cytoplasm with a cleavable signal sequence that targets them to the Sec translocon for translocation across the inner membrane. Molecular chaperones have been shown to bind nascent OMPs upon transport through the Sec translocon [31,32]. Chaperones, such as SurA and Skp, then aid in the transport of unfolded  $\beta$ -barrels across the periplasm to the outer membrane, where they are assembled by the  $\beta$ -barrel assembly machinery (BAM) complex (Figure 1.3). It is becoming clear that the recognition of target substrates by the BAM complex might be mediated by the binding of BamA to a C-terminal motif referred to as the  $\beta$ -signal [33]. The  $\beta$ -signal contains a highly conserved C-terminal phenylalanine residue that is critical for the assembly of OMPs *in vivo* [34,35], yet the additional features in the  $\beta$ -signal remain obscure.



**Figure 1.2: Translocation of lipoproteins to the outer membrane by the Lol pathway.**

Once processed, lipoproteins destined to the outer membrane are sorted by the localization of lipoprotein (Lol) system. The Lol machinery comprises of an ATP-binding cassette (ABC) transporter LolCDE at the inner membrane, a periplasmic chaperone LolA and an outer membrane receptor LolB. Lipoproteins are first localized to LolE and then transferred to LolA in an ATP dependant manner. The lipoprotein is then transported across the periplasm via its lipid moiety by the carrier protein LolA and subsequently transferred to its receptor protein LolB. The last step involves the insertion of the N-terminal lipid moiety of the lipoprotein into the outer membrane.



**Figure 1.3: Assembly of  $\beta$ -barrel proteins by the BAM complex.** Integral  $\beta$ -barrel membrane proteins (OMPs) are synthesized as an unfolded pre-protein in the cytoplasm with an N-terminal signal sequence for targeting to the SecYEG translocon. Upon transport through the Sec, the signal sequence is cleaved and periplasmic chaperones SurA, Skp or DegP bind the processed OMP for targeting to the  $\beta$ -barrel assembly machinery (BAM) complex. It has been reported that OMPs interact with the periplasmic N-terminal polypeptide transport-associated (POTRA) domains of BamA but the molecular details of barrel assembly remains unknown. Errant proteins that fail to engage periplasmic chaperones are degraded by proteases such as DegP.

Upon translocation into the periplasm via the SecYEG channel, the nascent OMP associates with periplasmic chaperones, such as SurA, Skp and DegP. These chaperones are thought to form two distinct but partially redundant pathways where SurA functions in one pathway, and Skp and DegP in the other [36]. This model is based on the observation that the loss of individual pathways is tolerated by the cell but the simultaneous loss of both pathways results in a synthetic lethal phenotype. It has been suggested that SurA acts as the major chaperone responsible for the biogenesis of most OMPs, with Skp and DegP functioning to rescue OMPs that have failed to engage the SurA pathway [37].

Genetic data has shown that deletion or depletion of *surA* or *skp* results in the loss of OMPs in the outer membrane and the accumulation of misfolded OMPs in the periplasm [37-39]. It has become clear that SurA and Skp can each interact with BamA, highlighting their importance in the correct targeting of OMPs to the outer membrane for assembly [37,40,41]. On the other hand, the chaperone function of DegP is still unclear with studies suggesting that it primarily functions as a protease to degrade unfolded OMPs in the periplasm [42,43].

### ***1.5 Assembly of $\beta$ -barrel proteins by the BAM complex***

Voulhoux *et al.*, (2003) showed for the first time that the assembly of OMPs into the outer membrane relied on the essential outer membrane protein 85 (Omp85) in *Neisseria meningitidis* and that depletion Omp85 (now known as BamA) lead to the accumulation OMPs in the periplasm [44]. BamA forms the essential core  $\beta$ -barrel of the BAM complex, with four accessory lipoproteins BamB, BamC, BamD and BamE. BamA belongs to the Omp85 family of proteins which are defined by a “D15 domain” or “bacterial surface antigen domain” and it has recently become clear that this family of proteins can be split into 10

subfamilies, based on distinct domains attached to the Omp85 barrel [45]. Omp85 homologues have also been identified in eukaryotic organelles derived from endosymbiotic bacteria, such as Sam50 in mitochondria and Toc75 or Oep80 in chloroplasts [46,47]. In addition to the integral  $\beta$ -barrel domain, BamA contains five N-terminal polypeptide transport-associated (POTRA) domains which are exposed to the periplasm [48]. The POTRA domains of BamA have been shown to interact with the BAM accessory lipoproteins and substrate proteins [49,50].

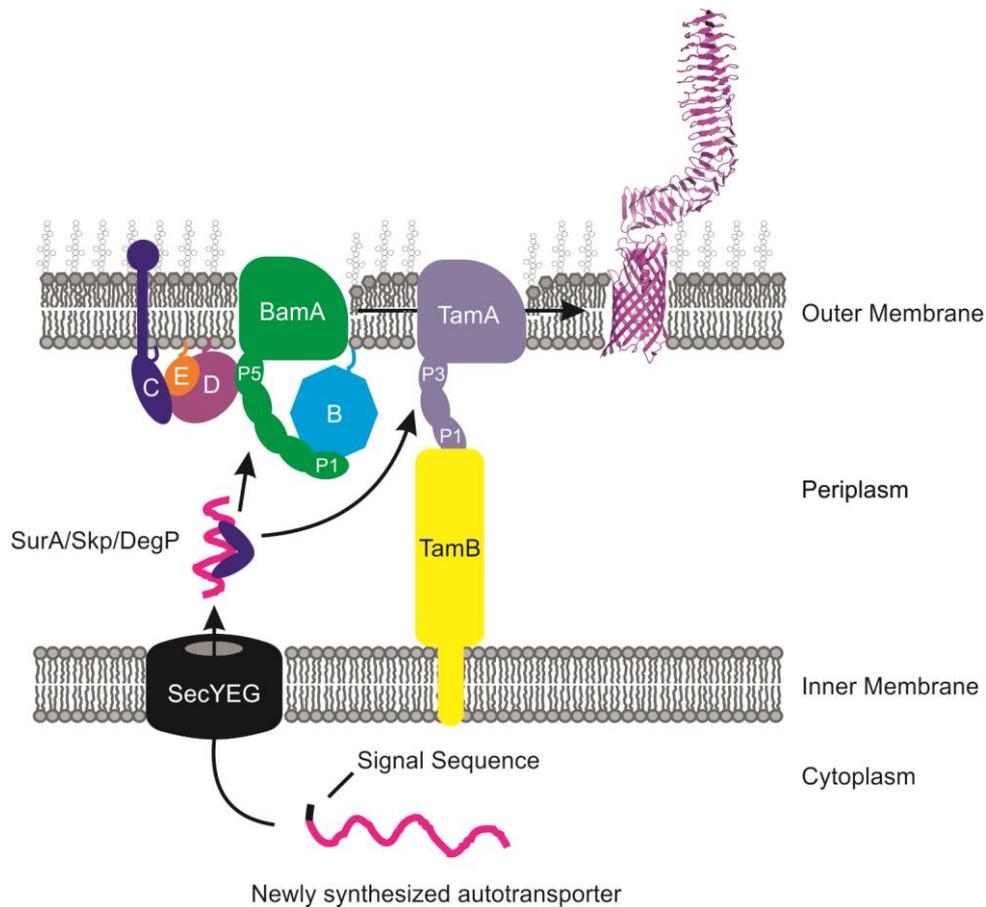
In *E. coli*, BamA interacts with 4 accessory lipoproteins BamB, BamC, BamD and BamE. The BAM lipoproteins are anchored to the inner leaflet of the outer membrane and form modules that interact with the POTRA domains of BamA. Investigations into the molecular architecture of the BAM complex revealed that it comprises of the BamAB and BamCDE modules. The BamCDE module interacts with POTRA 5 of BamA via BamD. BamB on the other hand interacts with BamA at POTRA 1 and 2. These modules can be reconstituted into liposomes from purified components [51]. The biochemically defined functions of the BAM lipoproteins remain unknown. Genetic studies have shown that *bamD* is essential [52,53] and that the deletion of any two non-essential lipoproteins (BamBCE) resulted in a synthetic or conditional lethal phenotype [11,54,55]. Additionally, *E. coli* cells containing single *bamB*, *bamC* or *bamE* gene deletions are viable but show a variety  $\beta$ -barrel assembly defects [54,56,57].

Recent structural studies revealed that BamA forms a 16 stranded  $\beta$ -barrel in which  $\beta$ -strands 1-4 and 16 are significantly shorter than the other strands [58,59] (Figure 1.5). Molecular dynamics suggested that these shorter  $\beta$ -strands would cause a distortion in the outer

membrane resulting in reduced lipid order and thickness in which the authors speculate could prime the outer membrane for OMP insertion [58]. Disruption of the membrane was predicted to occur between the first and last strands of the BamA barrel which are associated by weak hydrogen bonds. Cysteine crosslinking to ‘close’ the lateral opening of BamA caused it to become non-functional [60]. Further analysis revealed the presence of a putative exit pore of BamA between extracellular loops 1-3 and loop 6 which is found just above the lateral gate [59,60]. It is speculated that BamA assembles OMPs via the lateral gate of BamA by inserting  $\beta$ -strands directly into the outer membrane, by a process that may involve strand templating and  $\beta$ -augmentation [60].

### ***1.6 The TAM***

The translocation and assembly module (TAM) is comprised of two integral membrane proteins, the outer membrane protein TamA, and inner membrane protein TamB, which together span the bacterial periplasm [61]. Both TamA and TamB are broadly conserved in proteobacteria [45,62]. In *E. coli*, the genes encoding *tamA* and *tamB* are located in an operon encoding a third gene *ytfP*, but the function of YtfP is unknown. Unlike *bamA*, the deletion of the *tamA* gene does not result in cell death and viable *tamA* deletion strains have been engineered. The TAM was shown to facilitate the insertion of an autotransporter into liposomes or reconstituted planar membranes *in vitro*, and cells lacking TamA show a reduction in the steady state levels of some autotransporters *in vivo* and virulence of pathogens from genera including *Klebsiella*, *Proteus*, *Citrobacter* and *Salmonella* [61,63-65] (Figure 1.4).

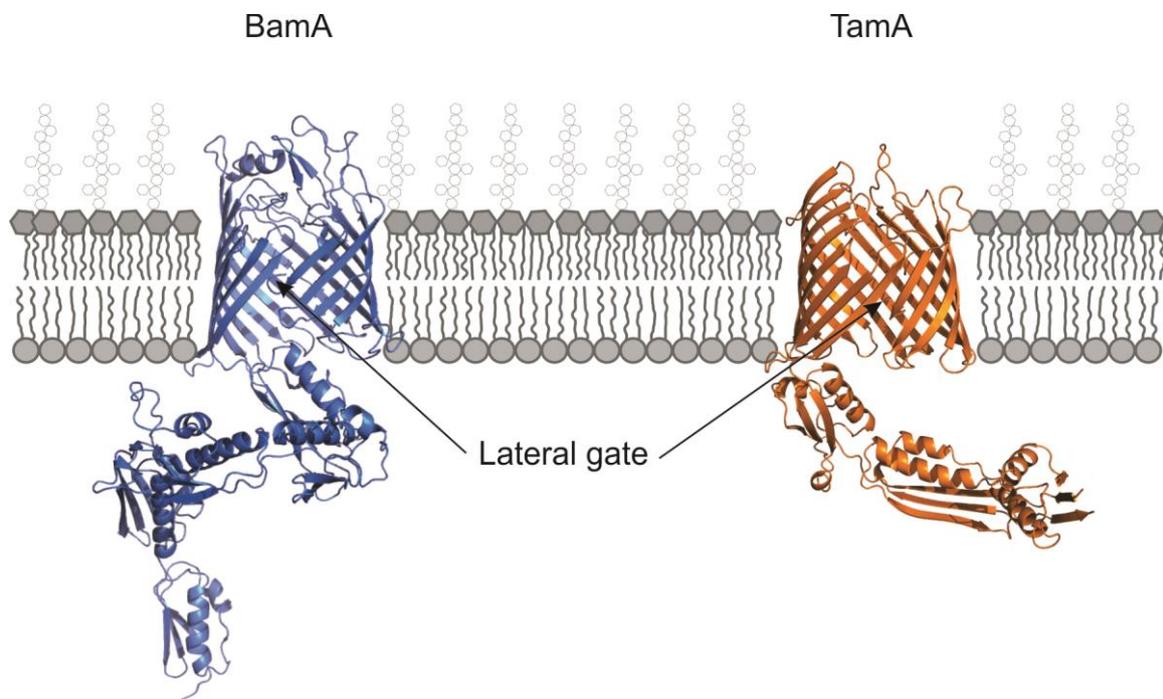


**Figure 1.4: Assembly of an autotransporter by the BAM and TAM complexes.**

Autotransporters are an important class of integral outer membrane proteins found in Gram-negative bacteria and comprise an N-terminal passenger domain which is translocated across the outer membrane and a C-terminal  $\beta$ -barrel. Autotransporters like other OMPs are synthesized as an unfolded pre-protein in the cytoplasm with an N-terminal signal sequence for targeting to the SecYEG translocon. Upon transport through the Sec, the signal sequence is cleaved and periplasmic chaperones such as SurA bind the processed autotransporter for targeting to the BAM complex for assembly of the C-terminal  $\beta$ -barrel. The TAM has also been implicated in the biogenesis of autotransporters but the molecular details of how TAM assembles autotransporters remains unclear.

Like BamA, TamA also belongs to the Omp85 family of proteins. However, TamA can be further distinguished from BamA and other Omp85 members due to the characteristics of its three POTRA domains [66]. It has been shown that POTRA 3 of TamA resembles the POTRA's of BamA and the other two POTRA domains of TamA differ significantly in their sequence and structural properties. A combination of biophysical and biochemical analysis has demonstrated that POTRA 1 is crucial for the interaction with TamB and loss of POTRA 1 results in the loss of the TAM complex [63,66].

Recently the structure of TamA was solved showing that it forms a 16 stranded  $\beta$ -barrel with three POTRA domains [67] (Figure 1.5). Like BamA,  $\beta$ -strands 1 and 16 of TamA show relatively weak hydrogen bonding and the adjacent strands are also shortened suggesting a similar mechanism of function to that of BamA. Based on the positioning of POTRA 2 and 3 it is speculated that this site of TamA would act to prime target substrates for intermolecular  $\beta$ -augmentation for assembly into the outer membrane [67]. Upon the recognition of target substrates the POTRA domains were found to shift relative to the outer membrane [63]. If TamB is rigidly anchored in the inner membrane, such movements of the TamA POTRA domains may cause a change in the orientation of the TamA barrel in relation to the outer membrane. Such movements may cause a disruption of the membrane which may facilitate the assembly of OMPs via TamA. However it may be possible that the movements of the TamA POTRA domains lead to a conformational change of TamB relative to the TamA. These biophysical studies may provide a mechanistic snap shot of the active and non-active forms of the TAM complex. However like the BAM complex, the exact mechanism of OMP assembly by the TAM remains to be demonstrated.

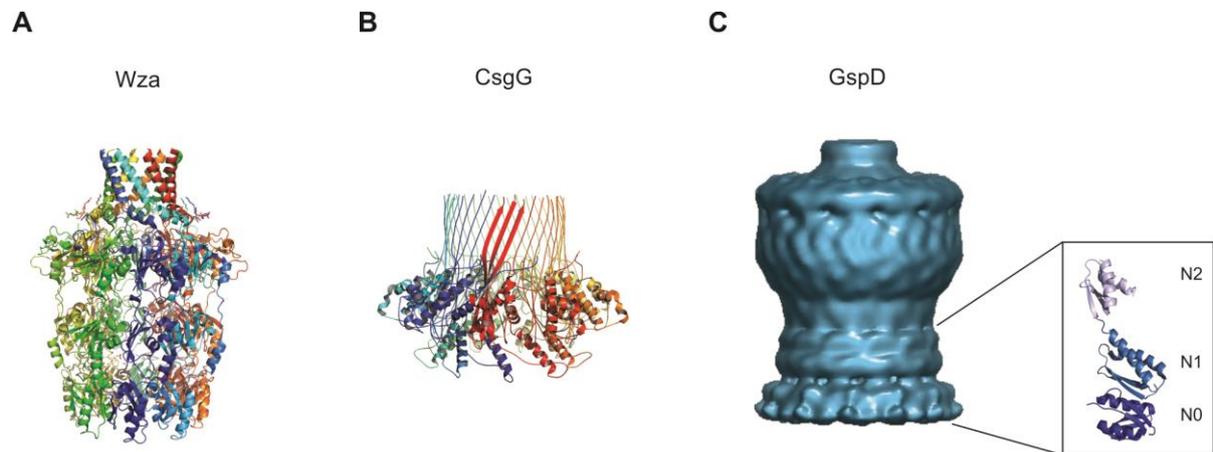


**Figure 1.5: Structure of Omp85 proteins BamA and TamA.** The crystal structures of BamA (PDB: 4K3B, [58]) and TamA (PDB: 4C00, [67]) revealed that the C-terminal domain consists of a 16 stranded  $\beta$ -barrel but differ in the nature of their N-terminal domains. BamA contains five periplasmic N-terminal polypeptide transport-associated (POTRA) domains whilst TamA contains three POTRA domains. The lateral gate between  $\beta$ -strands 1 and 16 of both BamA and TamA are indicated by arrows.

### ***1.7 Multimeric bacterial outer membrane proteins***

Proteinaceous substrates are secreted across the outer membrane by specific molecular machines referred to as secretion systems. Currently nine different secretion systems have been identified (Type 1-9 secretion systems) [6,68]. Depending on the system, substrates are secreted directly from the cytoplasm or from the periplasm to the extracellular milieu [6]. The secretion of various macromolecules and surface appendages across the outer membrane relies on the presence of translocation channels embedded into the outer membrane. Through the use of cryo-EM and other structural biology techniques, the translocation channels of some of these machines comprise of a large multimeric pore with each monomer contributing to the wall of the secretion channel. The translocation channels for the secretion of substrates of the type 2 secretion system (T2SS), type 3 secretion system (T3SS), type 4 pili (T4P) are referred to as secretins and have been the subject of extensive biochemical, genetic and structural analysis [69].

The secretion of group 1 and group 4 capsule to the cell surface also relies on the presence of a large oligomeric pore, formed from the protein Wza [70]. The crystal structure of the Wza octamer showed its transmembrane domain to be composed of 8  $\alpha$ -helices, making Wza the first outer membrane protein reported to adopt an  $\alpha$ -helical transmembrane domain (Figure 1.6a). The crystal structure of another oligomeric pore, the curli translocon CsgG (T8SS), was recently solved but this structure showed that the transmembrane domain of the pore is comprised of 36  $\beta$ -strands, with 4  $\beta$ -strands contributed by each of the 9 CsgG subunits [20,21] (Figure 1.6b). The mechanistic details of how these pores are assembled into the outer membrane remained unknown.



**Figure 1.6: Structures of multimeric secretion pores Wza, CsgG and GspD.** Ribbon structures of (A) Wza (PDB: 2J58, [19]) and (B) CsgG (PDB: 4UV3, [20]). (C) Cryo-EM structure of the GspD secretion channel [71]. The N-terminal domains of GspD are exposed to the periplasm and the crystal structure of the N0, N1 and N2 domains are shown inset (PDB: 3EZJ, [72])

## ***1.8 Secretins***

Secretins are a unique and important family of integral bacterial outer membrane proteins. They form translocation channels for the secretion of substrates of the T2SS, T3SS, T4P and for assembly of filamentous bacteriophage [71,73-75]. Secretins form large homologous multimeric complexes of 12-15 subunits with a total size ~ 0.5 – 1 MDa, with each monomer contributing to the wall of the secretion channel [71,74,76-78]. Once assembled, these structures are extremely stable and can resist dissociation in many detergents and in some cases heat treatment [79].

## ***1.9 Secretin architecture***

The structure of several secretin complexes have been determined by cryo-electron microscopy. In each case these studies have reported a stacked ring shaped complex with a large central cavity [71,74,76-78]. Structural predictions combined with cryo-electron microscopy evidence suggest a complete distinction between the architecture of secretins and the ‘classical’ group of OMPs. The ‘classical’  $\beta$ -barrel OMPs are found as either monomers or trimers, with each protomer forming a barrel to provide a channel across the outer membrane [80,81].

Reichow *et al.*, (2010) reconstructed the cryo-electron microscopy structure, to 19 Å resolution, of the secretin GspD from *Vibrio cholerae* [71] (Figure 1.6c). GspD appeared as a dodecomeric channel of ~155 Å in diameter and ~200 Å in length. Interestingly the inner chamber of GspD was observed to be ~ 55 Å in diameter, providing a pore large enough for the secretion of folded protein substrates. Analysis of the cross-section of GspD showed two gated regions separating the extracellular milieu, the extracellular chamber and the

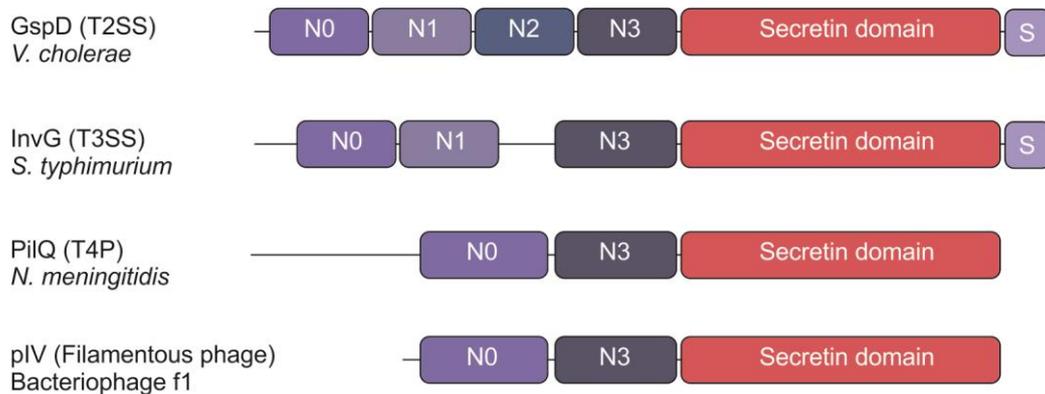
periplasmic entrance. This gated region is believed to ensure membrane integrity, but has a permeation cut-off similar to that of general porins and may undergo conformational changes to allow secretion of target substrates [82,83].

The crystal structure of the periplasmic N-terminal domains of secretins GspD and XcpQ from enterotoxigenic *Escherichia coli* (ETEC) and *Pseudomonas aeruginosa* respectively has been solved and showed a 'stacked' ring structure of three of the N-terminal domains [72,84] (Figure 1.6c). The nature of the C terminus is unknown however circular dichroism analysis of PulD revealed ~ 27% of the C terminus to contain  $\beta$  structures [77]. Whether the C-terminal domain spans the outer membrane in the form of  $\beta$  strands cannot be interpreted from these solution studies alone, and the cryo-electron microscopy structures of many secretins suggest instead that, in a membrane milieu, the C-terminal region adopts an alternative conformation.

### ***1.10 Secretin domains***

Biochemical and bioinformatic analysis of secretins revealed conserved domains amongst secretin family members (Figure 1.7). All secretin family members contain a large domain at the C-terminus referred to as the secretin domain. This domain is highly conserved amongst all secretin family members and structural and biochemical studies has shown that this domain is embedded in the outer membrane [71,77,85-87].

The N terminal region of many secretins is comprised of up to 4 sub-domains N0, N1, N2 and N3. This N-terminal region is highly variable and has relatively low sequence similarity



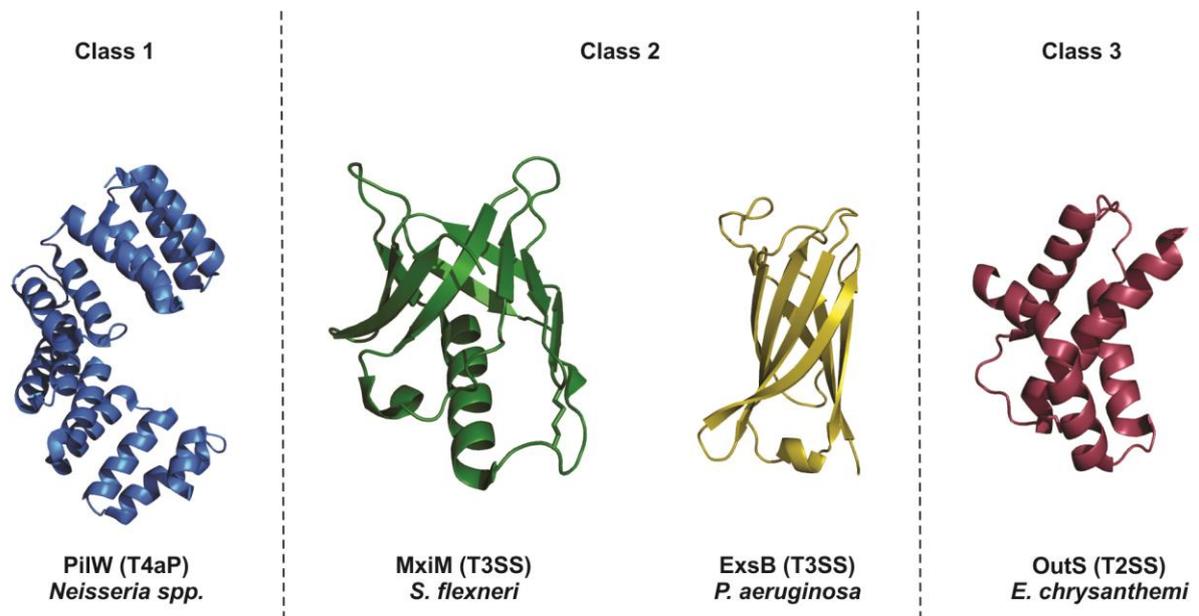
**Figure 1.7: Domain organisation of secretins.** The C-terminus of secretins consists of the membrane embedded secretin domain (Pfam PF00263) and in some cases a small unstructured domain at the far C-terminus referred to as the S domain which binds its cognate pilotin. The N-terminal domain which is exposed to the periplasm consists of up to 4 sub-domains comprising the N0 domain (Pfam PF07660) followed by one or several homologous domains referred to as N1-N3 (Pfam PF03958).

between different secretin subfamilies. For example, the N0 domains from different secretins adopt a similar structural fold despite low amino acid similarity. The N0 domain was shown to form a similar structural fold related to the signalling domain of TonB dependent receptors and bacteriophage tail proteins [72,84,88]. The N1 and N2 domains have been shown to form the same structural fold and it is predicted that the N3 sub-domain contains a similar structure [72,84]. These N-terminal regions have been shown to be involved in substrate recognition [89-91], to interact with other components of the secretion machinery [92,93] or provide gating of the secretion channel [71,82,83,86].

Biochemical studies have revealed several key residues that are vital for efficient multimerization [94-96]. In most cases, mutating residues or inserting linker regions located in the secretin and N3 domains ceased the ability of the secretin to form multimers. A small domain at the far C-terminus of some secretins has been characterised and is referred to as the S-domain and is responsible for secretin-pilotin binding [97-100]. The significance of this interaction will be discussed in greater detail below.

### ***1.11 Piloting and assembly of secretins to the outer membrane***

The stabilization, oligomerization or translocation of secretins to the outer membrane in some cases requires the interaction with auxiliary proteins referred to as a pilotin. Pilotins are typically small lipoproteins that are diverse in sequence, structure and functionality and as such have been classified into groups based on their structural characteristics [101] (Figure 1.8). Class 1 pilotins consist of  $\alpha$ -helical tetratricopeptide (TPR) motifs and are roughly twice the size of other pilotins. Class 2 pilotins are composed predominately of  $\beta$ -strands which adopt a cracked  $\beta$ -barrel or  $\beta$ -sandwich. Finally, class 3 pilotins form  $\alpha$ -helical bundles.



**Figure 1.8: Structures of the different pilotin classes.** Class 1 pilotins consist of  $\alpha$ -helical tetratricopeptide (TPR) motifs (PilW, PDB: 2VQ2 [103]). Class 2 pilotins are composed predominately of  $\beta$ -strands which adopt a cracked  $\beta$ -barrel or  $\beta$ -sandwich (MxiM, PDB: 1Y9L [104]; ExsB, PDB: 2YJL [105]). Finally, class 3 pilotins form  $\alpha$ -helical bundles (OutS, PDB: 3UTK [100]).

Pilotins have been identified for secretins of the T2SS, T3SS and T4aP [79,102,103], but have not yet been identified for competence systems, filamentous phage or T4bP.

As mentioned above, the S-domain is a small domain at the far C-terminus of some secretins which is critical for the interaction with its cognate pilotin [97-100]. Structural studies have shown that the S-domain is largely unstructured but adopts an  $\alpha$ -helical conformation upon binding its pilotin [98,100]. However, PilQ from *P. aeruginosa* contains a much shorter C-terminal tail adjacent to the secretin domain. Deletion of this region did not affect T4P function or PilQ assembly, suggesting the mechanism of pilotin binding is different to that of previously characterized secretins [106]. Biochemical mapping experiments showed PilQ mutants lacking the N0, N1, or the N-terminal segment of the secretin domain, resulted in a loss of T4P function and assembly of PilQ [106]. Whilst these domains are important intrinsic factors for PilQ assembly, they may also provide speculative binding regions for the interaction with its pilotin PilW. But this needs to be verified experimentally.

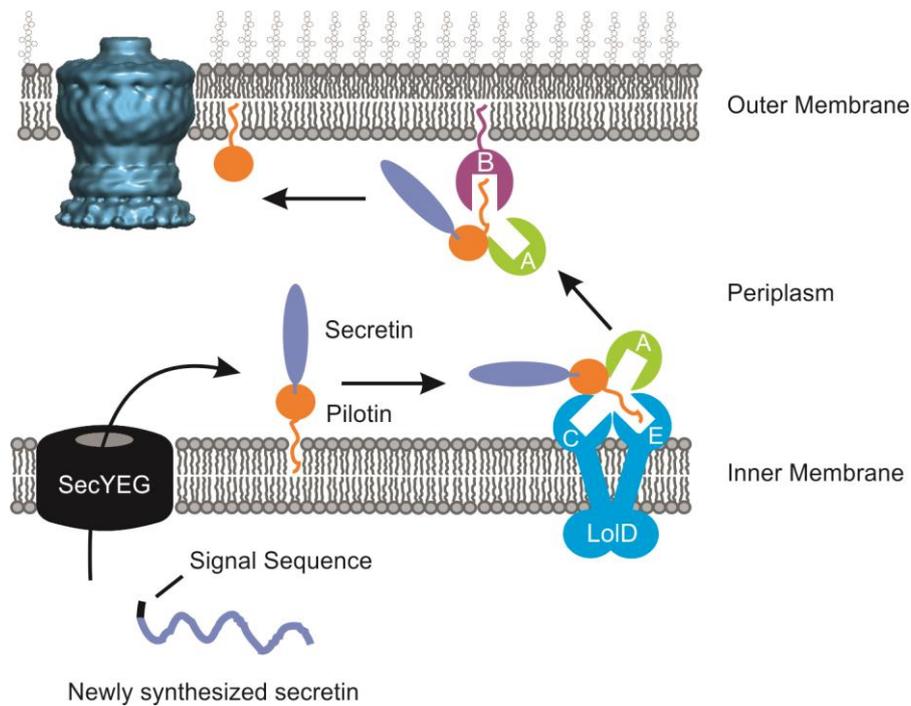
PulS is the archetypal pilotin for the T2SS secretin PulD of *K. oxytoca*. Hardie *et al.*, (1996) showed that PulS is required for the stabilization and translocation of PulD to the outer membrane [79]. Interestingly, in the absence of PulS, a large proportion of PulD was observed to form oligomers which were localised to the inner membrane instead of the outer membrane [107]. Given that PulS is a lipoprotein, Collin *et al.*, (2011) investigated the role of the Lol pathway in transporting possible PulS-PulD heterodimers to the outer membrane [108]. The dependence on the Lol pathway was shown in which a lipidated mutant of PulD lacking the S-domain could still assemble into multimers. Additionally, these PulD multimers were localised at the outer membrane suggesting the Lol pathway could replace the sorting

function of PulS. To confirm this, Collin *et al.*, (2011) created a LolA<sup>R43L</sup> variant which prevents the transfer of LolA<sup>R43L</sup> bound proteins to LolB. Both LolA<sup>R43L</sup>-PulS and LolA<sup>R43L</sup>-PulS-PulD complexes were localized to the periplasm and inner membrane respectively [108]. Taken together, these findings put forward the first detailed report of pilotin-secretin translocation to the outer membrane using a dedicated protein translocase, which is being referred to as the “piggy-back model” (Figure 1.9).

### ***1.12 Role of accessory factors in the assembly of secretins***

Several accessory proteins have been identified for secretins of the T2SS, T3SS, T4P and filamentous phage. Accessory proteins are not always associated with all secretion systems and those that have been characterized show variations in their function. Many of which are associated with the stability of the secretin pore or of the secretin monomer. T2SS proteins ExeA (GspA) and ExeB (GspB) from *Aeromonas hydrophila* and *Vibrio cholerae* are crucial for the targeting and assembly of secretin ExeD/GspD to the outer membrane [109,110]. ExeAB forms a complex at the inner membrane in which ExeA binds peptidoglycan and ExeB interacts with the N0N1 periplasmic domains of secretin ExeD [111,112]. However, homologs of ExeA or ExeB have not been identified in all characterized T2SSs [110] and in the case of the Out system from *Erwinia chrysanthemi*, only the ExeB homolog OutB can be identified [113].

In some cases, secretins are themselves lipoproteins and are self-targeted to the outer membrane, presumably by the Lol pathway [114-116]. Accessory proteins BfpG and TcpQ have been shown to play a role in the assembly of T4bP secretins BfpB and TcpC respectively [114,115]. These accessory proteins are not lipidated and instead are themselves



**Figure 1.9: “Piggyback” model of secretin translocation to the outer membrane.** The Lol pathway facilitates the translocation of lipoproteins across the periplasm to the outer membrane. A crucial step in the assembly of secretin pores is the translocation of secretin monomers to the outer membrane. The piggyback model involves the interaction of the secretin monomer with LolA via its lipoprotein pilotin. The secretin-pilotin heterodimer are then transported across the periplasm to LolB at the outer membrane for subsequent assembly into mature multimers.

“piggybacked” on their cognate secretin for transit to the outer membrane; a reversal of the fashion observed for the translocation of PulD to the outer membrane by its pilotin PulS.

The inner membrane lipoprotein MxiJ is an accessory factor involved in the assembly of T3SS secretin MxiD of *Shigella flexneri* [102]. Schuch and Maurelli, (2001) reported that MxiJ interacts with MxiD which in turn facilitates the stability and multimerization of the MxiD pore with a loss of both MxiJ and pilotin MxiM leading to the degradation of MxiD [102]. Finally, Jain *et al.*, (2011) showed that secretin PilQ from *N. meningitidis* suffered reduced stability and underwent a change in symmetry (14 to 19 fold) in the absence of its accessory protein PilP [117].

### ***1.13 Does the BAM complex play a role in secretin assembly?***

The folding and assembly of OMPs into the outer membrane is facilitated by the essential core  $\beta$ -barrel BamA and four accessory lipoproteins that make up the BAM complex. Several studies have aimed to determine the role of the BAM in assembling a variety of different secretins. Previous studies have shown that secretins PulD (*K. oxytoca* T2SS), XcpQ (*P. aeruginosa* T2SS), PscC (*P. aeruginosa* T3SS) and PilQ (*P. aeruginosa* T4P) can assemble into multimers in cells depleted of BamA [118,119]. However, the PilQ homolog from *N. meningitidis* requires BamA (Omp85) for assembly into the outer membrane [44].

### ***1.14 Aims of this study***

Given the different roles of auxiliary proteins and assembly factors in assembling secretins, the aims of the thesis were to:

1. Develop an *in vivo* assay to monitor the assembly of secretins and other multimeric secretion pores.
2. Identify the cognate pilotin for secretin GspD in *E. coli*.
3. Optimize and further characterize a *bamA* depletion strain of *E. coli* to monitor outer membrane protein assembly.
4. Determine what role if any the BAM complex plays in assembling secretin GspD and multimeric secretion channels Wza and CsgG.

## **CHAPTER 2: Identification of a novel pilotin AspS**

### ***2.1 INTRODUCTION***

The T2SS, also referred to as the general secretory pathway (GSP), is a multi-protein machine that spans both the inner and outer membranes of Gram-negative bacteria [120,121]. The T2SS facilitates the secretion of folded proteins, typically toxins or enzymes, from the periplasm to the extracellular milieu via a two-step process [120,121]. Secreted proteins are synthesized as a precursor with a cleavable N-terminal signal sequence targeting the protein to the Sec translocon or Tat complex for transport across the inner membrane into the periplasm [122,123]. The signal sequence is subsequently cleaved and the target substrate protein is translocated across the outer membrane in a folded state by the T2SS machinery [124,125].

Depending on the bacterial species, the T2SS comprises of 12-16 different proteins which are associated with either the inner membrane or the outer membrane [120,121]. The T2SS machine is made of two sub-complexes at the inner membrane. The first is built around the ATPase GspE, which forms a complex with GspL, GspM, GspF and GspC providing a structural platform for the machine. The second complex referred to as the pseudopilus is comprised of GspG, GspH, GspI, GspJ and GspK. This pseudopilus complex has been proposed to act as a piston or plug during the secretion of target substrates. Finally, the secretin GspD is the integral outer membrane protein of the T2SS and provides the pore for the transport of target substrates across the outer membrane. At the time of this study's

inception, secretins were thought to rely on the presence of a small lipoprotein termed pilotin (GspS) or other accessory factors for targeting to and assembly in the outer membrane.

In *E. coli* three distinct T2SSs have been characterized. A plasmid encoded T2SS found in enterohemorrhagic *E. coli* (EHEC) that secretes the complement cleaving protease StcE [126]. A second system (sometimes referred to as T2SS $\alpha$ ) found in the *E. coli* lab strain K-12 and other pathogenic *E. coli* isolates that secretes chitinase [127]. The third system (sometimes referred to as T2SS $\beta$ ) found in many pathogenic *E. coli* isolates that was initially identified in enterotoxigenic *E. coli* responsible for the secretion of the heat-labile toxin (LT) [128].

This study is interested in the “T2SS $\beta$ ” of enteropathogenic *E. coli* (EPEC). EPEC is a leading cause of infantile diarrhoea in developing countries [129] but unlike ETEC does not secrete the LT toxin [130,131]. EPEC uses this T2SS to secrete the mucin degrading enzyme SsIE [15,132], which enables the bacterium to translocate through the mucin layer to get access to host cells for adhesion and colonization [133]. The T2SS $\beta$  machinery of EPEC is comprised of 14 proteins; however a pilotin has not been identified for secretin GspD. In this chapter, the identification of a novel pilotin AspS, will be presented. Structural biology and bioinformatics was used to demonstrate that AspS forms a new class of pilotins that despite the structural differences interact with its cognate secretin in a similar way compared to previously characterized pilotins from other secretion systems.

## **2.2 METHODS and MATERIALS**

### **2.2.1 Bioinformatics**

#### *2.2.1.1 Hidden Markov Models*

Hidden Markov models (HMM) were generated and HMMER searches performed using HMMER v.2.4 [134] to search for the T2SS pilotin candidate in enteropathogenic *Escherichia coli*. HMMER searches performed using *Vibrio cholerae* O1 biovar El Tor N16961 were performed by Dr. Eva Heinz (Lithgow Laboratory, Monash University). A list of the protein sequences used to build the HMM can be found in Appendix A1.

#### *2.2.1.2 Sequence analysis predictions*

LipoP 1.0 [135] was used to predict lipoprotein signal peptides of YacC and YghG (AspS), ([www.cbs.dtu.dk/services/LipoP](http://www.cbs.dtu.dk/services/LipoP)). The conserved domain architecture tool (CDART) [136], (<http://www.ncbi.nlm.nih.gov/Structure/lexington/lexington.cgi>) was used to define the PulS\_OutS domains of Puls, OutS, EtpO and YacC.

#### *2.2.1.3 Sequence cluster visualization*

Similarity-based clustering analyses were performed by Dr. Eva Heinz (Lithgow Laboratory, Monash University) using the CLANS software [137], a graph-based sequence similarity visualization software based on sequence similarities obtained by BlastP p-values using BLAST 2.2.26 with default settings as implemented in the CLANS software.

### 2.2.2 Strains, plasmids and growth conditions

The bacterial strains and plasmids used in this chapter are listed in Table 2.1, using the parental strains enteropathogenic *E. coli* E2348/69, *E. coli* BL21 (DE3) (Invitrogen), Rosetta (DE3) (Novagen) and *E. coli* DH5 $\alpha$  (Invitrogen). Strains were grown in Luria Broth (LB, 10 g/L Tryptone, 5 g/L Yeast Extract, 5 g/L NaCl) or Casamino acid-yeast extract-salts (CAYE, [138,128]), as indicated, supplemented with the appropriate antibiotics (ampicillin 100  $\mu$ g/ml, kanamycin 30  $\mu$ g/ml or chloramphenicol 12.5  $\mu$ g/ml).

**Table 2.1: Strains and Plasmids used in Chapter 2**

<b>Strain</b>	<b>Description</b>	<b>Reference</b>
E2348/69	EPEC serotype O127:H6	[139]
E2348/69 $\Delta$ <i>gspD</i>	E2348/69 $\Delta$ <i>gspD</i>	[15]
E2348/69 $\Delta$ <i>sslE</i>	E2348/69 $\Delta$ <i>sslE</i>	[15]
E2348/69 $\Delta$ <i>yacC</i>	E2348/69 $\Delta$ <i>yacC</i>	This Study
E2348/69 $\Delta$ <i>aspS</i> ( <i>yghG</i> )	E2348/69 $\Delta$ <i>aspS</i> ( <i>yghG</i> )	This Study
E2348/69 $\Delta$ <i>gspD</i> $\Delta$ <i>aspS</i>	E2348/69 $\Delta$ <i>gspD</i> ::Cm <sup>r</sup> , $\Delta$ <i>aspS</i>	This Study
BL21 (DE3)	<i>E. coli</i> strain: F- <i>ompT hsdSB</i> (rB- mB-) <i>dcm gal</i> (DE3)	Invitrogen
BL21 (DE3) $\Delta$ <i>aspS</i> ( <i>yghG</i> )	BL21 (DE3) $\Delta$ <i>aspS</i>	This Study
BL21 (DE3) $\Delta$ <i>gspD</i> $\Delta$ <i>aspS</i> ( <i>yghG</i> )	BL21 (DE3) $\Delta$ <i>gspD</i> ::Cm <sup>r</sup> , $\Delta$ <i>aspS</i>	This Study
DH5 $\alpha$	F <sup>-</sup> <i>endA1 glnV44 thi-1 recA1 relA1 gyrA96 deoR nupG <math>\Phi</math>80dlacZ<math>\Delta</math>M15 <math>\Delta</math>(lacZYA-argF)U169, hsdR17(r<sub>K</sub><sup>-</sup> m<sub>K</sub><sup>+</sup>), <math>\lambda</math>-</i>	Invitrogen
<b>Plasmid</b>	<b>Description</b>	<b>Reference</b>
pBAD24	<i>ori</i> pMB1, Amp <sup>r</sup>	[140]

pET DUET-1	<i>ori ColE1, lacI</i> gene, Amp <sup>r</sup>	Novagen
pACYC DUET-1	<i>ori P15A, lacI</i> gene, Cm <sup>r</sup>	Novagen
pET DUET-1 GspD-C <sub>4</sub> AspS	E2348/69 <i>gspD-C<sub>4</sub></i> and <i>yghG</i> (AspS) were cloned into NcoI/HindIII and NdeI/XhoI sites of pET DUET-1 respectively	This Study
pET DUET-1 GspD-C <sub>4</sub> YacC	E2348/69 <i>gspD-C<sub>4</sub></i> and <i>yacC</i> were cloned into NcoI/HindIII and NdeI/XhoI sites of pET DUET-1 respectively	This Study
pET DUET-1 GspDΔS-C <sub>4</sub> AspS	E2348/69 <i>gspDΔS-C<sub>4</sub></i> and <i>yghG</i> (AspS) were cloned into NcoI/HindIII and NdeI/XhoI sites of pET DUET-1 respectively	This Study
pJP117	E2348/69 <i>gspD-C<sub>4</sub></i> is cloned into the NcoI/XbaI sites of pNM12 (pBAD24 modified to include MscI site).	This Study (Dr. Judyta Praszkiar Robins-Browne laboratory)
pJP168	P <sub>BAD</sub> promoter and <i>araC</i> of pBAD24 replaced by P <sub>tetA</sub> and <i>tetR</i>	[15]
pJP181	E2348/69 <i>gspD-C<sub>4</sub></i> was cloned into the NcoI/XbaI sites of pJP168.	This Study (Dr. Judyta Praszkiar Robins-Browne laboratory)
pFT-A	<i>ori</i> R101, <i>repA101ts</i> , flp, Amp <sup>r</sup>	[141]
pGEM-T-Easy	<i>ori</i> pMB1, Amp <sup>r</sup>	Promega
pJP133	pGEM-T-Easy containing the <i>yghG::Kn<sup>r</sup></i> cassette	This Study (Dr. Judyta Praszkiar Robins-Browne laboratory)
pKD3	FRT flanked Cm <sup>r</sup> gene, Cm <sup>r</sup> , Amp <sup>r</sup>	[142]
pKD4	FRT flanked Kan <sup>r</sup> gene, Kan <sup>r</sup> , Amp <sup>r</sup>	[142]
pKD46	λ Red recombinase, <i>ori</i> R101, <i>repA101ts</i> , Amp <sup>r</sup>	[142]

### 2.2.3 Generation of isogenic gene deletions

Bacterial mutants resulting from the deletion of the genes *gspD*, *yacC* or *yghG* (*aspS*) were constructed in E2348/69 and BL21 by allelic exchange with *gspD*::Cm<sup>r</sup>, *yacC*::Kan<sup>r</sup>, *yghG*::Kan<sup>r</sup>. These knockouts were generated utilising the  $\lambda$  Red recombinase system carried on plasmid pKD46 [142]. Phusion high-fidelity DNA polymerase (New England BioLabs) was used to amplify the FRT flanked Cm or Kn antibiotic cassette from pKD3 and pKD4 respectively and the corresponding upstream and downstream fragments *gspD*, *yacC* and *yghG*. The Cm<sup>r</sup> and Kn<sup>r</sup> cassettes were joined to the corresponding upstream and downstream flanks by overlapping PCR with Platinum Taq DNA polymerase (Thermo Fisher Scientific) using primer pairs GspD F/GspD R, YacC F/YacC R and YghG BL F/YghG BL F. These fragments were subsequently cloned into pGEM-T\_Easy to create donor plasmids. The above primer pairs were used to generate the PCR fragments required to construct the gene deletions according to Datsenko and Wanner (2000) [142]. When required the Kan<sup>r</sup> or Cm<sup>r</sup> genes were removed using the flanking FRT sites and FLP on plasmid pFT-A [141]. All mutations were confirmed by PCR analysis, using primers flanking the targeted regions. Primers used for the construction of listed isogenic gene deletions are detailed in Table 2.2.

**Table 2.2: Primers used to generate *gspD*, *yacC* and *yghG* isogenic gene deletions in EPEC E2348/69 and BL21 (DE3).**

Primer Name	Primer sequence	Details
YghG BL F	ACTCTTATTAATAACCTTAGAGATT ATTTACC	Used to amplify the <i>yghG</i> ::Kan <sup>r</sup> from plasmid pJP133
YghG BL R	GATAAATTCGTGCGTCACGAAAAAC	
YghG BL F Seq	ATGCTGCGGCCTGATATATGC	Sequencing primers
YghG BL R Seq	AGCAGCATCAGCCAGAACATC	

GspD F	CACTGATCCACGAGCAATGATTGC	Used to amplify the GspD upstream flank
GspD CK.R	GAAGCAGCTCCAGCCTACACACCCC TGTGCTTCCAGCAGGTTAAG	
GspD R	GTCGATTTGCCGCAGCGTAAAG	Used to amplify the GspD downstream flank
GspD CK.F	CTAAGGAGGATATTCATATGGTGAC CATTCTGCGTGACGGTATGG	
GspD UF Seq	GTCGGTTATGCAGTGAAGCCG	Sequencing primers
GspD DR seq	GTCCTCGGCAGAACCGAGATC	
YacC F	GGTACTTCCAGCCCGTGCCAG	Used to amplify the YacC upstream flank
YacC CK.R	GAAGCAGCTCCAGCCTACACACTTC ATTGCTTCTACCAGGGGCTTAAAG	
YacC R	ATAGATAAACTCGTCGCGCTCGGTG	Used to amplify the YacC downstream flank
YacC CK.F	CTAAGGAGGATATTCATATGGGCA TTGGCATTCCGGTC	
YacC Seq F	GGCAGGTGGATCAACGTTCAACG	Sequencing primers
YacC Seq R	GGTACTTCCAGCAGCATGGCAC	
Kan1	GTATCCATCATGGCTGATGC	Internal Kan <sup>r</sup> check primers, used with YacC and YghG seq primers
Kan2	CCGCTATCAGGACATAGCG	
Cm.C1	TTATACGCAAGGCGACAAGG	Internal Cm <sup>r</sup> check primers used with GspD seq primers
Cm.C2	GATCTTCCGTCACAGGTAGG	
pKD4F	TGTGTAGGCTGGAGCTGCTTC	Used to amplify FRT flanked Kan or Cm antibiotic cassette
pKD4R	CATATGAATATCCTCCTTAG	

#### 2.2.4 Cloning of *gspD*, *yacC* and *yghG* (*aspS*)

The open-reading frames for GspD, YacC and YghG (AspS) correspond to sequences in the Gene ID entries, *gspD* (7064260), *yacC* (7062744) and *yghG* (*aspS*) (7064258). Plasmids

pJP117 and pJP168 were constructed by Dr. Judyta Praszquier (Robins-Browne Laboratory, The University of Melbourne). The gene encoding *gspD-C<sub>4</sub>* was subcloned from pJP117 into pET DUET-1 using NcoI and HindIII restriction enzymes. Primers used for the amplification of *gspD*, *yacC* and *yghG* are listed in Table 2.3. The resulting PCR fragments were cloned into pET DUET-1 GspD-C<sub>4</sub> using NdeI and XhoI restriction enzymes to create pET DUET-1 GspD-C<sub>4</sub> YacC and pET DUET-1 GspD-C<sub>4</sub> YghG (AspS)

**Table 2.3: Primers used to construct the GspD, YacC and YghG (AspS) expressing plasmids.**

Primer Name	Primer sequence
YacC F	GGCGCTCATATGAAGACGTTTTTCAGAACAG
YacC R	CGGCTCTCGAGTTATTTGACGTAGGCAAGC
YghG F	GGCGCTCATATGTCGATAAAACAAATGC
YghG R	GCCGGCTCGAGTTATGCTTTGACTATTCC
GspD F	CGGTACCATGGTTTGCGGTGATATGACG
GspDΔS-C <sub>4</sub> R	GCCGGAAGCTTTTAACGGCCCGCCGGTTCCATGCAGCA GCCCCGGCAGCAGTTCAGAAAGCTGCCCGCACGCAGA ATGGTCGGACG

### 2.2.5 Sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE)

SDS-PAGE was performed under denaturing and reducing conditions as previously described [143], unless otherwise stated. Stacking buffer (375 mM Tris pH 6.8, 0.02 % SDS and 0.5 mM EDTA) and separating buffer (375 mM Tris pH 8.8, 0.1 % SDS and 0.5 mM EDTA) were used to make continuous 10% acrylamide gels with standard SDS running buffer (25 mM Tris, 192 mM glycine and 0.1% SDS) Table 2.4.

**Table 2.4: Solutions required to make 1 x 10% gradient SDS-PAGE gel.**

	<b>5% Stacking</b>	<b>10% Separating</b>
<b>dH<sub>2</sub>O</b>	4.45 ml	7.78 ml
<b>Gel Buffer</b>	750 µl	4.41 ml
<b>Acrylamide (40% 29:1; Biorad)</b>	750 µl	4.38 ml
<b>APS (20% w/v)</b>	60 µl	87.5 µl
<b>TEMED</b>	6 µl	10.5 µl

In order to visualize the oligomers for each protein, optimization of the SDS-PAGE system was achieved by running the protein samples through a 3% stacking gel and 3-14% separating gel. An SG50 Hoefer gradient pourer was used to create 2 mm thick 3-14% SDS-PAGE gels (Table 2.5).

**Table 2.5: Solutions required to make 1 x 3-14% gradient SDS-PAGE gel.**

	<b>3% Sep</b>	<b>14% Sep</b>	<b>3% Stack</b>
<b>dH<sub>2</sub>O</b>	9.33 ml	5.48 ml	7.89 ml
<b>Gel Buffer</b>	3.54 ml	3.54 ml	1.25 ml
<b>Acrylamide (40% 29:1; Biorad)</b>	1.05 ml	4.9 ml	750 µl
<b>APS (20% w/v)</b>	70.16 µl	70.16 µl	100 µl
<b>TEMED</b>	8.35 µl	8.35 µl	10 µl

Unless stated, traditional SDS sample buffer (50 mM Tris-Cl pH 6.8, 1% SDS, 10% glycerol, 100 mM DTT, 0.005% bromophenol blue) was added to protein samples and boiled for 10 min prior to loading. Gels were typically run at 60 V overnight.

### **2.2.6 Immunoblotting**

Proteins were transferred to nitrocellulose membranes from SDS-PAGE gels in transfer buffer (25 mM Tris, 190 mM glycine and 10% methanol). Proteins were transferred for 60 or 90 minutes at 1 A for 0.75 mm or 1.5 mm thick gels respectively. Membranes were blocked with Tris buffered saline (TBS) containing 0.1% Tween-20 and 5% skim milk powder. Membranes were subsequently incubated with primary antibodies raised against the desired protein of interest. Detection of the protein was obtained by using goat anti-rabbit/mouse secondary antibody conjugated to horse-radish peroxidase (HRP) and home-made ECL. Immediately before ECL detection, equal volumes of solution 1 was added to solution 2, mixed and added to the membrane. Excess ECL was removed and the membrane exposed to Fuji X-ray film and developed. Solution 1 is made by separately mixing 200 mg of luminol and 35 mg of p-coumaric acid in 4.5 ml and 2.5 ml dimethyl sulfoxide (DMSO) respectively. These two solutions were mixed and made up to 250 ml with 100 mM Tris, pH 9.35. Solution 2 consisted of 30  $\mu$ l of hydrogen peroxide ( $H_2O_2$ , 30% stock) in 250 ml 100 mM Tris, pH 9.35. Both solutions were stored at 4 °C.

### **2.2.7 Assays to measure GspD assembly**

E2348/69 and BL21 (DE3) cells transformed with *gspD-C<sub>4</sub>* expressing plasmids were grown in LB to  $OD_{600} = 0.6$  at 37 °C prior to induction with arabinose (0.1%) or IPTG (0.1mM) respectively for 2 hours at 37 °C. After 0, 15, 30, 60 and 120 minutes, 1 ml samples were

taken and cells were harvested by centrifugation. Cell pellets were resuspended in a non-standard lysis buffer (50 mM  $\text{KH}_2\text{PO}_4$  pH 7.8, 400 mM NaCl, 100 mM KCl, 10% glycerol, 1% DDM and 10 mM imidazole) proportional to the culture density (e.g. if  $\text{OD}_{600}$  is 0.600 then 60  $\mu\text{l}$  of sample buffer is used). Lumio Green Dye is a proprietary stain (Thermo Fisher Scientific) which binds proteins carrying the tetra cysteine sequence CCPGCC. Fifteen microlitres of whole cell lysate were added to 5  $\mu\text{l}$  of Lumio sample buffer and 0.2  $\mu\text{l}$  of Lumio Green dye and incubated at 70 °C for 10 minutes. Samples were allowed to cool to room temperature before the addition of 2  $\mu\text{l}$  Lumio enhancer solution (Thermo Fisher Scientific). Samples were incubated for 5 minutes at room temperature before loading onto a 3-14% SDS-PAGE. Tagged GspD was detected using fluorometry (Typhoon Trio, Argon Blue 488nm laser, 520nm BP40 filter) and analysed using ImageQuant (GE) and ImageJ.

### ***2.2.8 Isolation of total membranes***

An overnight culture was grown in 5 ml of LB at 37°C from a single colony. The starter culture was diluted 1:100 into 400 ml of LB with antibiotics as required. Cultures were grown to  $\text{OD}_{600} = 1.0$ . Note, expression of GspD-C<sub>4</sub> from pJP117 or pET DUET-1 GspD-C<sub>4</sub> AspS was induced with arabinose (0.1%) or IPTG (0.1mM) respectively when the culture reached an  $\text{OD}_{600}$  of 0.6. Cells were grown for further 1 hour at 37°C. Cells were harvested by centrifugation (5000 x g, 10 min, 4 °C), and resuspended in TS buffer (0.75 M sucrose/10mM Tris-HCl, pH 7.5). Lysozyme (50  $\mu\text{g}/\text{ml}$ ), PMSF (2mM) and 2 volumes of 1.65 mM EDTA, pH 7.5, were added sequentially before homogenizing cells with an EmulsiFlex (Avestin Inc.) at 15,000 psi. Membranes were collected by ultracentrifugation (38,000 rpm, 45 minutes, 4°C) and washed in TES (3.3 mM Tris pH 7.5, 1.1 mM EDTA, 0.25 M sucrose) buffer using a dounce. Membranes were pooled (~8 ml) with TES and collected by

centrifugation (38,000 rpm, 45min, 4°C) and then resuspended in 400 µl 25% (w/v) sucrose in 5 mM EDTA, pH 7.5 and stored at -80°C.

### ***2.2.9 Sucrose density fractionation***

Total membranes (~400 µl) were fractionated on a six-step sucrose gradient (1.9 ml of 35:40:45:50:55:60% (w/w) sucrose in 5 mM EDTA, pH 7.5) by ultracentrifugation in a SW40 Ti rotor (34,000 rpm, 17 hours, 4 °C) and 1 ml fractions were stored at - 80°C. Fifteen microlitre aliquots of each fraction were prepared using Lumio detection kit (described in Chapter 2.2.7) and loaded onto a 3-14% SDS-PAGE and analysed by fluorometry (described in Chapter 2.2.7). Western immunoblotting was used to determine the localisation of inner and outer membrane proteins F<sub>1</sub>β (serum dilution 1:8000) and BamA (serum dilution 1:5000) respectively.

### ***2.2.10 Secretome analysis***

Cultures were grown in 30 ml of CAYE media for 4 hours. Culture supernatant were isolated and passed through a 0.45 µm filter before the addition of trichloroacetic acid (TCA; 10% final concentration) and incubated on ice for 1 hour. Precipitated proteins were collected by centrifugation (15,000 rpm, 30 minutes, 4 °C) and protein pellets were washed twice with cold 100% methanol. Pellets were allowed to dry and resuspended in 50 µl SDS sample buffer. Forty microlitres of protein sample were loaded onto 3-14% gradient gels for analysis by SDS-PAGE and Coomassie brilliant blue staining.

### ***2.2.11 In-gel digestion of proteins and LC-MS/MS***

Coomassie-stained protein bands were excised and destained with 50 % methanol in 100 mM  $\text{NH}_4\text{HCO}_3$  (pH 8.5). The proteins were subsequently reduced with 2 mM TCEP-HCl (Thermo Scientific), carbamidomethylated with 20 mM iodoacetamide and digested with 100 ng trypsin. Tryptic peptides were extracted with 50 % acetonitrile (acidified with 1% formic acid), lyophilized in a vacuum concentrator and reconstituted in 15  $\mu\text{l}$  buffer A (0.1 % formic acid). Prior to mass spectrometry, the peptides were further purified and enriched using OMIX C18 Mini-Bed tips.

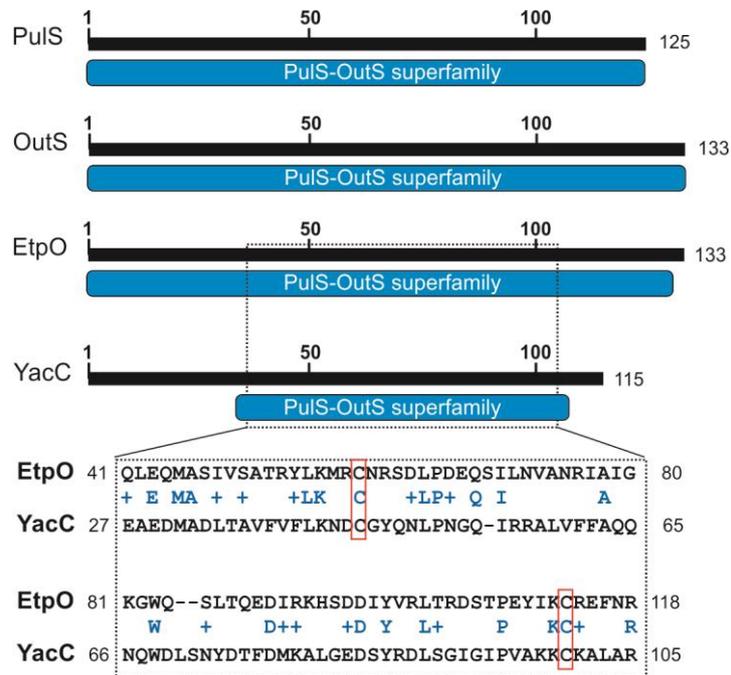
Using a Dionex UltiMate 3000 RSLCnano system equipped with a Dionex UltiMate 3000 RS autosampler, an Acclaim PepMap RSLC analytical column (75  $\mu\text{m}$  x 50 cm, nanoViper, C18, 2  $\mu\text{m}$ , 100 $\text{\AA}$ ) and an Acclaim PepMap 100 trap column (100  $\mu\text{m}$  x 2 cm, nanoViper, C18, 5  $\mu\text{m}$ , 100 $\text{\AA}$ ), the tryptic peptides were separated by increasing concentrations of 80 % Acetonitrile / 0.1 % formic acid at a flow of 250 nl/min for 30 min and analyzed with a QExactive Plus mass spectrometer.

The raw files were converted to the mgf file format using msConvert [144] and searched against the *E. coli* E2348/69 UniProtKB/SwissProt database (v2014\_07) using ProteinPilot software v5.0 (AB SCIEX) to obtain sequence information. Only peptides identified at a false discovery rate (FDR) of 1 % based on a decoy database were further analyzed.

## 2.3 RESULTS

### 2.3.1 *PulS-OutS* family HMM analysis detects *YacC* in the genome of EPEC E2348/69

The PulS-OutS Pfam protein family was initially defined by four protein sequences: PulS from *Klebsiella*, OutS from *Dickeya*, OutS from *Pectobacterium*, and EtpO from *E. coli* O157:H7 (Figure 2.1). A hidden Markov model (HMM) was constructed in order to have a highly-sensitive tool to detect more distantly related members of the PulS-OutS protein family encoded in the EPEC genome. The protein sequences used to build the PulS\_OutS HMM can be found in Appendix A2. Applying the PulS\_OutS HMM with a threshold cut-off E value of  $10e-3$ , a single statistically-significant hit (E value =  $1.10e-41$ ) was recovered, corresponding to the protein YacC. Sequence analysis showed YacC had limited (21%) sequence identity to the *E. coli* pilotin EtpO, and contained a partially conserved PulS\_OutS domain (Figure 2.1). In addition to YacC, the HMM analysis revealed several low confidence (E value greater than  $1.0e-05$ ) hits. A list of the protein hits obtained from the HMM in EPEC can be found in Appendix A2. One of which was the previously uncharacterized protein YghG. The score from the HMM is not statistically significant (E value =  $4.20e-03$ ), but the gene encoding YghG is located within the operon encoding for the T2SS of EPEC. Closer analysis of the sequence of YghG showed the characteristics of a lipoprotein: the bioinformatic tool LipoP predicts a signal peptidase II cleavage sequence between positions 24/25 of YghG, which produces the N-terminal sequence CASHN in the processed mature lipoprotein. For reasons described later, YghG and its homologs will be referred to as AspS (Alternate general secretion protein subunit S).



**Figure 2.1: YacC is a novel member of the PulS-OutS family of proteins.** (A) The conserved domain architecture tool (CDART) was used to define the PulS\_OutS domains of PulS from *Klebsiella oxytoca* 10-5250 (EHT07154.1), OutS from *Dickeya dadantii* 3937 (YP\_003883937.1), EtpO from enterohemorrhagic *E. coli* (EHEC) O157:H7 (CAA70966.1) and YacC (CAS07673.1) from EPEC. Numbers refer to the amino acid length of each protein sequence. The blue bar represents the PulS\_OutS domain found in these proteins. Sequence alignment shows the similarity between EtpO (CAA70966.1) and YacC (CAS07673.1) over 80 residues. Identical residues are highlighted between the two sequences, conserved substitutions are shown (+), and conserved cysteine residues are highlighted in red boxes.

The T2SS of EPEC secretes a mucin degrading enzyme, SslE, which belongs to the M60-like class of zinc metalloproteases, and shares the same domain architecture compared to other Accessory Colonising Factor D (AcfD) like proteins found in species of *Vibrio* and *Shigella*. EPEC, *V. cholerae* and some *Shigella* spp. have recognizable T2SSs, and yet no PulS-OutS like pilotin has been reported previously in these genome sequences. Analysis with the PulS-OutS HMM revealed that no high-scoring sequences were detected in the genome of *V. cholerae* O1 biovar El Tor N16961. There are, however, AspS related sequences. A BLAST search using the AspS sequence from EPEC as a query, the protein sequence VC1703 (NP\_231339.1) was detected in this *V. cholerae* genome and found to have high (52%) sequence similarity to AspS (Figure 2.2a). AspS-related protein sequences were also found in all strains of *V. cholerae* and other species of the genus *Vibrio*, and in *Shigella boydii* ATCC 9905 and *Shigella* sp. D9. All of these bacteria have clearly recognizable operons that would encode a T2SS. In EPEC, *Shigella boydii* ATCC 9905 and *Shigella* sp. D9 the gene encoding AspS is embedded within that operon (Figure 2.2b).

CLuster ANalysis of Sequences (CLANS; performed by Dr. Eva Heinz, Lithgow Laboratory, Monash University) was used to characterize the relationship of the PulS-OutS family of proteins to each other and to the YacC and AspS related proteins detected in BLAST searches. The analysis defined YacC related proteins as being a distinct grouping, and that this group of proteins are related to the PulS-OutS family of pilotins. It also showed that AspS shared no statistically significant relationship to the PulS-OutS family of proteins (Figure 2.3a) suggesting that EPEC encodes two previously uncharacterized proteins: one (YacC) with sequence characteristics common to previously-characterized T2SS pilotins, and another unrelated lipoprotein (AspS).

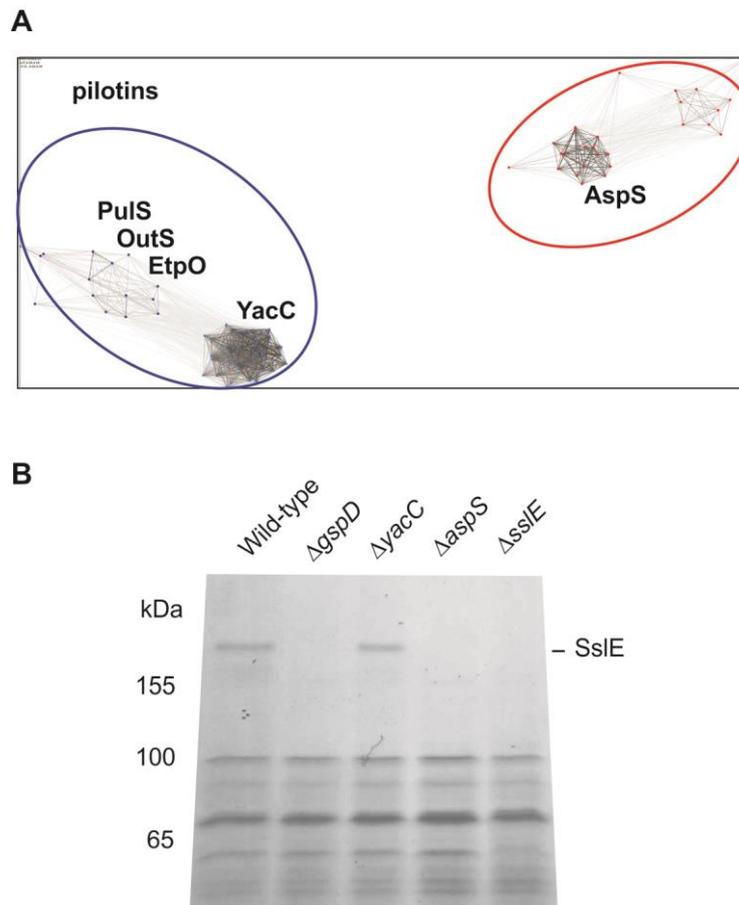


### ***2.3.2 AspS is required for T2SS function***

To determine whether the loss of either YacC or AspS has any phenotypic consequences to the function of the T2SS, single gene deletions (see strain list, Table 2.1) were constructed using the lambda Red recombinase method [142]. The secretion of SslE, the major substrate of the T2SS in EPEC, was monitored in these strains as a read out of T2SS activity. The parental EPEC strain and mutants lacking GspD, SslE, YacC or AspS were grown in CAYE media and the secreted proteins were isolated and analysed by SDS-PAGE and Coomassie blue staining. The identity of SslE was confirmed by mass spectrometry of a ~165 kDa protein band present in the secretome of wild-type EPEC and by its absence from the  $\Delta$ *sslE* mutant. While SslE was present in the secretome of cells lacking YacC, it was not secreted by the  $\Delta$ *aspS* strain (Figure 2.3b). The levels of proteins in the secretome remained unchanged in all strains.

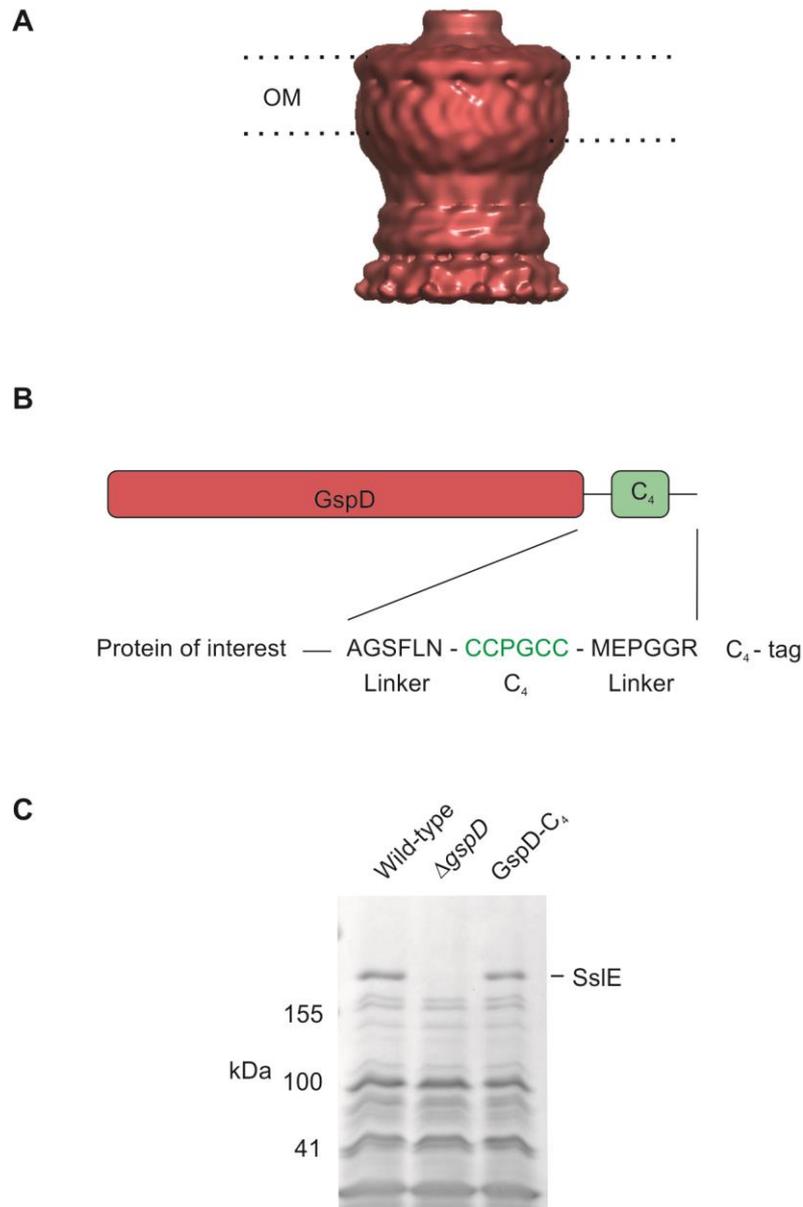
### ***2.3.3 Assays to monitor the kinetics of assembly of secretin GspD in EPEC***

To determine if either YacC or AspS functioned as a pilotin for the assembly of secretin GspD (Figure 2.4a), two additional deletion mutants in EPEC were engineered ( $\Delta$ *gspD* $\Delta$ *yacC* and  $\Delta$ *gspD* $\Delta$ *aspS*). Two GspD expressing plasmids (Table 2.1) were constructed carrying the *gspD* gene from EPEC modified with a C-terminal tetra-cysteine (FLAsH) tag (Figure 2.4b). This tagged GspD is now referred to as GspD-C<sub>4</sub>. The FLAsH tagged constructs of GspD allows the sensitive and selective labelling of GspD monomers and multimers after SDS-PAGE using the Lumio Green Detection Kit (Thermo Fisher Scientific). The FLAsH-tagged GspD was functional, since  $\Delta$ *gspD* mutants expressing GspD-C<sub>4</sub> were able to secrete SslE into the extracellular milieu at wild-type levels (Figure 2.4c).



**Figure 2.3: AspS belongs to a novel group of pilotins and is required for T2SS function.**

(A) CLANS analysis graphically depicts homology in large datasets of proteins using an all-against-all BLAST search. Lines are shown between the most similar sequences, using an E-value cut-off of  $1e-5$ . (B) Wild-type EPEC (WT) and the indicated mutants of EPEC were grown in CAYE media for 4 hours before TCA precipitation of bacterial supernatants. Supernatants were analysed by SDS-PAGE and Coomassie blue staining. Mass spectrometry was used to identify SslE.



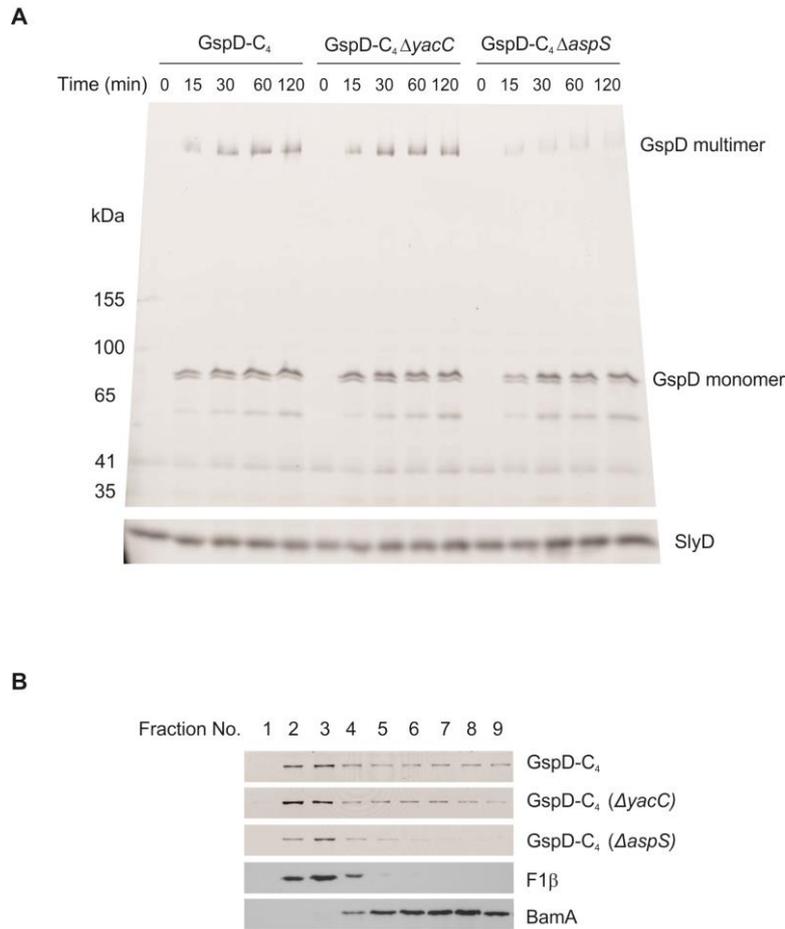
**Figure 2.4: GspD-C<sub>4</sub> is functional and is able to secrete SsIE.** (A) Cryo-EM structure of the *Vibrio cholerae* GspD secretion pore [71]. (B) Schematic of the C-terminal FAsH tagged GspD. (C) Wild-type EPEC (WT),  $\Delta$ gspD mutant EPEC, and the  $\Delta$ gspD mutant EPEC complemented with the plasmid encoding GspD-C<sub>4</sub> were grown in CAYE media for 4 hours before TCA precipitation of bacterial supernatants. Supernatants were analysed by SDS-PAGE and Coomassie blue staining.

The assembly of GspD-C<sub>4</sub> was monitored by expressing GspD-C<sub>4</sub> over time, taking samples at the indicated time points and analysing whole cell lysates by SDS-PAGE. GspD-C<sub>4</sub> expression was observed within 15 minutes of induction with arabinose. The monomeric form of GspD-C<sub>4</sub> is detected at early time-points, and multimers of GspD-C<sub>4</sub> form with a slight delay in kinetics (Figure 2.5a). The endogenous metallo-chaperone SlyD which binds to the Lumio reagent (Lumio Green Detection Kit user manual, Publication part number 25-0672) was used as an internal loading control. EPEC mutants lacking YacC assembled GspD-C<sub>4</sub> with the same kinetics as the complemented ‘wild-type’ strain. While monomeric GspD-C<sub>4</sub> was observed in  $\Delta aspS$  mutants, assembly into GspD-C<sub>4</sub> multimers was greatly reduced in the absence of AspS (Figure 2.5a).

Sucrose density fractionation was used to demonstrate that the multimers of GspD-C<sub>4</sub> were selectively present in the outer membrane. GspD-C<sub>4</sub> multimers were detected in the outer membrane fractions of ‘wild-type’ and  $\Delta yacC$  mutants, and were not present in the outer membranes of EPEC lacking AspS (Figure 2.5b). However, use of this EPEC system for sub-cellular fractionation proved to be non-ideal, as the over-expression of GspD-C<sub>4</sub> lead to ill-defined amounts of GspD-C<sub>4</sub> multimers in the inner membrane fractions. We hypothesized that this was due to the absence of enough pilotin molecules to transport all of the synthesised GspD to the outer membrane.

### ***2.3.4 Overexpression of AspS pilots GspD to the OM***

A second assembly assay was established to allow the dual expression of GspD with or without YacC or AspS using the model *E. coli* strain BL21 (DE3). Both *gspD* and *aspS* were knocked out in this background strain using the lambda Red recombinase and complemented

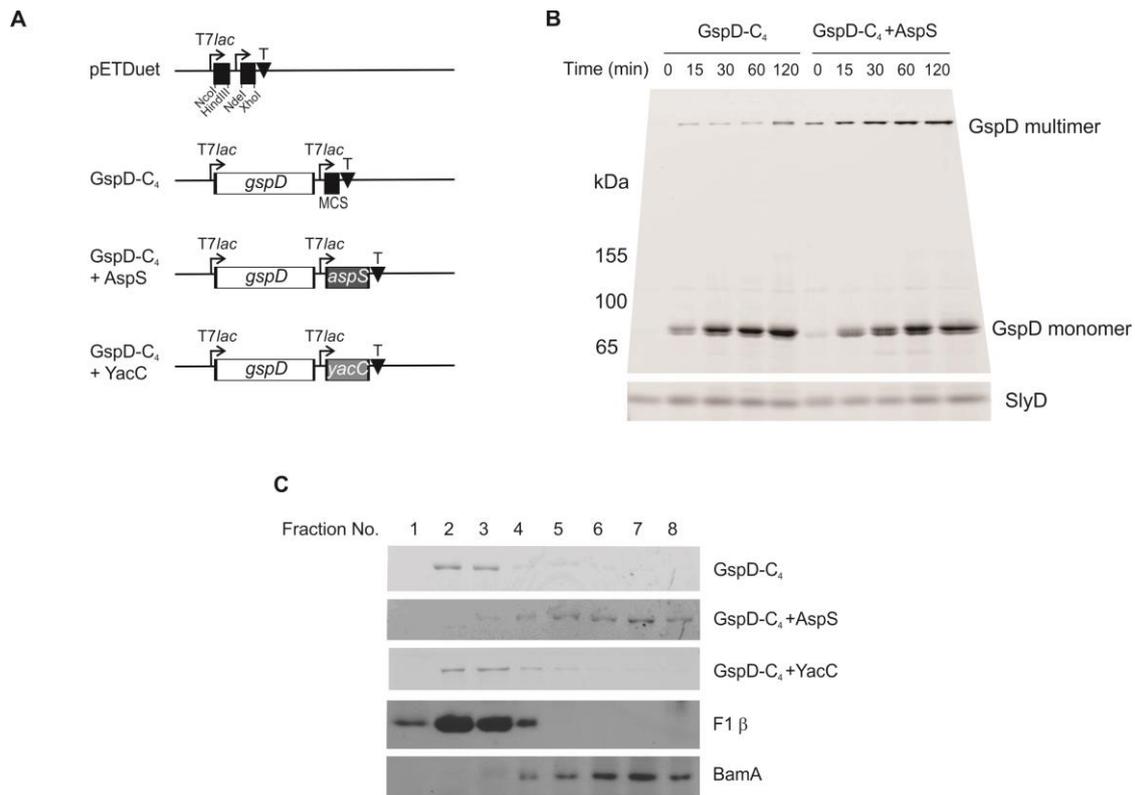


**Figure 2.5: Assay to measure the *in vivo* assembly of GspD in EPEC.** (A) The indicated strains of EPEC, harbouring the GspD-C<sub>4</sub> expressing plasmid were cultured in Luria Broth to an OD<sub>600</sub> ~ 0.6 and arabinose was then added to the culture (0.1%, final concentration). At the indicated time-points cell extracts were prepared from the culture, resuspended in sample buffer containing Lumio reagent and analysed by SDS-PAGE. The polyacrylamide gels were then imaged by fluorimetry. The 21 kDa protein SlyD was used as a loading control. (B) EPEC strains lacking *gspD*, *gspDyacC* or *gspDaspS*, were complemented with the plasmid encoding GspD-C<sub>4</sub> under control of the *tet* promoter and were cultured to an OD<sub>600</sub> ~1.0. Total membranes were isolated and then fractionated by sucrose density centrifugation. Fractions were analysed by SDS-PAGE for detection of GspD multimers with Lumio reagent and immunoblotting for the outer membrane protein BamA and the inner membrane F1β.

with pETDuet-1 vectors containing the coding sequences for GspD-C<sub>4</sub>, with or without putative pilotins YacC or AspS (Figure 2.6a). In conjunction with the results observed in EPEC, low amounts of GspD-C<sub>4</sub> multimers were observed in the absence of AspS in BL21 (DE3) cells (Figure 2.6b). However a much more rapid assembly of the GspD multimers were observed in cells over expressing AspS. Importantly, sucrose density gradients revealed that in the absence of AspS all detectable GspD-C<sub>4</sub> was associated with the inner membrane, co-migrating with inner membrane protein F<sub>1</sub>β (Figure 2.6c). This is consistent with previous observations that the *K. oxytoca* secretin PulD assembles into the inner membrane in the absence of its pilotin PulS [107]. In cells over expressing AspS, all of the multimeric GspD-C<sub>4</sub> was detected in the outer membrane fractions (Figure 2.6c). It can be concluded that AspS is the pilotin for the EPEC secretin GspD. Conversely, co-expression of YacC had no effect on the assembly or trafficking of GspD into the outer membrane. Bioinformatics analyses do not detect a second T2SS secretin encoded in the EPEC genome (data not shown), leaving the possibilities that YacC functions as a pilotin for an unrelated group of secretins or other membrane proteins found in other *E. coli* strains [127,145], or performs an entirely different function.

### ***2.3.5 The structure of AspS distinguishes it from the PulS-OutS family of proteins and interacts with the S-domain of its cognate secretin***

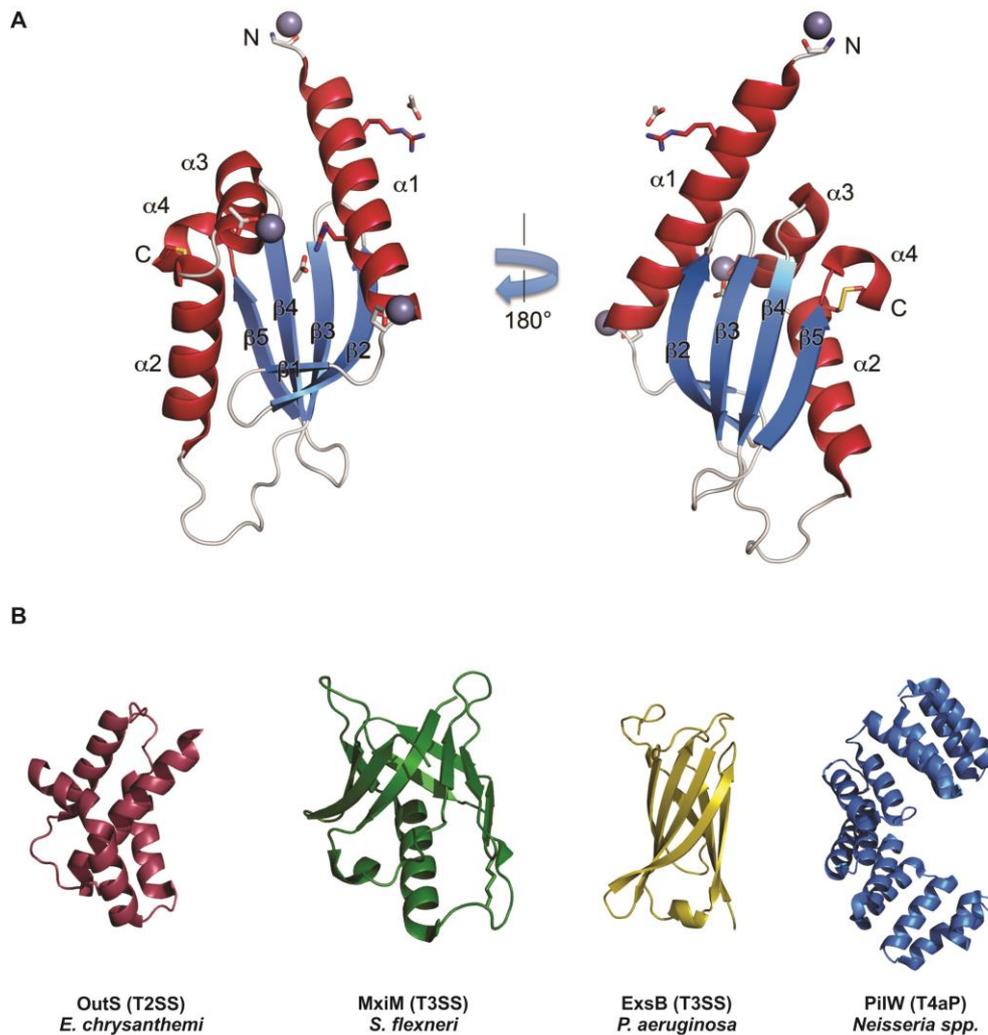
The AspS pilotin from *V. cholerae* (residues 6-114, numbered from the acylated Cys<sup>1</sup>) was expressed and purified in soluble form from the periplasm of *E. coli* str. Rosetta (DE3). AspS from *V. cholerae* yielded crystals diffracting to high resolution allowing the structure to be solved to 1.48 Å. AspS expression, purification and structure determination was performed by Assistant Professor Konstantin Korotkov (University of Kentucky).



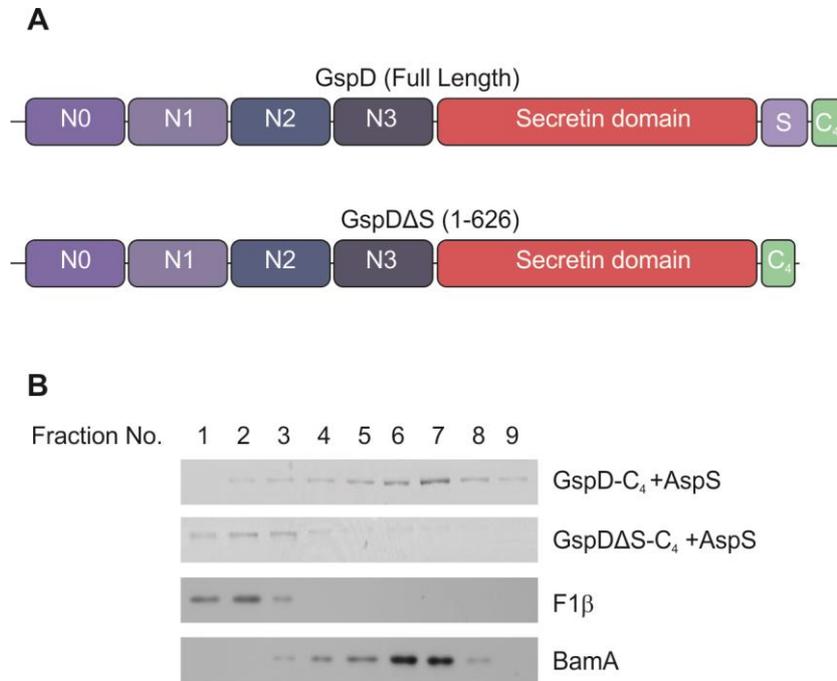
**Figure 2.6: GspD requires AspS for targeting to the outer membrane.** (A) The expression cassette in pETDuet plasmids GspD-C<sub>4</sub>, GspD-C<sub>4</sub>+AspS and GspD-C<sub>4</sub>+YacC are represented diagrammatically. The pETDuet-1 vector (Novagen) has two multi-cloning sites (MCS) represented as black squares, NcoI and HindIII sites were used to clone the open-reading frame corresponding to GspD-C<sub>4</sub> and NdeI and XhoI sites were used to clone the open-reading frame corresponding to AspS or YacC. The T7 terminator sequence (T) in the plasmid is represented by a black triangle. (B) *E. coli* BL21 (DE3) ( $\Delta gspD\Delta aspS$ ) harbouring GspD-C<sub>4</sub>, GspD-C<sub>4</sub>+AspS or GspD-C<sub>4</sub>+YacC expressing plasmids were cultured to an OD<sub>600</sub> of ~ 0.6 before the addition of 0.1 mM IPTG to the culture. At the indicated time-point cell extracts were prepared from the cultures using Lumio, analysed by SDS-PAGE and imaged by fluorimetry. (C) The strains described above were cultured to an OD<sub>600</sub> ~1 before the isolation of total membranes and sucrose density centrifugation. Fractions were analysed by SDS-PAGE for detection of GspD multimers with Lumio reagent and immunoblotting for BamA and F<sub>1</sub>β.

The AspS structure forms an  $\alpha/\beta$  domain consisting of a 5-stranded  $\beta$ -sheet and 4  $\alpha$ -helices (Figure 2.7a). The N-terminal helix  $\alpha_1$  is followed by antiparallel  $\beta$ -strands  $\beta_1$ ,  $\beta_2$  and  $\beta_3$ . The helices  $\alpha_2$  and  $\alpha_3$  are orientated across  $\beta$ -strands  $\beta_4$  and  $\beta_5$ , which are followed by the C-terminal helix  $\alpha_4$ . Two conserved cysteine residues, Cys<sup>74</sup> and Cys<sup>111</sup>, form a disulfide bond between helix  $\alpha_4$  and helix  $\alpha_2$ . The structure of AspS is distinct from the previously characterised pilotins of the T2SS, T3SS and T4P (Figure 2.7b).

The difference in structure of AspS to other characterised pilotins raised the question of where in the GspD protein was the binding site for AspS. To determine whether the S-domain of GspD was required for AspS targeting to the outer membrane the pETDuet-1 system described above, was used co-express a truncated form of GspD (GspD $\Delta$ S) lacking the predicted S-domain with AspS (Figure 2.8a). Fractionation of membranes from *E. coli* str. BL21 (DE3) expressing full length or truncated forms of GspD with AspS showed that GspD $\Delta$ S was not delivered to the outer membrane (Figure 2.8b). Thus, the piloting function of AspS depends on the presence of the S-domain.



**Figure 2.7: AspS provides a novel structure for pilotin like proteins.** (A) Ribbon representation of the structure of *V. cholerae* AspS.  $\alpha$ -helices are in crimson and  $\beta$ -strands, are in light blue.  $Zn^{2+}$  ions are shown as grey spheres. Acetate ions are shown in stick representation. Residues coordinating  $Zn^{2+}$  and acetate ions are in stick representation with oxygen and nitrogen atoms colour-coded red and blue, respectively. The position of the disulphide bond between Cys74-Cys111 is shown in yellow. (B) Ribbon structures of representative members of the three characterized classes of pilotins OutS (PDB: 3UTK, [100]), MxiM (PDB: 1Y9L, [104]), ExsB (PDB: 2YJL, [105]) and PilW (PDB: 2VQ2, [103]).



**Figure 2.8: The S-domain is required for AspS dependant targeting of GspD to the outer membrane.** (A) Schematic showing the domains of full length and a S-domain truncation mutant of GspD. (B) *E. coli* BL21 (DE3) ( $\Delta gspD, \Delta aspS$ ) harbouring GspD-C<sub>4</sub>+AspS or GspD $\Delta$ S-C<sub>4</sub>+AspS were cultured to an OD<sub>600</sub> of ~ 0.6 and IPTG was added to the culture (0.1 mM, final concentration). Total membranes were isolated and separated by sucrose density centrifugation. Fractions were analysed by SDS-PAGE for detection of GspD multimers with Lumio reagent and immunoblotting for BamA and F1 $\beta$ .

## **2.4 DISCUSSION**

### **2.4.1 Targeting secretins to the outer membrane**

A “piggyback” model has been proposed for the targeting of secretin PulD to the outer membrane of *K. oxytoca* by its cognate pilotin PulS. This targeting first relies on the selective binding of the S-domain of the secretin by its cognate pilotin at the periplasmic face of the inner membrane, and secondly, the Lol dependant translocation of the pilotin-secretin heterodimer to the outer membrane. This has also been observed for secretins of the T3SS and T4aP that also engage a pilotin for translocation to the outer membrane [97,146,147]. It was considered widely in the field that only members of the PulS-OutS family of proteins would function as pilotins for T2SS secretins. However, in organisms like *V. cholerae* and *P. aeruginosa* that contain very clear T2SS operons, no obvious PulS-OutS like pilotins have been identified in their genomes. In the case of *V. cholera* and *A. hydrophila*, accessory factors ExeA (GspA) and ExeB (GspB) are crucial for the targeting of secretin ExeD (GspD) to the outer membrane [109,110]. ExeAB forms a complex at the inner membrane in which ExeA binds peptidoglycan and ExeB interacts with the NON1 periplasmic domains of secretin ExeD [111,112]. This translocation process by which ExeAB mediates ExeD translocation and assembly is unknown, but it appears that this process is functionally distinct compared to the ‘piggyback’ model. ExeAB may involve the remodelling of the peptidoglycan to aid the pilotin dependant targeting of ExeD to the outer membrane.

The targeting of GspD to the outer membrane by AspS is in line with the requirements of the piggyback model with the presence of the S-domain of GspD being crucial for its translocation to the outer membrane. This interaction was also shown directly in an *in vitro*

binding assay where the S-domain of GspD fused to a maltose-binding protein carrier bound with AspS (Experiment performed by Assistant Professor Konstantin Korotkov, [145]). A study by Strozen et al., (2012) also showed the pilotin function of AspS, with *aspS* mutants of enterotoxigenic *E. coli* showing decreased levels of GspD and a reduced ability to secrete LT into the extracellular milieu [148]. Moreover, Strozen et al., (2012) showed that AspS is a lipoprotein with mutants of AspS (S25A, A26D or S27D) no longer able to be processed into a mature form or translocate to the outer membrane respectively [148]. These Lol evasive mutants of AspS also resulted in the loss of both monomeric and multimeric GspD and a reduction in LT secretion [148].

#### **2.4.2 Secretin/Pilotin interactions**

With the growing amount of sequence and structural data, it is becoming increasingly clear just how variant are the pilotins from different bacterial strains and different secretion machines [101]. The structures of several pilotins have now been solved and show vast structural differences between them (Figure 2.7a, b). For example the structures PulS, OutS and EtpO all show an all-helix bundle [149,100,150], whereas, MxiM and ExsB display cracked  $\beta$ -barrel and  $\beta$ -sandwich folds respectively [104,98,105]. PilW and PilF are different again displaying an all-helical tetratricopeptide repeat folds [103,151,147]. The sequence and structural differences of AspS to other pilotins adds further insight and complexity to the characterization of pilotin classes.

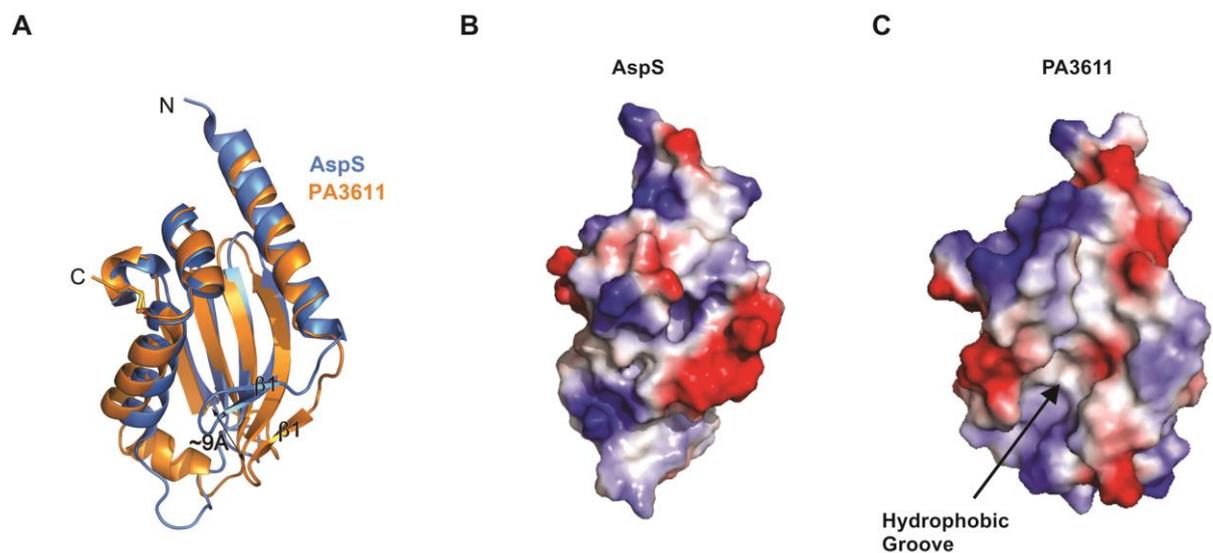
Structural studies have given insight into the mechanism of interaction of the pilotin with the natively disordered S-domain of its cognate secretin [103,152,153]. Given the novel structure of AspS it is tempting to speculate the nature of its interaction with the S-domain of GspD.

Using PDBeFold [154] a structural homolog of AspS was identified in *P. aeruginosa*. PA3611 shares 18% sequence identity for 96 aligned residues with AspS and also contains the conserved disulphide bond. The PA3611 structure features an extra  $\alpha$  helix after  $\beta$ -strand  $\beta$ 3 and shows a separation of  $\beta$ -strands  $\beta$ 1 and  $\beta$ 2 away from helix  $\alpha$ 2 compared to AspS (Figure 2.9a, [155]). This open outward conformation of  $\beta$ -strands  $\beta$ 1 and  $\beta$ 2 leads to the formation of a crevice on the surface of PA3611 compared to AspS (Figure 2.9b, c). It is plausible that such a shift of  $\beta$ -strands  $\beta$ 1 and  $\beta$ 2 in AspS will expose a similar groove on the protein surface which may suggest a possible mechanism of binding the S-domain of its secretin. But this structural analysis needs to be clarified experimentally.

Closer analysis of the primary sequence of PA3611 revealed that unlike AspS, PA3611 contains a type I signal sequence that would target the protein to the periplasm, without the sequence characteristics to suggest that it is a lipoprotein. This observation is particularly interesting as the secretin HxcQ from *P. aeruginosa* is itself a lipoprotein, capable of self-targeting to the outer membrane [116]. Whilst the function of PA3611 is unknown, it is tempting to speculate that PA3611 binds to lipoprotein secretins such as HxcQ to stabilize and protect the secretin monomer from proteolysis.

### ***2.4.3 What is the role of YacC?***

It is still of great interest to determine the role of YacC and address whether it is a pilotin for another secretin or outer membrane protein in *E. coli*. HMM analysis performed by Nermin Celik (Lithgow Laboratory, Monash University) revealed the presence of 3 secretins, GspD (T2SS), EscC (T3SS) and RcpA/geneIV (Flp pilus or filamentous phage extrusion) in the



**Figure 2.9: Structural homology between AspS and PA3611.** (A) Overlay of AspS (blue) and PA3611 (orange) structures. (B) Electrostatic surface potential of the AspS structure (positive = blue; negative = red) showing a closed conformation. (C) Electrostatic surface potential of the PA3611 structure (positive = blue; negative = red) showing the presence of a hydrophobic groove.

genome of EPEC and an additional secretin BfpB (type IV bundle-forming pili) on the EPEC virulence plasmid pMAR2. Interestingly, BfpB belongs to the very small group of secretins (like HxcQ) that are themselves lipoproteins and are self-targeted to the outer membrane, presumably by the Lol pathway. Biochemical studies have shown that BfpB interacts with a small accessory protein BfpG in the periplasm and that the presence of BfpG is crucial for either the assembly or stability of the BfpB secretion pore [115]. However, a classical pilotin has not yet been identified for EscC and RcpA/geneIV and it is possible that YacC may function as the pilotin for these secretins.

Genetic and biochemical studies have identified two distinct T2SSs in several *E. coli* strains [127,128]. These different T2SSs referred to as T2SS $\alpha$  (i.e. *Klebsiella* type) and T2SS $\beta$  (i.e. *Vibrio* type) are defined based on their secretin, pilotin and substrate characteristics [145,148]. Whilst EPEC contains a complete T2SS $\beta$  operon, it has a severely truncated T2SS $\alpha$  operon (only containing *gspO*). On the other hand *E. coli* lab strain K12 and other pathogenic isolates (ETEC H10407, UPEC UT189, APEC O1 and ExPEC IHE3034) encode a complete T2SS $\alpha$  with the exception of a gene encoding a noticeable PulS-OutS like pilotin. Interestingly, all of these strains encode *yacC*. Whilst AspS is the cognate pilotin for T2SS $\beta$ /*Vibrio*-type secretins, it is plausible that YacC is the cognate pilotin for secretins belonging to the T2SS $\alpha$ /*Klebsiella*-type family from *E. coli*.

## CHAPTER 3: Depletion of BamA.

### 3.1 INTRODUCTION

BAM complex function depends on the essential Omp85 protein BamA and, in *E. coli*, several strains have been engineered to conditionally repress expression of the *bamA* gene. Depletion of BamA impacts the steady-state levels of integral membrane proteins of  $\beta$ -barrel topology, including proteins like LamB, OmpA, OmpF [11,156], as well as  $\beta$ -barrel protein of more complex topologies such as TolC [156], fimbrial ushers [157], autotransporters [158], trimeric autotransporters [159] and intimin [160]. Conditional BamA depletion strains have also been engineered in other bacterial lineages, including *Pseudomonas* [119,161], *Borrelia* [162] and *Thermus* [163].

Past studies on the effect of BamA on secretin assembly have used steady-state levels of the secretin as a proxy measure for expression [44,118,119]. Given that secretins and other multimeric secretion pores are typically very stable in an assembled form, the assembly of newly synthesized secretion channels might not be measurable against a background of pre-assembled pores during the time required to deplete *bamA*. In this chapter the optimal depletion time course will be established wherein *bamA* levels are depleted, OMP assembly is significantly reduced yet bacterial cell viability has not yet been impacted. The interactions between BAM lipoproteins at the outer membrane after BamA depletion will also be analysed.

## **3.2 METHODS and MATERIALS**

### **3.2.1 Strains and Growth Conditions**

A BamA depletion strain of *E. coli* MC4100A, WT *E. coli* BW25113 and BW25113  $\Delta$ *bamC* were used to monitor outer membrane protein assembly. Strains were grown in Luria Broth (LB) supplemented with the appropriate antibiotics (kanamycin 30  $\mu$ g/ml) and 0.02 % (w/v) arabinose or 0.02 % (w/v) glucose if needed at 37 °C.

### **3.2.2 Measuring BamA depletion**

To express BamA, cells were grown in LB supplemented with 0.02 % (w/v) arabinose. BamA levels can be repressed by growing cells in LB containing 0.02 % glucose. To measure the levels of BamA, 1 ml samples were taken during 3 h of BamA expression or depletion and cells were harvested by centrifugation. Cell pellets were resuspended in SDS sample buffer and 15  $\mu$ l of whole cell lysates were analysed by 10% SDS-PAGE and immunoblotting against BamA (1:2500) and F<sub>1</sub> $\beta$  (1:8000). For all experiments requiring the depletion of BamA, cells were grown as described above prior to further assembly assays or cellular fractionation.

### **3.2.3 Fluorescence Microscopy**

Wild-type MC4100A and BamA depletion strains were grown as described in Chapter 3.2.1 and after 45 min, 2h 45 min and 4 h 45 min post BamA depletion 300  $\mu$ l aliquots of cells were pelleted, washed in saline (0.9% NaCl) and resuspended in 300  $\mu$ l saline + 3  $\mu$ l of LIVE/DEAD BacLight (Thermo Fisher Scientific). Cultures were incubated at 37°C for 15

min and 2.5  $\mu$ l of cell suspension were spotted onto microscope slides. Images were obtained using an Olympus microscope (IX81 inverted Olympus) using the FitC and TritC filters. Data was analysed using Cell M (Olympus) and ImageJ.

#### ***3.2.4 Isolation of total membranes after BamA expression or depletion***

Cells were grown in LB media supplemented with 0.02 % arabinose for BamA expression or 0.02 % glucose for BamA depletion for 3 h at 37 °C. Cells were then harvested and total membranes were prepared as described in Chapter 2.2.8.

#### ***3.2.5 Sucrose density fractionation***

The separation of bacterial inner and outer membranes prepared in Chapter 3.2.4 was performed as described in Chapter 2.2.9. Fifteen micro litres of each fraction were analysed by SDS-PAGE and subsequent western immunoblotting for F<sub>1</sub> $\beta$  and BamA (serum dilutions outlined in Table 3.2).

#### ***3.2.6 Transmission Electron Microscopy***

MC4100A BamA depletion strains were grown in 30 ml of LB as described above. After 3 h of depletion, cells were harvested by centrifugation (4000 x g, 5 min) and fixed in 300  $\mu$ l of fixation solution (2.5% glutaraldehyde and 0.1 M sodium cacodylate) for 60 min. Cells were collected and washed twice with 0.1 M sodium cacodylate followed by two washes with 1% osmium tetroxide. Cells were washed with distilled water and dehydrated by subsequent washes with increasing concentrations of acetone (50, 70, 90 and 100%). Resin was infiltrated into the samples with successive washes with EPON (25, 50, 75 and 100%) and

allowed to polymerise for 48 h at 60° C. All washes were performed in a BioWave microwave (Pelco). 70 nm thick sections were cut and stained with 2% Uranyl acetate for 15 minutes followed by 3 minutes in Reynold's Lead citrate solution [164]. Sections were imaged using a Hitachi H7500 transmission electron microscope and analysed using Digital Micrograph (Gatan).

### **3.2.5 SDS-PAGE**

See Chapter 2.2.5.

### **3.2.8 Blue Native-PAGE (BN-PAGE)**

Native protein complexes were resolved by BN-PAGE with BN-gel buffer (200 mM n-amino caproic acid, 150mM Bis Tris pH 7.0) using BN-PAGE cathode buffer (50 mM Tricine, 15 mM Bis-Tris pH 7.0, 0.02% Brilliant Blue G250, 0.03% DDM) and BN-PAGE anode buffer (50 mM Bis-Tris pH7.0). The protocol required to make one BN-PAGE gel is described in Table 3.1 using a BN-PAGE specific SG50 Hoefer gradient pourer (to prevent SDS contamination).

Protein samples were prepared by adding ACA750 buffer (750 mM n-amino caproic acid, 50 mM Bis-Tris 0.5 mM EDTA pH 7.0; to a final volume of 18 µl) to 50 µg of total membranes and solubilised with 2 µl 10% DDM for 20 min on ice with gentle mixing. Samples were cleared by centrifugation (20,000 x g, 10 min, 4 °C) to pellet any non-solubilised membranes. The supernatant was added to 5 µl 5 x BN-sample buffer (5% Brilliant blue G, 500mM amino caproic acid in 100mM Bis-Tris pH 7.0) prior to loading. BN-PAGE gels were run overnight

at 120 V, 4 °C. The next day the cathode buffer was replaced with fresh buffer without Brilliant Blue G250 and run at 800 W until the detergent front has run off the gel.

**Table 3.1: Solutions required to make 1 x 5-16% gradient BN-PAGE gel.**

	<b>5% Sep</b>	<b>16% Sep</b>	<b>3% Stack</b>
<b>dH<sub>2</sub>O</b>	8.09 ml	1.6 ml	5.89 ml
<b>3 x Gel Buffer</b>	5 ml	5 ml	3.33 ml
<b>Glycerol</b>	-	2.8 ml	-
<b>Acrylamide (40% 37.5:1; Biorad)</b>	1.88 ml	6 ml	750 µl
<b>APS (20% w/v)</b>	25 µl	25 µl	27 µl
<b>TEMED</b>	4 µl	4 µl	6 µl

### **3.2.9 Immunoblotting**

Proteins were transferred from SDS-PAGE and BN-PAGE gels to nitrocellulose and activated PVDF membranes respectively in transfer buffer as described in Chapter 2.2.6. Prior to the transfer of a BN-PAGE gel; the gel was first incubated in 1x SDS running buffer for 10 min to denature proteins in the gel. After the transfer, the membrane was rinsed in destain and then 100% methanol to remove residual coomassie and to allow the visibility of protein markers. The membrane was blocked and probed with antibodies for detection as per Chapter 2.2.6. The antibodies used in this chapter are described in Table 3.2.

**Table 3.2: Antibodies used in Chapter 3**

<b>Antibody</b>	<b>Serum Dilution</b>	<b>Source</b>
BamA peptide TVDTDTQRVPGSPDC (Rabbit)	1:2500	Lithgow Laboratory (Felicity Alcock)
BamA POTRA 1-5 (Mouse)	1:20000	Buchanan Laboratory [165]
BamB (Mouse)	1:30000	Buchanan Laboratory [165]
BamC (Mouse)	1:30000	Buchanan Laboratory [165]
BamD (Mouse)	1:30000	Buchanan Laboratory [165]
BamE (Mouse)	1:5000	Buchanan Laboratory [165]
F1 $\beta$ (Mitochondrial; Rabbit)	1:8000	Lithgow Laboratory [166]
TamA (Rabbit)	1:5000	Lithgow Laboratory [61]
Goat anti-Mouse IgG, HRP conjugate	1:20000	Sigma A4416
Goat anti-Rabbit IgG, HRP conjugate	1:20000	Sigma A6154

**3.2.10 BamC Co-immunoprecipitation**

Cells (50 ml) were grown as described in Chapter 3.2.1 for 3 hours before harvesting by centrifugation (15 min, 1500 x g, 22 °C). Cells were washed in 25 ml PBS and pelleted by centrifugation (15 min, 1500 x g, 22 °C) before being resuspended in 10 ml of PBS. Cells were spheroplasted with 10 ml of ice cold spheroplasting buffer (0.75 M sucrose, 50 mM Tris, pH 7.8, 0.6 mg/ml lysozyme, 6 mM EDTA) and incubated at room temperature for 90 minutes or until complete spheroplasting could be observed. Cells were lysed with 1% Triton X-100 and incubated at 4 °C for 15 min. Add 15 mM MgCl<sub>2</sub> and remove unlysed cells by

centrifugation (15 min, 6000 x g, 4 °C). Fifteen ml of the cleared lysate was added to 3 µl of BamC antibody (see table 3.2) and incubated for 1 h at room temperature with gentle mixing. Whilst the lysate and antibody mixture are incubating, 60 µl protein G-Sepharose beads was washed with 500 µl IP buffer (Sigma-Aldrich) by centrifugation (2 min, 6000 x g, 4 °C). The supernatant was removed by careful pipetting before adding the lysate:antibody solution and the mixture was incubated for 1 h at room temperature on a rotary wheel. The beads were pelleted by centrifugation (2 min, 6000 x g, 4 °C) and the supernatant discarded before washing the beads 3 times with 500 µl of IP buffer and once with 1 ml PBS. Proteins were eluted from the beads by adding 50 µl of 2x SDS-sample buffer and boiled for 10 min. Beads were pelleted as described above and 20 µl of the supernatant was analysed by SDS-PAGE and western blotting for each component of the BAM complex (antibodies listed in table 3.2).

### ***3.2.11 In-gel digestion of proteins and LC-MS/MS***

See Chapter 2.2.11.

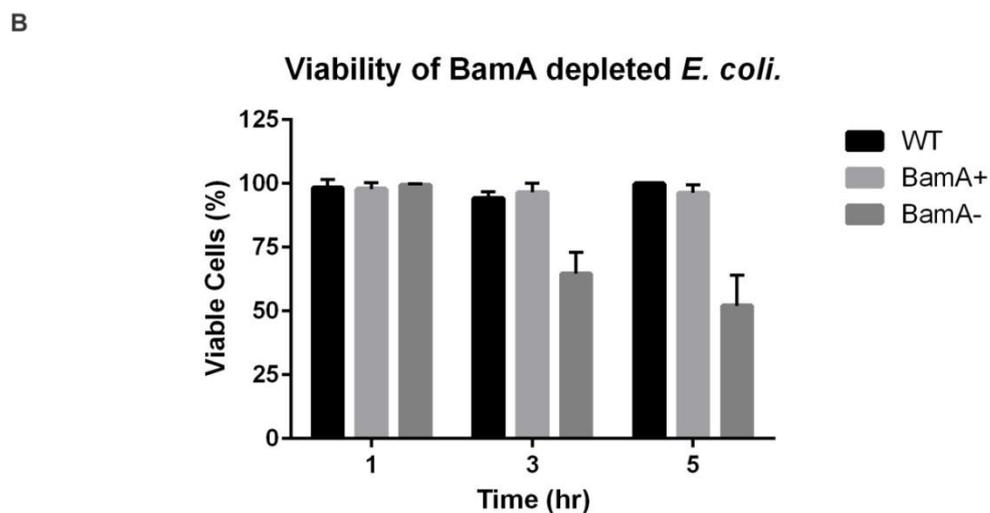
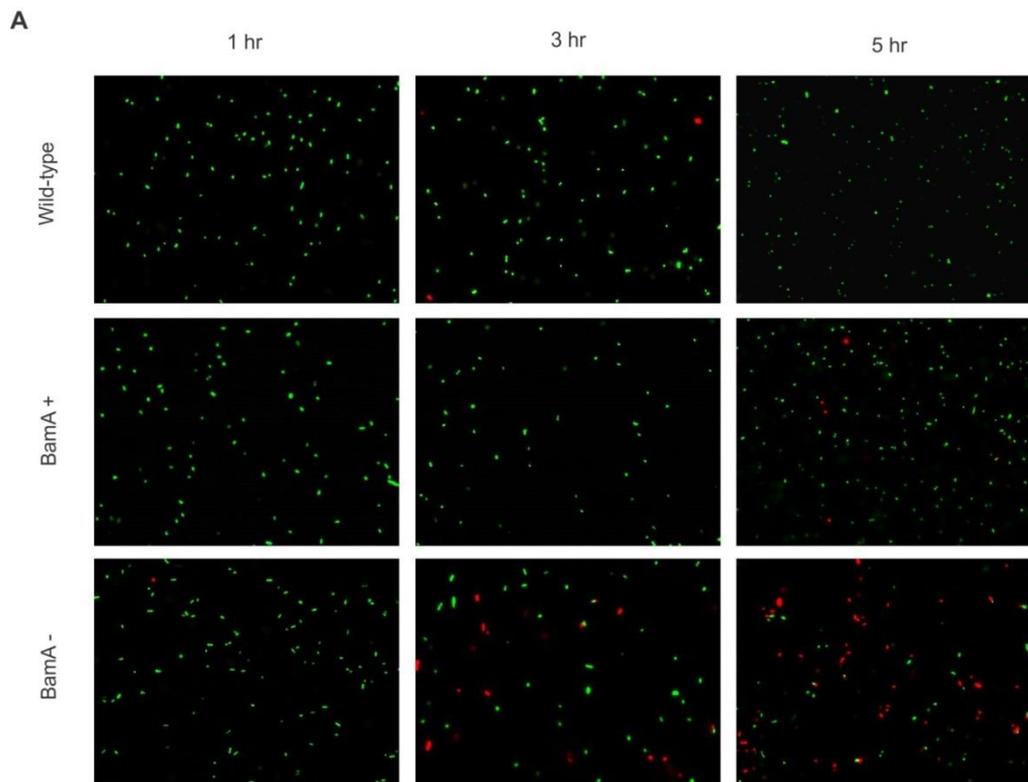
### **3.3 RESULTS**

#### **3.3.1 BamA depletion leads to growth defects and membrane damage**

The *bamA* gene is essential for cell viability, and depletion of BamA prevents the assembly of  $\beta$ -barrel proteins in *E. coli* [11,156-160]. We utilized a similar BamA conditional depletion strain in an *E. coli* MC4100A background to test the role of BamA in assembling multimeric outer membrane proteins [159] (Figure 3.1a). Expression of BamA was induced with 0.02% arabinose (BamA+) and depleted with 0.02% glucose (BamA-). Cell growth was monitored (Figure 3.1b) and cell extracts were prepared at time points after the addition of arabinose or glucose and subject to immunoblotting with antibodies recognising BamA or the control protein F<sub>1</sub> $\beta$  (Figure 3.1c). BamA levels remained constant for 3 h in the cells grown in arabinose, while a depletion of BamA is seen within 1 h in the cells grown in glucose containing medium.

Cell viability was monitored by fluorescence microscopy using the LIVE/DEAD BacLight reagent (Thermo Fisher Scientific) (Figure 3.2a). Cells with intact membranes stain green (SYTO9), whilst dying cells with compromised membranes stain red (propidium iodide). Wild-type and BamA+ cells remained green throughout 5 h of growth, with the BamA- cells progressively becoming red indicating increasing cell death over time, with 65% viable cells after 3 hours and 50% viable cells after 5 h of depletion (Figure 3.2b). Thus, 3 hours of growth in the presence of glucose was established as a time-point at which BamA is undetectable by western blot, while approximately two-thirds of the cell population remains viable.





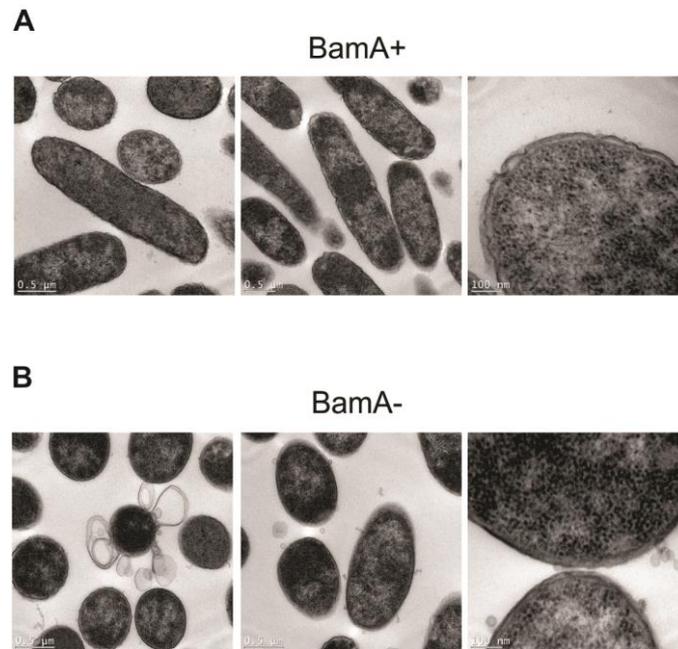
**Figure 3.2: Depletion of BamA leads to a loss in cell viability.** (A) The LIVE/DEAD BacLight (Life Technologies) was used to monitor cell viability in BamA depleted cells (green = viable cells, red = dead cells). (B) Graphical representation of the percentage viable cells counted from 6 independent images from each time point (error bars represent standard deviation of the mean).

Given the diminished levels of  $\beta$ -barrel proteins in the BamA depleted cells, we sought to assess the morphology of the cells. Transmission electron microscopy of cultures after growth in glucose revealed no obvious changes in morphology with a small proportion (~10%) of cells exhibiting blebbing of the outer membrane (Figure 3.3a, b).

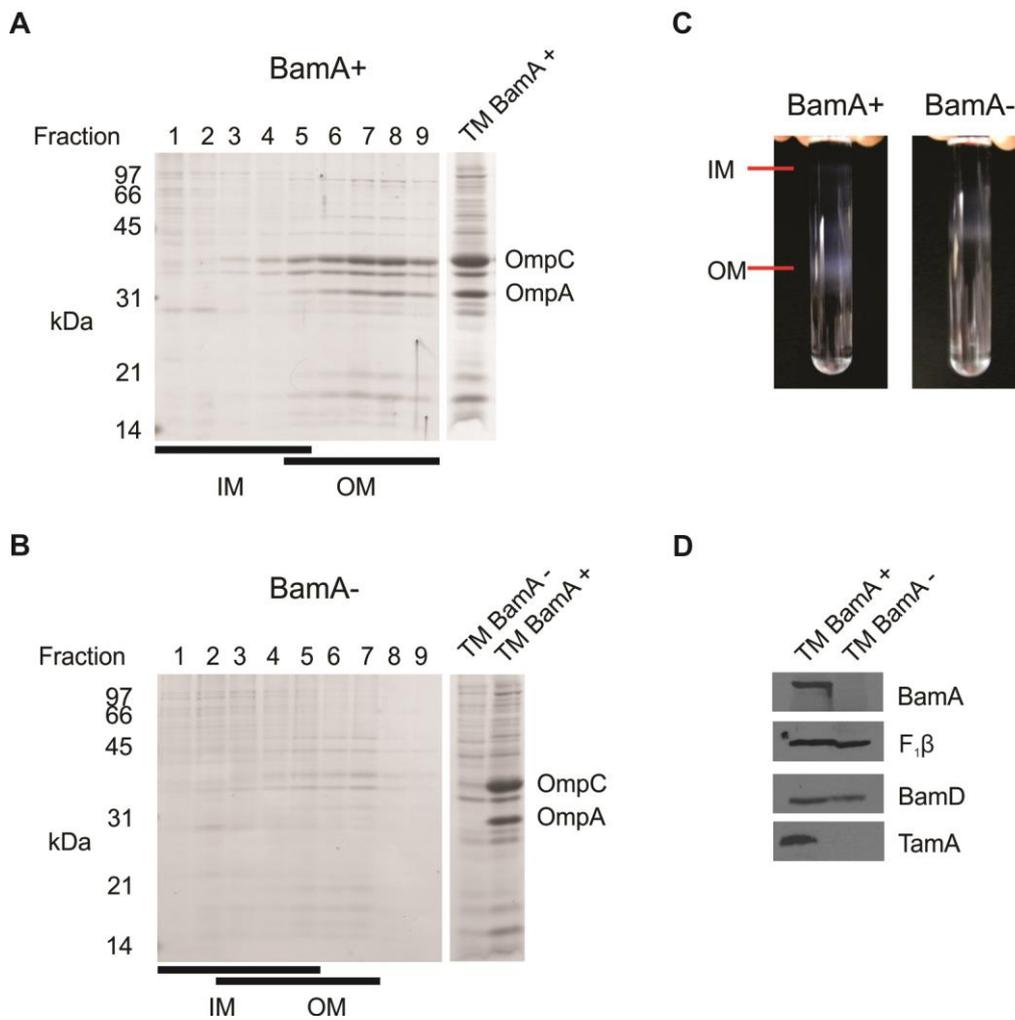
### ***3.3.2 Depletion of BamA inhibits the assembly of $\beta$ -barrel proteins***

Given that  $\beta$ -barrel proteins are substrates of the BAM [167-169] and are subject to degradation by periplasmic proteases like DegP if not folded properly [170], inactivation of the BAM complex would lead to a decrease in abundance of the  $\beta$ -barrel proteins within the cell. To confirm that BamA function had been lost we compared the membrane proteome from cells isolated after 3 hours of BamA depletion to un-depleted cells. Total membranes were subjected to sucrose density gradient fractionation and fractions were separated by SDS-PAGE and subsequent Coomassie brilliant blue staining. The levels of major outer membrane proteins OmpC and OmpA in the BamA depletion were reduced to around 5% of the levels in the bamA expression membranes (Figure 3.4a, b), and did not migrate as far into the gradient compared to membranes from BamA expressing cells, suggesting that the membranes had an altered density due to the lower levels of proteins in the outer membrane. This was also observed directly after separation of total membranes by sucrose density fractionation with the cloudy membrane bands showing different patterning from cells grown in the presence of arabinose or glucose (Figure. 3.4c).

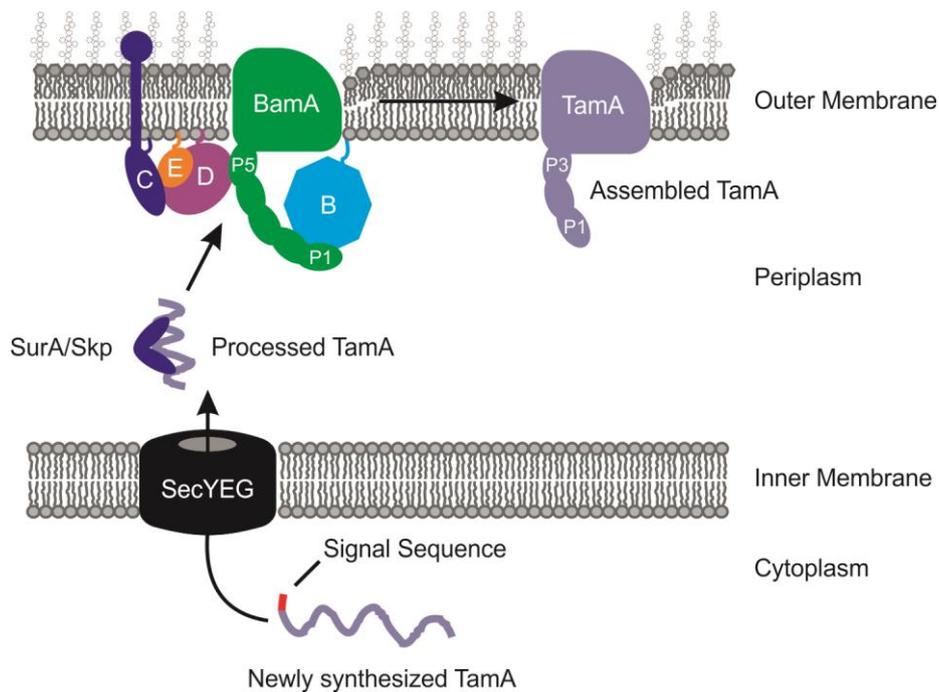
Western blot analysis also showed that TamA was undetectable after 3 hr of BamA depletion, suggesting a master-slave relationship in which BamA mediates the assembly of TamA into the outer membrane (Figure 3.4d, Figure 3.5). The levels of inner membrane protein F<sub>1</sub> $\beta$



**Figure 3.3: Cell morphology of BamA depleted *E. coli*.** Cell morphology of BamA depletion strains were analysed after 3 h of growth in arabinose (BamA+, A) or glucose (BamA-, B) by transmission electron microscopy. Scale bars represent 0.5  $\mu\text{m}$  and 100 nm for higher magnification images.



**Figure 3.4: Membrane proteome of BamA depleted *E. coli*.** Total membranes were isolated from *E. coli* MC4100A BamA depletion strains after 3 h of growth in LB containing arabinose (A) or glucose (B), and subjected to sucrose density fractionation. Fractions were analysed by 10% SDS-PAGE and Coomassie brilliant blue staining. Mass spectrometry was used to identify the abundant proteins OmpC and OmpA. (C) The membrane profile from BamA+ and BamA- membranes immediately after ultracentrifugation of sucrose density gradients. (D) Western blots indicate the levels of BamA, TamA, the lipoprotein BamD and the inner membrane protein F<sub>1</sub>β.



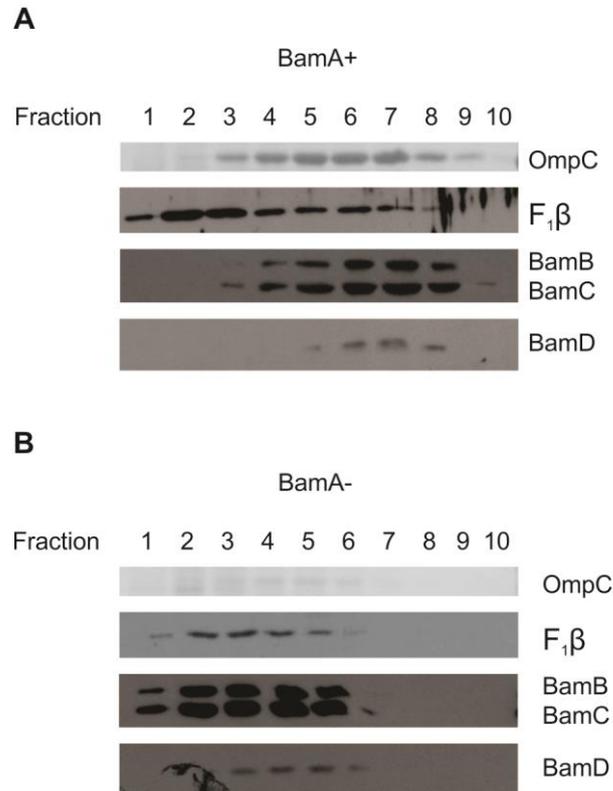
**Figure 3.5: Model of TamA assembly into the outer membrane.** TamA is synthesized as an unfolded pre-protein in the cytoplasm with an N-terminal signal sequence for targeting to the SecYEG translocon. Upon transport through the Sec, the signal sequence is cleaved and the processed TamA is targeted to the  $\beta$ -barrel assembly machinery (BAM) complex for assembly into the outer membrane.

serve as a loading control for membranes in this experiment. Furthermore, the level of the lipoprotein BamD remained unchanged in the extracts, while the migration of BamD to a region of lower density was consistent with the changed migration of the outer membranes (Figure 3.4d).

### ***3.3.3 BamC and BamD remain modular after BamA depletion.***

The BAM complex is considered to be a modular machine, and it has been reported that the BAM complex can be purified as BamAB and BamCDE modules [171,51]. These modules can be reconstituted together *in vitro* into an active form of the BAM “holo-complex” [171,51]. This raised the prospect that depletion of BamA might release the lipoproteins as modules *in vivo*. To address this prospect, sucrose density fractionation was used to demonstrate localisation of BamB, BamC and BamD in BamA<sup>+/-</sup> cells. BamB, BamC and BamD localised to the outer membrane of BamA expressing cells and migrated to an intermediate section of the gradient containing both inner and outer membrane proteins after BamA depletion (Figure 3.6a, b).

To assess the quaternary structure of the BAM components after BamA depletion *in vivo*, total membranes from BamA <sup>+/-</sup> cells were solubilized in DDM and subjected to blue native-polyacrylamide gel electrophoresis (BN-PAGE) and analysed by western blot against antibodies raised against each BAM component (details of antibodies listed in Table 3.2). The mature BAM complex was observed with an apparent migration of approximately 300 kDa in BamA expressing cells and was significantly reduced in BamA<sup>-</sup> membranes (Figure 3.7). Depletion of BamA resulted in the loss of the BamAB sub-complex: BamB was

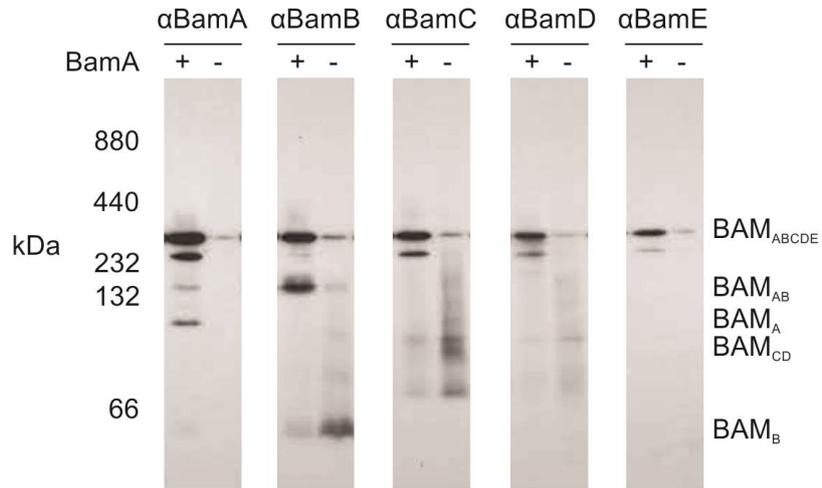


**Figure 3.6: Bam lipoproteins are still targeted to the outer membrane after BamA depletion.** Total membranes were isolated from *E. coli* MC4100A BamA depletion strains after 3 h of growth in LB containing arabinose (A) or glucose (B), and subjected to sucrose density fractionation. Fractions were analysed by 10% SDS-PAGE and Coomassie brilliant blue staining for the identification of OmpC and immunoblotting against BamB, BamC, BamD and F<sub>1</sub>β.

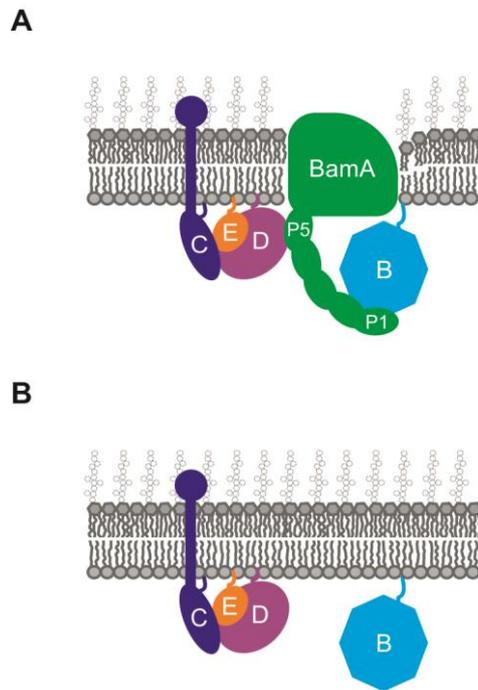
detected instead migrating at approximately 50 kDa, which is the molecular size of a BamB monomer.

The BamC and BamD each migrated as a series of forms in the range 70-100 kDa in membranes isolated from BamA- cells (Figure 3.7, Figure 3.8). This is largely consistent with the size of BamC (35 kDa), BamD (25 kDa) and BamE (12 kDa). As an independent assessment of whether BamC and BamD interact after depletion of BamA, co-immunoprecipitation assays using antibodies raised against BamC were established (Figure 3.9). In the WT BW25113 and MC4100A BamA expressing cells all of the BAM components could be identified after immunoprecipitation with BamC antibodies. The BW25113 *bamC* mutant extract serves as a control, demonstrating that none of the BAM proteins are precipitated in the absence of BamC. After BamA depletion, only BamD could be detected in association with BamC.

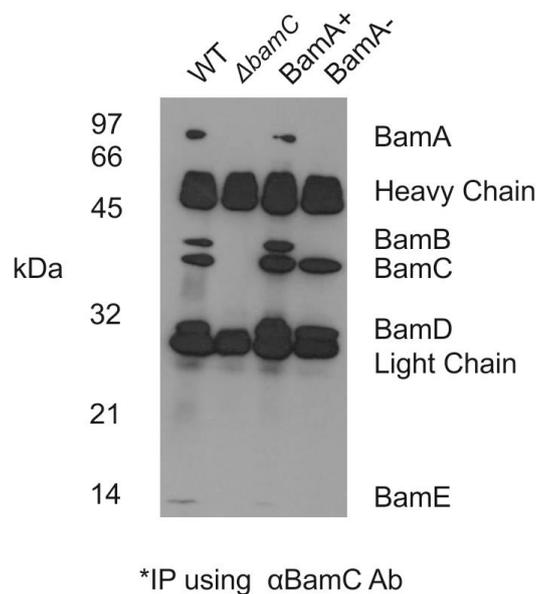
In conclusion, a three hour depletion regime for this strain generates a condition wherein all of the BamA has been removed from the outer membrane, other OMPs including TamA have been depleted, lipoproteins such as the BamB, BamCDE modules are still present in the outer membrane, and the *bamA*-depleted cells remain largely viable and suitable for protein assembly assays.



**Figure 3.7: BamC and BamD remain modular after BamA depletion.** Total membranes isolated from *BamA*<sup>+</sup> or *BamA*<sup>-</sup> cells were solubilised in 1% DDM and analysed by BN-PAGE followed by immunoblotting against each protein of the BAM complex.



**Figure 3.8: Modular nature of the BAM complex.** Schematic of the BAM complex before (A) and after (B) depletion of BamA, showing the BamCDE and BamB modules.



**Figure 3.9: Depletion of BamA does not inhibit the interaction between BamC and BamD.** Co-immunoprecipitation of BamC in BW25113 (WT), *ΔbamC* and MC4100A BamA expressing or depletion cells using antibodies raised against BamC. Samples were analysed by SDS-PAGE followed by Western blotting against each protein of the BAM complex.

### **3.4 DISCUSSION**

#### **3.4.1 Depletion of BamA for outer membrane assembly assays**

The function of the BAM complex is dependent on the essential Omp85 protein BamA and, in *E. coli*, depletion of BamA impacts the assembly of integral membrane proteins of  $\beta$ -barrel topology [11,156-160]. Several studies analysing the role of BamA in assembling secretins into the outer membrane relied on steady state levels of the secretin to measure of the assembly of these secretion pores [44,118,119]. Secretins PulD (*K. oxytoca* T2SS), XcpQ (*P. aeruginosa* T2SS), PscC (*P. aeruginosa* T3SS) and PilQ (*P. aeruginosa* T4P) do not require BamA for assembly into multimers [118,119], whilst, BamA (Omp85) is required for the assembly of the PilQ homolog from *N. meningitidis* [44]. However, secretins assemble into stable pores in the outer membrane and thus rendering it difficult to distinguish between pre-assembled multimers to newly synthesized proteins after depletion of BamA. We therefore wanted to optimize the conditions in which BamA was physiologically depleted but the cells were still viable, able express proteins and target them to the outer membrane for assembly.

The BamA depletion strain used in this study was engineered by Lehr *et al.*, (2010) to investigate the role of BamA in assembling trimeric autotransporter YadA [159]. Cells were grown as described in Chapter 3.2.1, with immunoblots showing that BamA was efficiently depleted and growth curves indicating that the cells were sick with the addition of glucose to the media. Fluorescent microscopy using the LIVE/DEAD BacLight dye showed that after 3 h of depletion 65% of the cells were still viable and this reduced to 50% after 5 h. This provided a window in which target substrates can be expressed over time to monitor their assembly in BamA depleted conditions, in which the majority of cells are still viable. Despite

a proportion of the cells demonstrating damaged membranes TEM revealed no obvious changes in morphology with approximately 10% of cells showing blebbing of the outer membrane. The proteome of the cellular envelope was analysed by sucrose density fractionation and showed that after 3 hours of BamA depletion, approximately 90% of the outer membrane proteome was lost, presumably degraded by periplasmic proteases such as DegP [170]. However, the levels of inner membrane protein (F1 $\beta$ ) and outer membrane lipoproteins (BamBCD) were unchanged after BamA depletion. Taken together, this 3 hour regime was sufficient to physiologically deplete cells of BamA allowing for the subsequent assembly of target substrate proteins into the outer membrane.

### ***3.4.2 Hierarchical assembly of outer membrane proteins***

This study has showed for the first time that the assembly of TamA into the outer membrane is dependent on BamA. This observation proposes a hierarchy of protein assembly factors at the outer membrane, as the depletion of BamA leads to a loss of the TAM and presumably a reduction in the levels of TAM substrates. The function of the TAM depends on a distinct member of the Omp85 protein family, TamA, and an inner membrane protein TamB which are broadly conserved in proteobacteria [45,62]. The TAM was shown to facilitate the insertion of an autotransporter into liposomes *in vitro*, and cells lacking TamA show a reduction in the steady state levels of autotransporters *in vivo* and virulence of pathogens from genera including *Klebsiella*, *Proteus*, *Citrobacter* and *Salmonella* [61,63-65]. In the case of autotransporters, where both the BAM and TAM have been implicated in their assembly [61,63,159,172-174], it is often assumed that there is a fork in the assembly pathway where some autotransporters use the BAM and others use the TAM. Alternatively, it is perhaps more probable that both the BAM and TAM work synergistically, where BAM is

the major contributor to outer membrane protein assembly with the TAM aiding the assembly of proteins with more complicated topologies, such as autotransporters.

### **3.4.3 Modular nature of the BAM complex**

The BAM complex is considered to be a modular machine, and it has been reported that the BAM complex can be purified as BamAB and BamCDE modules. These modules can be reconstituted together *in vitro* into an active form of the BAM “holo-complex” [51,171]. Depletion of BamA lead to the loss of the BamAB module resulting in increased levels of BamB detected by BN-PAGE. Consistent with observations from *in vitro* studies with liposomes not containing BamA, the BamCDE remained modular after depletion of BamA [51,171]. In both cases all of the BAM lipoproteins tested were targeted to the outer membrane, suggesting that Lol pathway is still functional after depletion of BamA.

The *in vivo* roles of the BAM lipoproteins are unknown. Genetic studies have shown that *bamD* is essential [52,53] and that the deletion of any two non-essential lipoproteins (BamBCE) resulted in a synthetic or conditional lethal phenotype [11,54,55]. *E. coli* cells with single gene deletions of *bamB*, *bamC* or *bamE* are viable and show a variety  $\beta$ -barrel assembly defects in *bamB* and *bamE* mutants [54,56,57]. Recently, it has been shown that the BAM lipoproteins (BamD and BamB) play a role in facilitating the folding of the BamA barrel [175,176]. Moreover, in a  $\Delta$ *bamB*  $\Delta$ *bamE* conditional mutant, BamA levels were significantly reduced and the folding of the BamA barrel is compromised [55]. Given the observation of these modules after depletion of BamA it could be speculated that the folding and assembly of BamA mimics that observed from *in vitro* studies.

It is now clear that BamC adopts a novel lipoprotein topology in which a C-terminal structured domain is exposed to the cell surface and the unstructured N-terminus domain located in the periplasm [165]. Recent TIRF super-resolution microscopy further supports this topology for BamC [177], but it is still unclear how BamC adopts this transmembrane topology. Recently it was shown that lipoprotein RcsF adopts a similar but distinct topology in which the N-terminus is attached to the cell surface and the folded C-terminal domain is located in the periplasm in which the transmembrane domain is threaded through the lumen of  $\beta$ -barrels assembled by the BAM complex [178]. It would be interesting to determine whether BamC uses a similar mechanism to reach the cell surface and if it can thread through the BamA barrel to adopt this topology. The question then arises as to whether the topology of newly synthesized BamC changes after depletion of BamA, and the significance of any possible interactions between BamA and BamC on the cell surface.

## **CHAPTER 4: Multimeric outer membrane proteins GspD, Wza and CsgG do not require BamA or TamA for assembly.**

### ***4.1 INTRODUCTION***

The correct folding and insertion of proteins destined to the outer membrane is an essential process for the integrity and functionality of the bacterial outer membrane. The BAM complex is a multi-protein machine responsible for the correct folding and insertion of  $\beta$ -barrel proteins into the bacterial outer membrane [44,11]. A second machine, the TAM, was shown to facilitate the insertion of  $\beta$ -barrels of more complicated topologies such as some autotransporters into liposomes *in vitro* [63], and resulted in the reduction in the steady state levels of some autotransporters *in vivo* [61].

Collin *et al.*, (2007) investigated the dependence on BamA for the multimerization of PulD [118]. In an *E. coli* strain, where BamA expression was placed under the control of a repressible promoter, PulD was associated with the outer membrane in a multimeric form even after the strain had been grown under ‘BamA-repressive’ conditions. As a control, the known  $\beta$ -barrel LamB failed to oligomerize into its mature trimer. A separate study in *P. aeruginosa* showed that secretins XcpQ (T2SS), PscC (T3SS) and PilQ (T4P) can assemble into multimers in cells depleted of BamA [119]. In contrast, a study by Voulhoux *et al.*, (2003) using showed that secretin PilQ from *N. meningitidis* was unable to form oligomers efficiently in BamA-depleted cells causing the accumulation of monomeric PilQ [44].

The use of BamA depletion strains have been used to assay assembly of secretins, however the role of the BAM in assembling other multimeric secretion pores like Wza and CsgG remains unknown. This chapter will focus on using an optimized regime of *bamA* depletion described in Chapter 3, which also results in the loss of TamA, with assembly assays to highlight the role of the BAM complex or other accessory assembly factors in the assembly and insertion of multimeric pores GspD, Wza and CsgG.

## 4.2 METHODS and MATERIALS

### 4.2.1 Strains, Plasmids and Growth Conditions

The bacterial strains and plasmids used in this chapter are listed in Table 4.1, using the parental strains *E. coli* MC4100A, *E. coli* BL21 (DE3) (Invitrogen), *E. coli* BW25113 and *E. coli* DH5 $\alpha$  (Invitrogen). Strains were grown in LB supplemented with the appropriate antibiotics (ampicillin 100  $\mu$ g/ml, kanamycin 30  $\mu$ g/ml or chloramphenicol 25  $\mu$ g/ml).

**Table 4.1: Strains and Plasmids used in Chapter 4**

Strain	Description	Reference
BL21 (DE3)	<i>E. coli</i> strain: F <sup>-</sup> <i>ompT hsdSB</i> (rB- mB-) <i>dcm gal</i> (DE3)	Invitrogen
MC4100A	<i>E. coli</i> strain: F <sup>-</sup> <i>araD139</i> $\Delta$ ( <i>argF-lac</i> )U169 <i>rpsL150</i> (Str <sup>R</sup> ) <i>relA1 flbB5301 deoC1 ptsF25 rbsR</i>	[179]
MC4100A BamA depletion	<i>att-B::</i> ( <i>araC</i> <sup>+</sup> , P <sub>BAD</sub> , <i>bamA</i> <sup>+</sup> )	[159]
BW25113	$\Delta$ ( <i>araD-araB</i> )567 $\Delta$ <i>lacZ4787</i> (::rrnB-3) $\lambda$ <i>rph-1</i> $\Delta$ ( <i>rhaD-rhaB</i> )568 <i>hsdR514</i>	[180]
BW25113 $\Delta$ <i>wza</i>	$\Delta$ ( <i>araD-araB</i> )567 $\Delta$ <i>lacZ4787</i> (::rrnB-3) $\lambda$ <i>rph-1</i> $\Delta$ ( <i>rhaD-rhaB</i> )568 <i>hsdR514</i> $\Delta$ <i>wza-760::kan</i>	[180]
BW25113 $\Delta$ <i>csgG</i>	$\Delta$ ( <i>araD-araB</i> )567 $\Delta$ <i>lacZ4787</i> (::rrnB-3) $\lambda$ <i>rph-1</i> $\Delta$ ( <i>rhaD-rhaB</i> )568 <i>hsdR514</i> $\Delta$ <i>csgG778::kan</i>	[180]
DH5 $\alpha$	F <sup>-</sup> <i>endA1 glnV44 thi-1 recA1 relA1 gyrA96 deoR nupG</i> $\Phi$ 80d <i>lacZ</i> $\Delta$ M15 $\Delta$ ( <i>lacZYA-argF</i> )U169, <i>hsdR17</i> (r <sub>K</sub> <sup>-</sup> m <sub>K</sub> <sup>+</sup> ), $\lambda$ -	Invitrogen
Plasmid	Description	Reference
pBAD24	<i>ori</i> pMB1, Amp <sup>r</sup>	[140]
pET DUET-1	<i>ori ColE1</i> , <i>lacI</i> gene, Amp <sup>r</sup>	Novagen
pACYC DUET-1	<i>ori P15A</i> , <i>lacI</i> gene, Cm <sup>r</sup>	Novagen

pJP168	P <sub>BAD</sub> promoter and <i>araC</i> of pBAD24 replaced by P <sub>tetA</sub> and <i>tetR</i>	[15]
pJP181	E2348/69 <i>gspD-C<sub>4</sub></i> was cloned into the <i>NcoI/XbaI</i> sites of pJP168.	This Study
pBAD <sub>tet</sub> Wza	<i>E. coli</i> st. K12 <i>wza-C<sub>4</sub></i> was cloned into the <i>NcoI/HindIII</i> sites of pJP168.	This Study
pBAD <sub>tet</sub> Wza C21A	<i>E. coli</i> st. K12 <i>wza-C<sub>4</sub></i> C21A (cysteine at position 21 was replaced with alanine) was cloned into the <i>NcoI/HindIII</i> sites of pJP168.	This Study
pBAD <sub>tet</sub> Wza*	<i>E. coli</i> st. K12 <i>wza-C<sub>4</sub></i> * (native Wza signal sequence replaced by a PelB signal sequence) was cloned into the <i>NcoI/HindIII</i> sites of pJP168.	This Study
pBAD <sub>tet</sub> CsgG	<i>E. coli</i> st. K12 <i>csgG-C<sub>4</sub></i> was cloned into the <i>NcoI/HindIII</i> sites of pJP168.	Dr. Iain Hay (Lithgow Laboratory)
pET DUET-1 CsgG-C <sub>4</sub>	<i>E. coli</i> st. K12 <i>csgG-C<sub>4</sub></i> was cloned into <i>NcoI/HindIII</i> of pET DUET-1	Dr. Iain Hay (Lithgow Laboratory)
pET DUET-1 CsgG-C <sub>4</sub> CsgE	<i>E. coli</i> st. K12 <i>csgG-C<sub>4</sub></i> and <i>csgE</i> were cloned into <i>NcoI/HindIII</i> and <i>NdeI/XhoI</i> sites of pET DUET-1 respectively	Dr. Iain Hay (Lithgow Laboratory)
pET DUET-1 CsgG-C <sub>4</sub> CsgF	<i>E. coli</i> st. K12 <i>csgG-C<sub>4</sub></i> and <i>csgF</i> were cloned into <i>NcoI/HindIII</i> and <i>NdeI/XhoI</i> sites of pET DUET-1 respectively	Dr. Iain Hay (Lithgow Laboratory)
pACYC DUET-1 CsgE CsgF	<i>E. coli</i> st. K12 <i>csgE</i> and <i>csgF</i> were cloned into <i>NcoI/HindIII</i> and <i>NdeI/XhoI</i> sites of pET DUET-1 respectively	Dr. Iain Hay (Lithgow Laboratory)

#### 4.2.2 Cloning of *wza* and *csgG*

The open-reading frames for, Wza, CsgG, CsgE and CsgF correspond to sequences in the Genbank entries, *wza* (946558), *csgG* (945619), *csgE* (945711) and *csgF* (945622). The GspD-C4 expressing plasmid was constructed previously (Chapter 2.2.2). The primers used to make the Wza-C<sub>4</sub> constructs used in this chapter are listed in table 4.2. Quick change mutagenesis was performed to generate Wza-C<sub>4</sub> C21A mutant the using the primer pair WzaC21A F/ WzaC21A R according to Kunkel, 1985 [181]. The CsgG-C<sub>4</sub>, CsgE and CsgF expressing constructs were made by Dr. Iain Hay (Lithgow Laboratory, Monash University)

**Table 4.2: Primers used to construct the Wza expressing plasmids.**

<b>Primer Name</b>	<b>Primer sequence</b>
Wza F	CGGTACCATGGTGAAATCCAAAATGAAATTG
Wza R C <sub>4</sub> tag	GCCGGAAGCTTTTAACGGCCGCCCGGTTCCATGCAGCA GCCCAGGCAGCAGTTCAGAAAGCTGCCCGCCAGTTAT GAATGTCGCTGG
WzaC21A F	TGATAAGCGGTGCGACAGTACTTCC
WzaC21A R	GGAAGTACTGTTCGCACCGCTTATCA
PelB ss Wza C <sub>4</sub> F	GCGCCGCCATGGAATACCTGCTGCCGACCGCTGCTGCT GGTCTGCTGCTCCTCGCTGCCAGCCGGCGATGGCCAC AGTACTTCCGGGCAGCAATATG
C <sub>4</sub> tag R	CGGCGCAAGCTTTTAACGGCCGCCCGGTTTC

#### **4.2.3 SDS-PAGE**

See Chapter 2.2.5.

#### **4.2.4 Immunoblotting**

See Chapter 2.2.6.

#### **4.2.5 GspD assembly assays in MC4100A BamA depletion strain**

MC4100A cells transformed with *gspD-C<sub>4</sub>* expressing plasmids were grown as described in Chapter 3.2.1 prior to induction with anhydrotetracycline (AhT, 10 ng/ml). After 0, 15, 30, 60

and 120 minutes, 1 ml samples were taken and prepared as per Chapter 2.2.7. Samples were analysed by 3-14% SDS-PAGE and fluorometry (described in Chapter 2.2.7).

#### ***4.2.6 Wza assembly assays in MC4100A BamA depletion strain***

MC4100A cells transformed with *wza-C<sub>4</sub>* expressing plasmids were grown as described in Chapter 3.2.1 prior to induction with 10 ng/ml AhT for 40 minutes at 37 °C. After 0, 10, 20, 30 and 40 minutes, 1 ml samples were taken and cells were harvested by centrifugation. Cell pellets were resuspended in a non-standard SDS-PAGE lysis buffer and samples were prepared according to (Chapter 2.2.7) with one exception. After the addition of the Lumio green dye lysates were incubated on ice to prevent the dissociation of Wza multimers. Samples were analysed by 3-14% SDS-PAGE and fluorometry (described in Chapter 2.2.6).

#### ***4.2.7 CsgG assembly assays in MC4100A BamA depletion strain***

MC4100A cells transformed with *csgG-C<sub>4</sub>* expressing plasmids were grown as described in Chapter 3.2.1 prior to induction with 10 ng/ml AhT for 40 minutes at 37 °C. After 0, 10, 20, 30 and 40 minutes, 1 ml samples were taken and cells were harvested by centrifugation. Cell pellets were resuspended in a non-standard SDS-PAGE lysis buffer and samples were prepared according to (Chapter 2.2.7) with one exception. After the addition of the Lumio green dye lysates were incubated on ice to prevent the dissociation of CsgG multimers. Samples were analysed by 3-14% SDS-PAGE and fluorometry (described in Chapter 2.2.6).

#### ***4.2.8 Isolation of total membranes***

Cells were grown in LB media supplemented with 0.02% arabinose for BamA expression or 0.02 % glucose for BamA depletion for 3 h at 37 °C. AhT (10 ng/ ml) was then added to the cultures for the expression of GspD, Wza and CsgG and incubated for 1 h at 37 °C. Cells were then harvested and total membranes were prepared as described in Chapter 2.2.8.

#### ***4.2.9 Sucrose density fractionation***

The separation of bacterial inner and outer membranes was performed as described in Chapter 2.2.9. Thirty microlitres of each fraction were prepared using Lumio detection kit (described in Chapter 2.2.7) and loaded onto a 3-14% SDS-PAGE and analysed by fluorometry (described in Chapter 2.2.7). Western immunoblotting was used to determine the localisation of outer membrane protein BamB (1:30,000) and Coomassie staining to monitor the proteins profile of membrane fractions.

#### ***4.2.10 Urea Extraction of peripherally associated membrane proteins***

Total membranes (50 µg) were treated with 1 ml of 5M urea or 1 x PBS and allowed to mix on a rotary wheel for 1 h at room temperature. Insoluble and extracted fractions were separated by ultracentrifugation (100,000 x g, 50 minutes, 4 °C). The soluble fraction containing extracted proteins were precipitated with TCA (10% final concentration) on ice for 30 min. Precipitated proteins were collected by centrifugation (15,000 rpm, 20 min, 4 °C) and protein pellets were washed twice with cold 100% acetone. Pellets from the soluble fraction were dried at 37 °C and both soluble and insoluble fractions were resuspended in 15 µl of SDS sample buffer + 5 µl of Lumio sample buffer as described in Chapter 2.2.7. An equivalent amount of total membranes (50 µg) was prepared by adding 5 µl of Lumio sample

buffer and SDS sample buffer to make up 20  $\mu$ l final volume. Samples separated by 3-14% SDS-PAGE and analysed by fluorometry and immunoblotting for F<sub>1</sub> $\beta$  (1:8000).

#### ***4.2.11 Wza complementation assay***

The AhT inducible Wza-C4 plasmid was transformed into the E coli BW25113 *wza* isogenic deletion mutant from the KEIO collection (JW2047-1) [180]. Colanic acid levels of strains were assayed by measuring the content of methylpentose (fucose) in the in the culture supernatant as described by Obadia et al., (2007) [182]. Cells were grown in LB with 10 ng/ml AhT for 24 h at 22 °C and harvested by centrifugation at 12,000 x g. Supernatants were collected and the cell pellets washed with 1 volume of saline (150 mM NaCl), cells were harvested as above and the supernatant pooled with the previous supernatant. The secreted polysaccharides were precipitated with 3 volumes of ethanol and incubated at 4 °C overnight. Insoluble material was collected by centrifugation at 12,000 x g and dried at 65 °C. The pellet was resuspended in 10 ml H<sub>2</sub>O. Quantification of colanic acid content of the obtained material was carried out according to Dische & Shettles (1951) [152]. 50  $\mu$ l of the preparation were diluted to 1 ml with distilled water, and mixed with 4.5 ml of H<sub>2</sub>SO<sub>4</sub>:H<sub>2</sub>O (6:1; v/v) and incubated at 100 C for 20 min. The difference in absorbance at 396 nm and 427 nm was measured both before and after addition of 100  $\mu$ l of cysteine hydrochloride. The difference in Abs<sub>396</sub>-Abs<sub>427</sub> after the addition of cysteine hydrochloride can be attributed to fucose. 10-100 ug/ml L-fucose was used as a standard. Fucose contents were normalised to the cellular dry weight of an equivalent volume of cells.

#### ***4.2.12 CsgG complementation assay***

The AhT inducible CsgG-C4 plasmid was transformed into the *E. coli* BW25113 *csgG* isogenic deletion mutant from the KEIO collection (JW1020-1) [180]. Curli fibre secretion was monitored using a Congo red staining assay. Cells were grown in LB overnight at 37 °C and 10 µl was subsequently spotted onto modified LB agar (5 g/L tryptone, 2.5 g/L yeast extract, 2% agar, 50 µg/ml Congo red, 5 µg/ml Coomassie brilliant blue G250 and 10 ng/ml AhT). Plates were incubated at 26 °C for 48 h to induce curli expression.

#### ***4.2.13 AraB structural modelling***

Phyre 2.0 [183] was used to create a structural model of AraB using the *E. coli* K12 AraB sequence (NCBI-Protein ID NP\_414605).

#### ***4.2.14 In-gel digestion of proteins and LC-MS/MS***

See Chapter 2.2.11.

## 4.3 RESULTS

### 4.3.1 *GspD does not require BamA for assembly into bacterial outer membranes*

To determine if BamA assists in the assembly of a T2SS secretin GspD, the assay developed in Chapter 2.2.7 was adapted to monitor the assembly of membrane proteins after *bamA* depletion. GspD-C<sub>4</sub> was expressed under the control of the AhT inducible promoter for 2h post BamA depletion. GspD-C<sub>4</sub> was observed using the Lumio Green reagent within 15 minutes of induction with 10 ng/ml AhT (Figure 4.1). The monomeric form of GspD-C<sub>4</sub> is detected in whole cells at early time points with HMW multimeric GspD-C<sub>4</sub> forming gradually over time. Unexpectedly the multimeric GspD-C<sub>4</sub> appeared earlier and had higher levels in *bamA*-depleted cells.

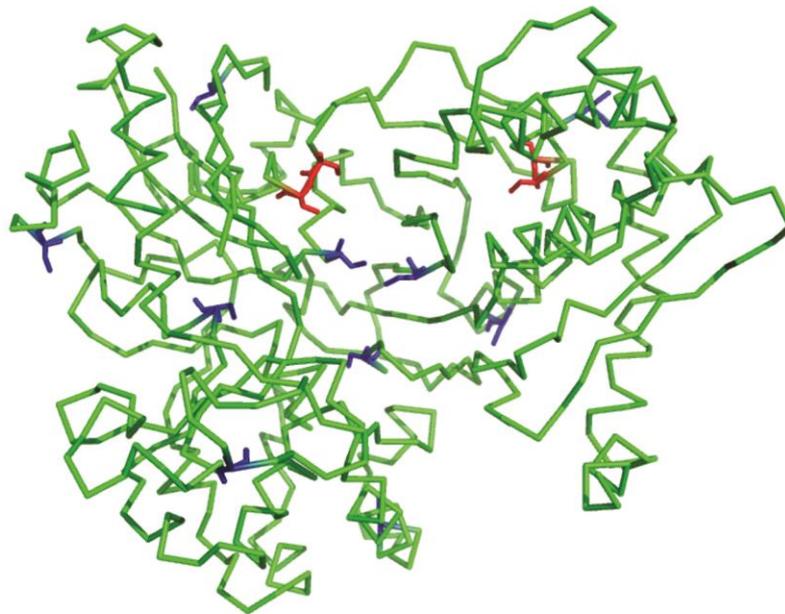
A previously undescribed protein of ~60kDa was also observed in BamA expressing (arabinose treated) cells using the Lumio dye. Mass Spectrometry analysis revealed this protein to be L-ribulokinase (AraB). The gene encoding AraB belongs to the AraBAD operon which is up regulated in the presence of arabinose. Why would AraB react with Lumio Dye? Investigation of the primary sequence of AraB revealed two cysteine pairs which could bind the Lumio Green dye (Figure 4.2) and SlyD, an endogenous metallo-chaperone that reacts with the Lumio reagent, was also observed and served as an internal loading control.

To determine if the GspD-C<sub>4</sub> multimers were present in the outer membrane of cells lacking BamA, total membranes were prepared from cells minimally expressing GspD-C<sub>4</sub> and its cognate pilotin AspS from a *bamA* depletion and expression background, and subjected to



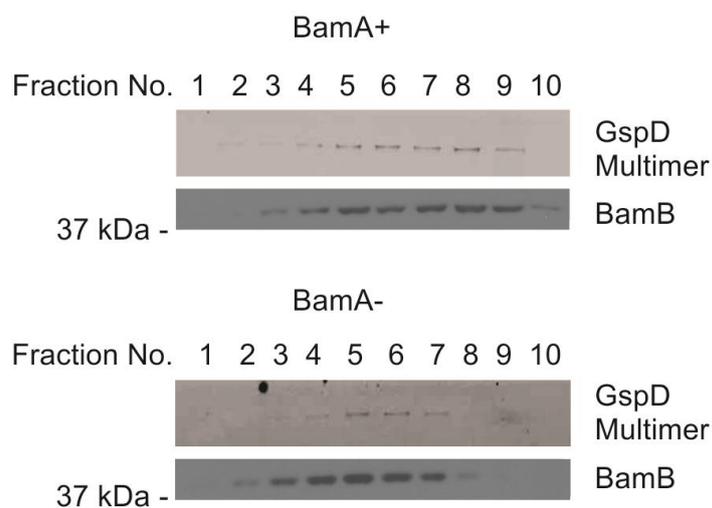
**A****AraB**

MAIAIGLDFGSDSVRALAVDCATGEEIATSVEWYPRWQKGQFC DAPNNQFRHHPRDYIES  
 MEAALKTVLAELSVEQRAAVVGIGVDSTGSTPAPIDADGNVLA LRPEFAENPNAMFVLWK  
 DHTAVEEAEETRIC HAPGNVDYSRYIGGIYSSEWFWAKILHVTRQDSAVAQSAASWIEL  
 CDWVPALLSGTTRPQDIRRGFC SAGHKS LWHESWGGLPPASFFDELDPILNRHLPSPLFT  
 DTWTADIPVGTLP EWAQRLGLPESVVISGGAFD HMGAVGAGAQPNALVKVIGTSTDI  
 LIADKQSVGERAVKGI CGQVDG SVVPGFIGLEAGQSAFGDIYAWFGRV LGWPLEQLAAQH  
 PELKTQINASQKQLLPALTEAWAKNPSLDHLPVVLDFWNGRRTPNANQRLKGVI TDLNLA  
 TDAPLLFGGLIAATAFGARAIMECF TDQGI AVNNVMALGGIARKNOVIMQACCDV LNRPL  
 QIVASDQCCALGAAIFAAVA AKVHADIPSAQQKMASAVEKTLQFC SEQAQRFEQLYRRYQ  
 QWAMSAEQHYLPTSAPAQAAQAVATL

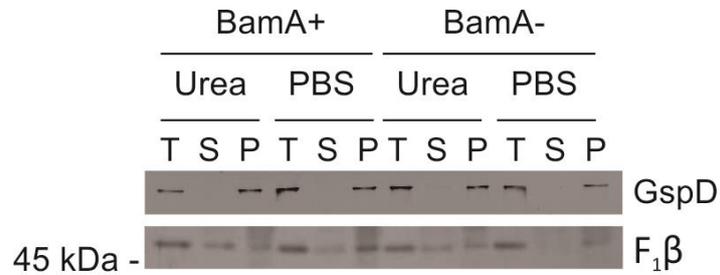
**B**

**Figure 4.2: Structural model of AraB showing Cysteines that may interact with the Lumio Green dye.** (A) Protein sequence of AraB from *E. coli* K12 (NCBI-Protein ID NP\_414605) showing cysteine pairs (red boxes) and single cysteines (blue boxes). (B) Phyre 2.0 [183] was used to create a structural model of AraB from *E. coli* K12. Cysteines pairs (red) or individual cysteines (blue) that may bind the Lumio dye are highlighted on the stick model of the AraB structure.

sucrose density fractionation. Under both growth conditions multimeric GspD-C<sub>4</sub> was found in the outer membrane fractions suggesting that BamA does not aid in the AspS dependent targeting of GspD-C<sub>4</sub> to the outer membrane (Figure 4.3). Representative Coomassie stained SDS-PAGE gels of these fractions are shown in Appendix A3. Correctly assembled integral membrane proteins are highly resistant to extraction by Urea [44,118]. Multimeric GspD-C<sub>4</sub> remained integrally associated with the outer membrane upon treatment with 5M urea in both BamA<sup>+</sup> and BamA<sup>-</sup> containing membranes, whereas inner membrane protein F<sub>1</sub>β was extracted with urea (Figure 4.4).



**Figure 4.3: GspD is targeted to the outer membrane after BamA depletion.** Total membranes were prepared from cells expressing GspD-C<sub>4</sub> and its cognate pilotin AspS after 3 hours of BamA depletion and separated by sucrose density fractionation. Fractions were prepared by adding Lumio sample buffer and analysed by 3-14 % SDS-PAGE and fluorimetry.

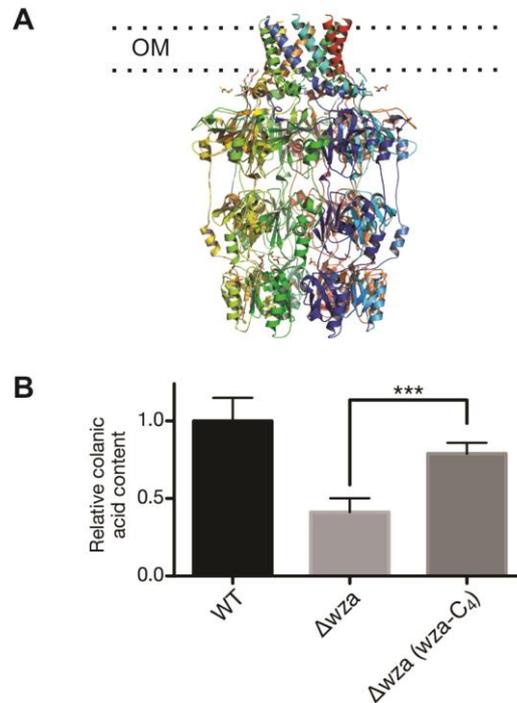


**Figure 4.4: GspD is integrated to the outer membrane after BamA depletion.** Total membranes from BamA + or BamA – cells expressing GspD-C<sub>4</sub> were treated with 5 M urea or PBS. Extracted supernatant material (S) was separated from the integral membrane fraction pellet (P) by ultracentrifugation. A sample of starting material was used as a total membrane reference (T).

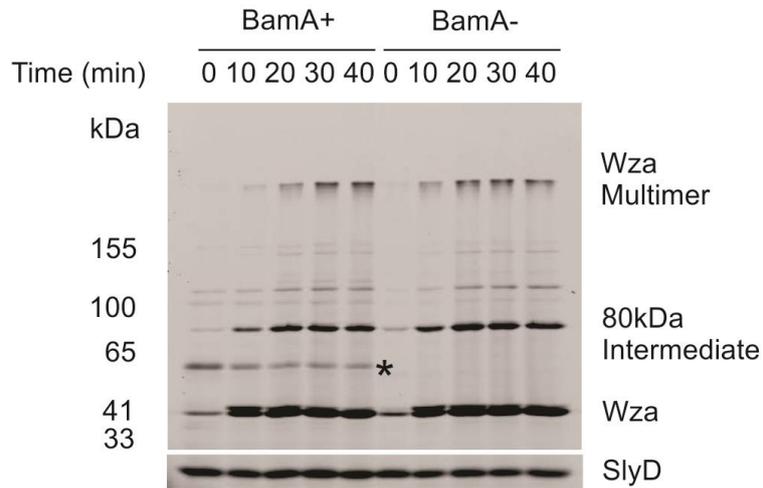
### ***4.3.2 Wza does not require BamA for assembly into bacterial outer membranes***

The Wza secretion pore is an octameric  $\alpha$ -helical barrel with the bulk of the protein forming a large periplasmic vestibule [19] (Figure 4.5a). In order to visualise the assembly of Wza, a FLAsH tag was engineered at the C-terminus of Wza (Wza-C<sub>4</sub>), and expressed via the AhT inducible system described above. The presence of extracellular polysaccharide (EPS) in the culture media was used as a read-out for the functionality of Wza-C<sub>4</sub>. Tagged Wza was shown to be functional, by complementation of the EPS (colanic acid) secretion defect of  $\Delta wza$  cells (Figure 4.5b). The Wza-C<sub>4</sub> monomer migrates at ~40kDa and was detected at early time points with a precursor form also apparent (Figure 4.6). The HMW Wza-C<sub>4</sub> multimer appears at later time points with a ~80 kDa possible assembly intermediate form of Wza-C<sub>4</sub> observed during the time course. Depletion of BamA did not alter the ability of Wza-C<sub>4</sub> to assemble into multimers. To determine if the targeting of Wza-C<sub>4</sub> is not affected in BamA-depleted cells, total membranes were prepared as described above and subjected to sucrose density fractionation. Multimeric Wza-C<sub>4</sub> was present in the outer membrane fractions (Figure 4.7, Appendix A3) and was resistant to urea washing in both BamA<sup>+</sup> and BamA<sup>-</sup> conditions (Figure 4.8).

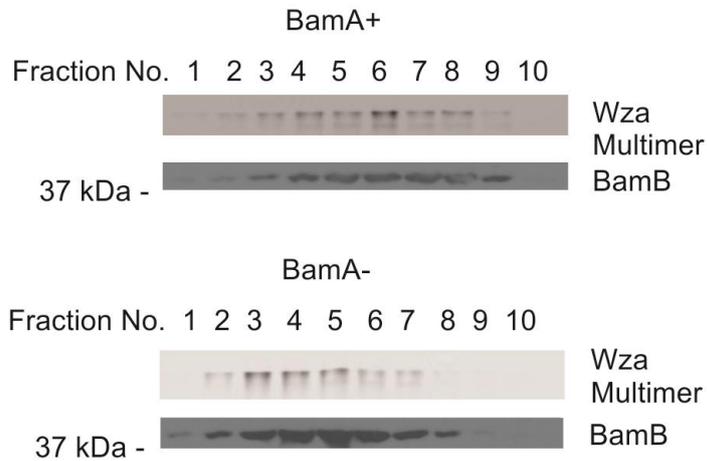
Wza is a lipoprotein, and is associated with the outer membrane via both its N and C termini. The C-terminus forms a  $\alpha$ -helical transmembrane domain, whilst the acylated N-terminus is anchored to the inner leaflet of the outer membrane. It is plausible that the N-terminal anchor of Wza stabilises the Wza multimer if the C-terminus is not correctly assembled into the outer membrane. A non-acylated form of Wza, where the protein is no longer anchored to the outer membrane by its N-terminus, was engineered to determine if a non-lipidated form of Wza could assemble independent of BamA. A substitution mutant of the N-terminal cysteine to alanine (C21A) was constructed and found its expression to be toxic in cells depleted of



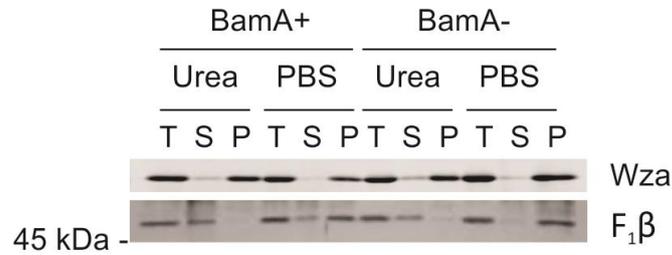
**Figure 4.5: Wza-C<sub>4</sub> is functional and able to secrete EPS.** (A) Crystal structure of the capsular polysaccharide secretion pore Wza (PDB: 2J58, [19]). (B) The tetracysteine-tagged Wza (Wza-C<sub>4</sub>) is capable of complementing a *wza* deletion mutant. The level of fucose (a component of colanic acid polysaccharide) in the culture supernatant was measured and normalised to the cellular dry weight. The relative levels as compared to wild-type (WT) are presented. \*\*\* p=0.0007.



**Figure 4.6: Wza does not require BamA for assembly into multimers.** *E. coli* MC4100A BamA depletion strains harbouring the Wza-C<sub>4</sub> expression plasmid were cultured in LB containing arabinose or glucose for 3 h before the addition of 10 ng/ml AhT to the culture. At the indicated time points cells were harvested, resuspended in Lumio sample buffer and separated by 3-14 % SDS-PAGE. The polyacrylamide gels were then imaged by fluorimetry. Wza and control protein SlyD are shown, and the protein identified as AraB is indicated with an asterisk.



**Figure 4.7: Wza is targeted to the OM after BamA depletion.** Total membranes were prepared from cells expressing Wza-C<sub>4</sub> after 3 hours of BamA depletion and separated by sucrose density fractionation. Fractions were prepared by adding Lumio sample buffer and analysed by 3-14 % SDS-PAGE.



**Figure 4.8: Wza is integrated to the OM after BamA depletion.** Total membranes from BamA + or BamA – cells expressing Wza-C<sub>4</sub> were treated with 5 M urea or PBS. Extracted material (S) was separated from the integral membrane fraction (P) by ultracentrifugation. An additional sample was used as a total membrane reference (T).

BamA (data not shown). Therefore, a Wza chimera (Wza\*-C<sub>4</sub>) was constructed in which its native type II signal sequence was substituted with the type I signal sequence of PelB (Figure 4.9a). Low levels of Wza\*-C<sub>4</sub> were observed from total membranes isolated from both BamA expression or depletion strains when compared to cells expressing wild-type Wza-C<sub>4</sub>. Multimeric Wza\*-C<sub>4</sub> was observed in the outer membrane of cells depleted of BamA (Figure 4.9b) suggesting that while the N-terminus is important for multimer stability, the transmembrane domain of Wza alone does not require BamA for assembly into the outer membrane.

**A**



**B**



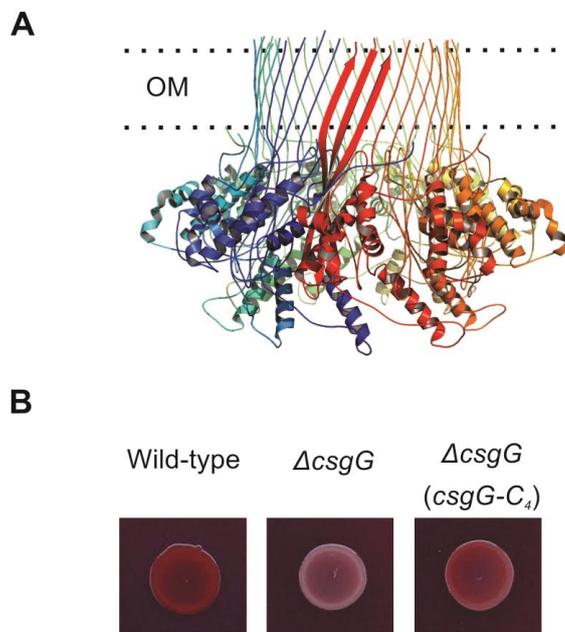
**Figure 4.9: Assembly of a non-acylated chimera of Wza does not require BamA.** (A) Schematic of FAsH tagged Wza and Wza\* (PeIBss-Wza 22-379). (B) Total membranes were prepared from cells expressing Wza-C<sub>4</sub> or Wza\*-C<sub>4</sub> after 3 hours of BamA depletion and separated by sucrose density fractionation. Total membranes and fractions corresponding to the outer membrane were prepared by adding Lumio sample buffer and analysed by 3-14 % SDS-PAGE.

### ***4.3.3 CsgG does not require BamA for assembly into bacterial outer membranes***

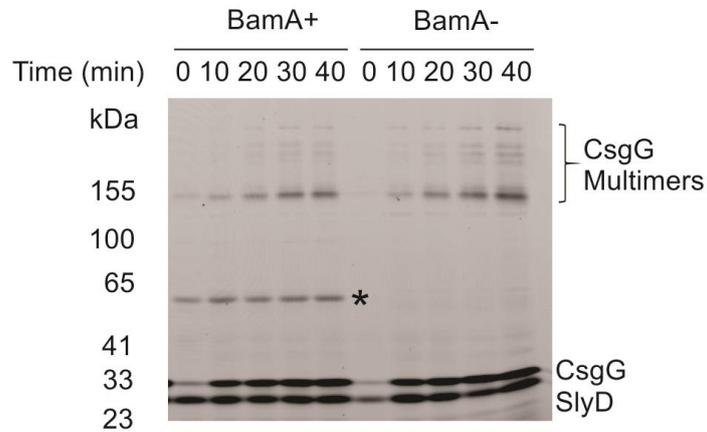
The curli secretion pore CsgG is comprised of 9 monomers that form a 36 stranded  $\beta$ -barrel with a large periplasmic vestibule [20] (Figure 4.10a). To determine if the CsgG barrel requires the BAM for its assembly a F<sub>l</sub>AsH tag was engineered at the C-terminus of CsgG (CsgG-C<sub>4</sub>), with protein expression placed under the control of the AhT inducible promoter described above. To demonstrate that the F<sub>l</sub>AsH tagged CsgG was still functional the ability to secrete curli fibres was assessed with a Congo red binding assay. Colonies from WT cells stained red whereas the  $\Delta$ csgG mutant produced white colonies when grown on modified LB agar containing Congo red. Complementation was observed from the  $\Delta$ csgG mutant expressing CsgG-C<sub>4</sub> with colonies showing a pink colour (Figure 4.10b). Assembly assays showed that monomeric CsgG-C<sub>4</sub> migrated at ~30 kDa and was observed rapidly after induction with AhT (Figure 4.11). In both Bam<sup>+/-</sup> conditions a predominant 150 kDa band appears during later time points with several other HMW bands observed over time. Multimeric CsgG-C<sub>4</sub> localised at the outer membrane (Figure 4.12, Appendix A3) showing integral topology in both BamA<sup>+</sup> and BamA<sup>-</sup> conditions (Figure 4.13) suggesting BamA is not required CsgG assembly into the outer membrane.

### ***4.3.4 CsgE and CsgF increase the assembly efficiency of CsgG***

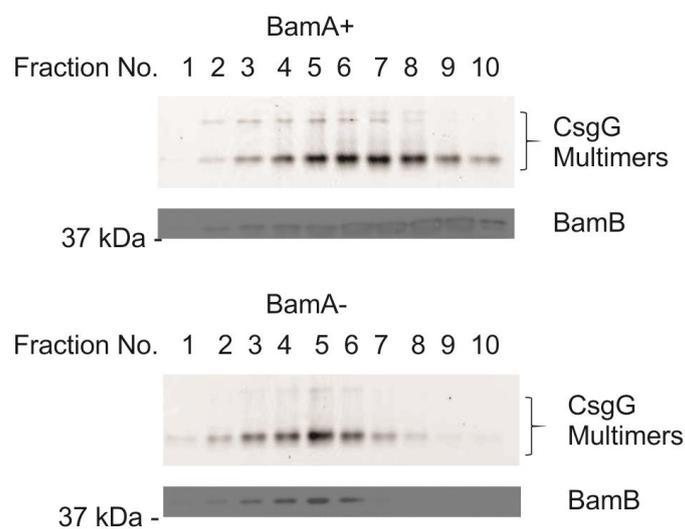
Curli biogenesis in *E. coli* is dependent on 2 distinct operons (Figure 4.14a) [184]. The 150 kDa intermediate of CsgG has been previously observed in cells overexpressing CsgG, lacking individual proteins involved in curli biosynthesis (CsgABEF) and upon heat treatment over 55°C of whole cell extracts [185]. CsgE and CsgF are two accessory proteins required for the targeting of curli subunits CsgA and CsgB respectively to the curli secretion apparatus and have been shown to interact with CsgG [186,153]. To determine if co-expression of CsgE or CsgF drives the formation of the 150 kDa CsgG intermediate to the



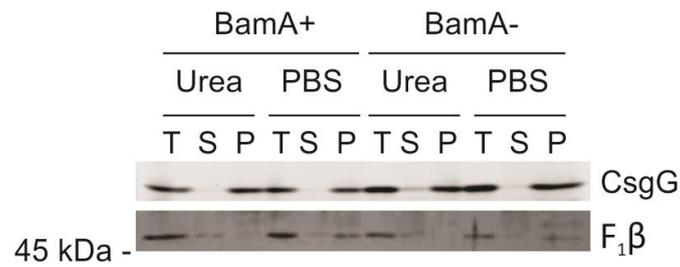
**Figure 4.10: CsgG-C4 is functional allowing the secretion of curli fibers.** (A) Crystal structure of the nanomeric curli biogenesis channel CsgG (PDB: 4UV3, [20]). (B) The tetracysteine-tagged CsgG (CsgG-C<sub>4</sub>) is capable of complementing a *csgG* deletion mutant. Congo red staining of WT and  $\Delta csgG$  strains harbouring the indicated plasmids after 48 h of growth at 26 °C on modified LB agar plates.



**Figure 4.11: CsgG can assemble into multimers after BamA depletion.** *E. coli* MC4100A BamA depletion strains harbouring the CsgG-C<sub>4</sub> expression plasmid were cultured in LB containing arabinose or glucose for 3 h before the addition of 10ng/ml AhT to the culture. At the indicated time points cells were harvested, resuspended in Lumio sample buffer and separated by 3-14 % SDS-PAGE. The polyacrylamide gels were then imaged by fluorimetry. CsgG and control protein SlyD are shown, and the protein identified as AraB is indicated with an asterisk.

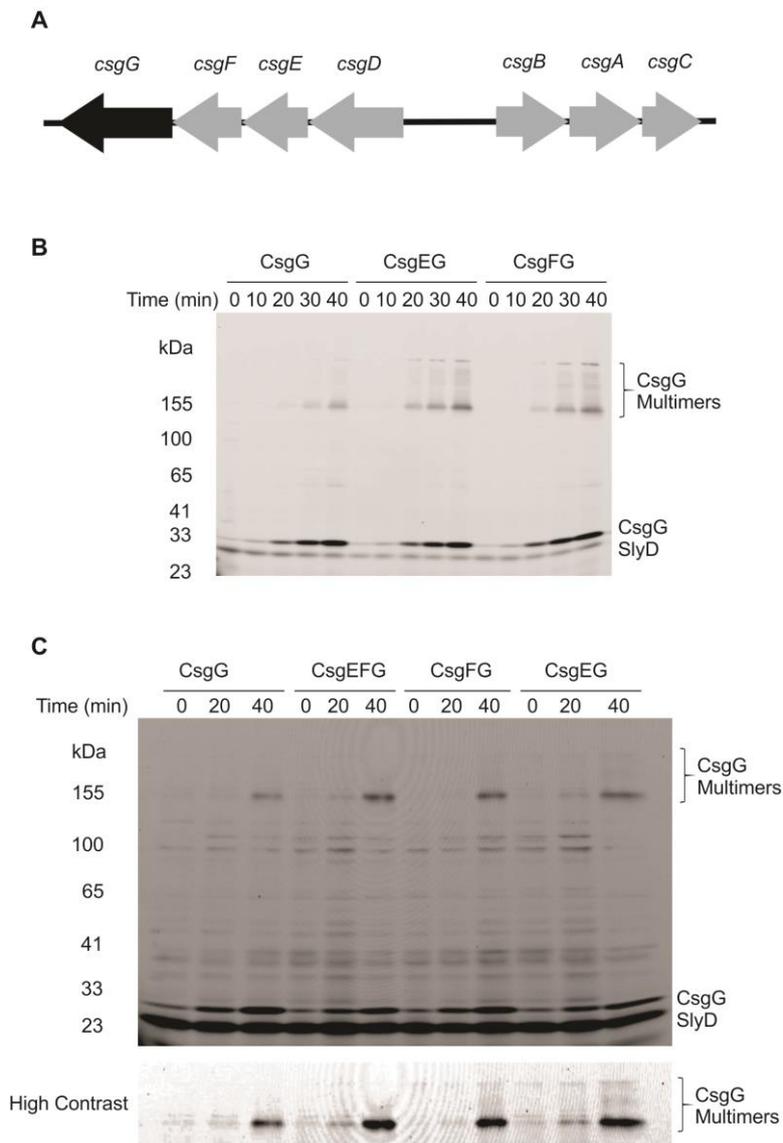


**Figure 4.12: CsgG is targeted to the OM after BamA depletion.** Total membranes were prepared from cells expressing CsgG-C<sub>4</sub> after 3 hours of BamA depletion and separated by sucrose density fractionation. Fractions were prepared by adding Lumio sample buffer and analysed by 3-14 % SDS-PAGE.



**Figure 4.13: CsgG is integrated to the OM after BamA depletion.** Total membranes from BamA + or BamA – cells expressing CsgG-C<sub>4</sub> were treated with 5 M urea or PBS. Extracted material (S) was separated from the integral membrane fraction (P) by ultracentrifugation. An additional sample was used as an input reference (T).

mature CsgG multimer, CsgG-C<sub>4</sub> and either CsgE or CsgF were cloned into pET DUET-1 for expression in the *E. coli* strain BL21 (DE3). Expression of either CsgE or CsgF improved CsgG-C<sub>4</sub> assembly with increased amounts of both the 150 kDa and HMW forms of CsgG-C<sub>4</sub> observed over time (Figure 4.14b). As CsgE and CsgF each increased the amount of CsgG-C<sub>4</sub> multimers formed, expression of both proteins together was tested for an additive effect on CsgG assembly. The genes encoding CsgE and CsgF were cloned into pACYC DUET-1 and transformed into BL21 (DE3) containing pET DUET-1 CsgG-C<sub>4</sub> to allow the expression of all 3 proteins. Co-expression of CsgEF did not enhance CsgG assembly above that of cells expressing individually expressed CsgE or CsgF (Figure 4.14c).



**Figure 4.14: CsgE and CsgF both increase the assembly efficiency of CsgG into multimers.** (A) Genetic organization of the genes required for curli biogenesis. (B) BL21 (DE3) strains harbouring plasmids expressing GspD, CsgE-CsgG or CsgF-CsgG were grown to OD ~0.6 before the addition of 0.1 mM IPTG. Whole cell lysates were prepared from cells harvested at the indicated time points and samples analysed by 3-14% SDS-PAGE. (C) BL21 (DE3) strains harbouring plasmids expressing GspD, CsgE-CsgF-CsgG, CsgE-CsgG or CsgF-CsgG were grown to OD ~0.6 before the addition of 0.1 mM IPTG. Whole cell lysates were prepared from cells harvested at the indicated time points and samples analysed by 3-14% SDS-PAGE.

## **4.4 DISCUSSION**

### **4.4.1 New architectures for outer membrane proteins**

The prevailing dogma has been that the proteome of the bacterial outer membrane consisted of integral  $\beta$ -barrels and peripheral lipoproteins. These  $\beta$ -barrel proteins are assembled into the outer membrane by the BAM and the TAM complexes, by a proposed mechanism of  $\beta$ -augmentation catalyzed by Omp85 proteins BamA and TamA [67,60,58]. Peripheral membrane proteins are targeted to the outer membrane by the acylated N-terminus which is targeted by the Lol machinery [187,4]. Upon docking to the LolB receptor at the outer membrane, these acyl-groups attach the lipoproteins to the inner leaflet of the outer membrane. It has become clear that a third group of proteins exist in the bacterial outer membrane.

Electron microscopy has revealed the presence of more complicated outer membrane proteins consisting of large homo-oligomeric pores, with extended periplasmic vestibules, that provide a channel for the secretion of proteins, polysaccharides and amyloid fibres. These pores include secretins in the case of T2SS, T3SS and T4P, Wza, GfcC or PelC in the case of polysaccharide export machines, and CsgG in the case of curli fibre extrusion. The X-ray structure of the capsular polysaccharide secretion pore Wza shed further light on the nature of these channels showing an  $\alpha$ -helical topology for an integral bacterial outer membrane protein [19,188]. Recently the structure of CsgG was solved and unlike Wza showed  $\beta$ -stranded topology of the transmembrane segments of the translocation channel [20,21]. Given this structural diversity it was intriguing to speculate whether these secretion pores are assembled by the BAM or TAM into the outer membrane; in the case of secretins, several studies have tried to address this with differing conclusions [118,119,44]. In any case, there

are many obstacles these proteins must overcome for assembly into the outer membrane. These critical steps of assembly are addressed in the following section.

#### ***4.4.2 Targeting of monomers to the outer membrane***

First, each monomer must be translocated to the outer membrane before efficient assembly into a functional pore. In the case of secretins like PulD, a ‘piggyback’ model has been proposed in which the secretin-pilotin heterodimer are translocated to the outer membrane by the Lol pathway [108,189]. Some secretins are themselves lipoproteins and are self-targeted to the outer membrane presumably by the Lol pathway [116,115]. Mutations that prevent the lipidation of pilotins result in the targeting of the pilotin to the periplasm rather than the outer membrane [97,146-148] and this is also the case for the non-lipidated CsgG mutant [20]. Alternatively, non-acylated mutants of Wza [190] and a Wza-related protein PelC [191] are targeted partly to the periplasm with some pores remaining at the outer membrane showing reduced stability. This is consistent with work in this thesis, showing that a non-acylated mutant of Wza (Wza\*) is still targeted to the outer membrane in BamA expressing or depleted cells. A potential explanation for this observation is that an unidentified pilotin-like protein exists for Wza and that the acylated N-terminus does not compete against this protein for translocation by the Lol pathway.

An additional critical step is for the unfolded protein in transit to avoid degradation. Previous studies have shown that secretins can be subject to proteolysis in the periplasm if not protected by the binding of its cognate pilotin or accessory protein and that these proteins remain a part of the secretin pore once assembled into the outer membrane [79,86,87,102,189]. The presence of accessory factors has also been observed to cause structural rearrangements of the secretion pore [192]. In a *pilP* deletion mutant of *N.*

*gonorrhoea*, PilQ showed an altered symmetry (14 to 19 fold) and reduced stability resulting in PilQ becoming susceptible to degradation [117]. Both CsgE and CsgF are periplasmic chaperones required for the targeting of curli subunits CsgA and CsgB respectively to the curli secretion apparatus and have been shown to interact with CsgG [153,186]. A 150 kDa intermediate of CsgG has been previously observed in cells lacking individual proteins involved in curli biosynthesis (CsgABEF) whilst overexpressing CsgG [185]. Over expression of CsgE and CsgF increases the amount of CsgG oligomers including this 150 kDa assembly intermediate into the outer membrane. Furthermore, the cryo-EM structure of the CsgG-CsgE complex showed that CsgE forms an oligomeric cap on the periplasmic side of the CsgG pore [20]. It is possible that the impact of the small periplasmic proteins CsgE and CsgF on enhancing the assembly of CsgG could be similar to the stabilisation of some secretins by their cognate pilotin or accessory proteins.

#### ***4.4.3 Integration of the secretion pore into the outer membrane***

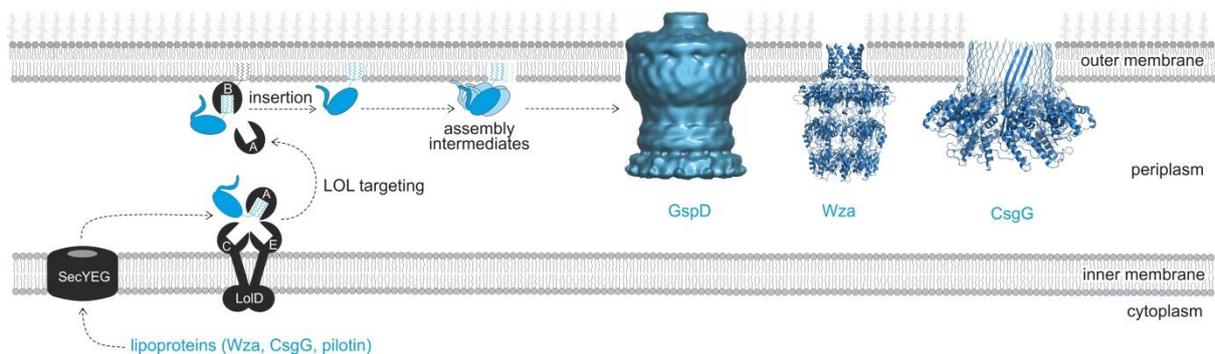
The final critical step in the assembly process is the integration of corresponding transmembrane segments into the plane of the outer membrane. The BAM and the TAM complexes facilitate the assembly of  $\beta$ -barrel proteins, perhaps by a proposed mechanism of beta augmentation [67,60,58]. Previous studies have shown that secretins PulD (*K. oxytoca* T2SS), XcpQ (*P. aeruginosa* T2SS), PscC (*P. aeruginosa* T3SS) and PilQ (*P. aeruginosa* T4P) do not require BamA for assembly into multimers [118,119]. In contrast, the PilQ homolog from *N. meningitidis* fails to assemble if *bamA* is shut down [44]. The assays described in this thesis showed that secretion pores GspD, Wza and CsgG could assemble into the outer membrane in the absence of BamA. Urea extraction assays also showed that these pores were integrally associated into the outer membrane in BamA depleted conditions

suggesting that BamA is not required for the assembly or insertion of these pores into the outer membrane.

A combination of biochemical, biophysical and structural studies have suggested that the assembly of these secretion pores may involve key conformational changes of the multimer. Structural studies on CsgG have demonstrated that it can exist in two variant conformations: a soluble pre-integration form without exposed segments and the membrane-integrated pore in which  $\beta$ -strands from each of the nine monomers form the transmembrane channel [20]. Independent studies on Wza and the homologous protein GfcC from *E. coli* demonstrate that these oligomeric pores also exist as membrane-integrated and pre-integrated species [19,193]. The crystal structure of GfcC showed that the C-terminal  $\alpha$ -helix packs against the protein in a closed conformation that would prevent its integration into the outer membrane [193]. The observation of pre- and post-integration forms of CsgG and the similar structural comparisons between Wza and GfcC, put forward a working model for the assembly of homo-oligomeric secretion pores into bacterial outer membranes.

Assembly intermediates have been observed for secretins, Wza and CsgG here and by others [194-196,20], encouraging speculation that these represent these pre-integration forms. An extension of this concept is that for secretins, pilotins can shield the presumptive transmembrane segment, releasing it during membrane integration [79,102]. It has also been reported that the efficiency of PulD assembly is enhanced into outer membrane vesicles (OMVs) pre-loaded with PulS compared to 'empty' OMVs [189]. A speculative working model for the assembly of these pores is first the recognition of either the cognate pilotin, or the protein subunit itself, by LolA leading to the delivery of the monomer to the outer membrane. Upon delivery to the outer membrane and in a concentration dependant fashion,

the subunits initiate oligomerization promoting a proposed change in conformation of the pore to drive integration of amphipathic segments of the protein into the outer membrane for secretins and CsgG (Figure 4.15) [194,195,20]. It is also possible that a yet to be identified assembly machinery catalyzes the oligomerization of monomers mediating the assembly of these diverse secretion pores into the outer membrane. It is possible that common machinery exists for all of these proteins, or individual factors that drive the assembly of these pores. In any case, with assays now available to measure the assembly process screens for putative assembly machinery and their mechanistic detail can now be pursued.



**Figure 4.15: A model for the assembly of homo-oligomeric secretion pores.** The first crucial step in the assembly of secretion pores is the translocation of each monomer to the outer membrane. In the case of secretins, this process often requires the presence of a pilotin. Lipoproteins Wza, CsgG and the pilotins that bind secretins such as GspD, are transported into the periplasm via the SecYEG translocon. Because secretion channels such as Wza and CsgG are lipoproteins, each monomer is bound by LolA and transferred to LolB, for anchoring into the inner leaflet of the outer membrane. Secretion pore subunits such as GspD interact with the LolA and LolB via their cognate lipoprotein pilotin. It remains to be determined whether some secretion pore subunits such as XcpQ [119] use alternative targeting mechanisms, or that other exogenous factors assist the assembly and membrane integration of multimeric pores. The final step would involve a conformational change of the pre-integration pore or assembly intermediate to allow the insertion of transmembrane segments of the oligomer into the outer membrane to form a functional secretion pore. Such pre-integration species have been structurally characterized in the case of CsgG [20] and the Wza homolog GfcC [193].

## CHAPTER 5: Conclusion

This study has led to the identification of the novel pilotin AspS, demonstrating that AspS forms a new class of pilotins. Assays monitoring the assembly of GspD showed that AspS is required for the targeting of GspD to the outer membrane and for the efficient assembly of GspD multimers. The crystal structure shows that AspS adopts a novel structural fold for pilotin like proteins, and despite these structural differences, AspS still interacts with the S-domain of GspD for targeting to the outer membrane.

A *bamA* depletion strain of *E. coli* was used to assess the role of BamA in assembling oligomeric outer membrane proteins GspD, Wza and CsgG. Optimization of the *bamA* depletion strain in *E. coli* MC4100A showed that after 3 hours of BamA depletion, OMP assembly is significantly reduced yet bacterial cell viability remains relatively high. Analysis of the proteome of the outer membrane revealed that the depletion of BamA also resulted in the loss of TamA, the outer membrane component of the TAM complex. The levels of BamD did not change and the BAM lipoproteins remained modular at the outer membrane after BamA depletion.

Time course assays showed that GspD, Wza and CsgG can assemble into multimers in the outer membrane in the absence of the BAM complex or the TAM. These assays also showed that periplasmic proteins CsgE and CsgF could enhance the assembly efficiency of both the mature CsgG multimer and a possible CsgG assembly intermediate. The use of these assays

can now be used to interrogate the role of specific assembly pathways or screen for putative assembly machinery involved in the biogenesis of outer membrane proteins.

## REFERENCES:

1. Silhavy TJ, Kahne D, Walker S (2010) The bacterial cell envelope. *Cold Spring Harbor perspectives in biology* 2 (5):a000414. doi:10.1101/cshperspect.a000414
2. Cronan JE (2003) Bacterial membrane lipids: where do we stand? *Annual review of microbiology* 57:203-224. doi:10.1146/annurev.micro.57.030502.090851
3. Dalbey RE, Wang P, Kuhn A (2011) Assembly of bacterial inner membrane proteins. *Annual review of biochemistry* 80:161-187. doi:10.1146/annurev-biochem-060409-092524
4. Okuda S, Tokuda H (2011) Lipoprotein sorting in bacteria. *Annual review of microbiology* 65:239-259. doi:10.1146/annurev-micro-090110-102859
5. Ruiz N, Kahne D, Silhavy TJ (2006) Advances in understanding bacterial outer-membrane biogenesis. *Nature reviews Microbiology* 4 (1):57-66. doi:10.1038/nrmicro1322
6. Gerlach RG, Hensel M (2007) Protein secretion systems and adhesins: the molecular armory of Gram-negative pathogens. *International journal of medical microbiology : IJMM* 297 (6):401-415. doi:10.1016/j.ijmm.2007.03.017
7. Nikaido H (2003) Molecular basis of bacterial outer membrane permeability revisited. *Microbiology and molecular biology reviews : MMBR* 67 (4):593-656
8. Denks K, Vogt A, Sachelaru I, Petriman NA, Kudva R, Koch HG (2014) The Sec translocon mediated protein transport in prokaryotes and eukaryotes. *Molecular membrane biology* 31 (2-3):58-84. doi:10.3109/09687688.2014.907455
9. Collinson I, Corey RA, Allen WJ (2015) Channel crossing: how are proteins shipped across the bacterial plasma membrane? *Philosophical transactions of the Royal Society of London Series B, Biological sciences* 370 (1679). doi:10.1098/rstb.2015.0025
10. Berks BC (2015) The twin-arginine protein translocation pathway. *Annual review of biochemistry* 84:843-864. doi:10.1146/annurev-biochem-060614-034251
11. Wu T, Malinverni J, Ruiz N, Kim S, Silhavy TJ, Kahne D (2005) Identification of a multicomponent complex required for outer membrane biogenesis in *Escherichia coli*. *Cell* 121 (2):235-245. doi:10.1016/j.cell.2005.02.015
12. Malojcic G, Andres D, Grabowicz M, George AH, Ruiz N, Silhavy TJ, Kahne D (2014) LptE binds to and alters the physical state of LPS to catalyze its assembly at the cell surface. *Proceedings of the National Academy of Sciences of the United States of America* 111 (26):9467-9472. doi:10.1073/pnas.1402746111
13. Matsuyama S, Yokota N, Tokuda H (1997) A novel outer membrane lipoprotein, LolB (HemM), involved in the LolA (p20)-dependent localization of lipoproteins to the outer membrane of *Escherichia coli*. *The EMBO journal* 16 (23):6947-6955. doi:10.1093/emboj/16.23.6947
14. Du D, van Veen HW, Luisi BF (2015) Assembly and operation of bacterial tripartite multidrug efflux pumps. *Trends in microbiology* 23 (5):311-319. doi:10.1016/j.tim.2015.01.010
15. Baldi DL, Higginson EE, Hocking DM, Praszquier J, Cavaliere R, James CE, Bennett-Wood V, Azzopardi KI, Turnbull L, Lithgow T, Robins-Browne RM, Whitchurch CB, 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*. *Infection and immunity* 80 (6):2042-2052. doi:10.1128/IAI.06160-11
16. Tokuda H, Matsuyama S, Tanaka-Masuda K (2007) Structure, Function, and Transport of Lipoproteins in *Escherichia coli*, p 67-79. In Ehrmann M (ed). ASM Press, Washington, D.C. doi:DC. doi: 10.1128/9781555815806.ch4

17. Michaelis S, Chapon C, D'Enfert C, Pugsley AP, Schwartz M (1985) Characterization and expression of the structural gene for pullulanase, a maltose-inducible secreted protein of *Klebsiella pneumoniae*. *Journal of bacteriology* 164 (2):633-638
18. d'Enfert C, Chapon C, Pugsley AP (1987) Export and secretion of the lipoprotein pullulanase by *Klebsiella pneumoniae*. *Molecular microbiology* 1 (1):107-116
19. Dong C, Beis K, Nesper J, Brunkan-Lamontagne AL, Clarke BR, Whitfield C, Naismith JH (2006) Wza the translocon for *E. coli* capsular polysaccharides defines a new class of membrane protein. *Nature* 444 (7116):226-229. doi:10.1038/nature05267
20. Goyal P, Krasteva PV, Van Gerven N, Gubellini F, Van den Broeck I, Troupiotis-Tsailaki A, Jonckheere W, Pehau-Arnaudet G, Pinkner JS, Chapman MR, Hultgren SJ, Howorka S, Fronzes R, Remaut H (2014) Structural and mechanistic insights into the bacterial amyloid secretion channel CsgG. *Nature* 516 (7530):250-253. doi:10.1038/nature13768
21. Cao B, Zhao Y, Kou Y, Ni D, Zhang XC, Huang Y (2014) Structure of the nonameric bacterial amyloid secretion channel. *Proceedings of the National Academy of Sciences of the United States of America* 111 (50):E5439-5444. doi:10.1073/pnas.1411942111
22. Sankaran K, Wu HC (1994) Lipid modification of bacterial prolipoprotein. Transfer of diacylglyceryl moiety from phosphatidylglycerol. *The Journal of biological chemistry* 269 (31):19701-19706
23. Yamaguchi K, Yu F, Inouye M (1988) A single amino acid determinant of the membrane localization of lipoproteins in *E. coli*. *Cell* 53 (3):423-432
24. Narita S, Tokuda H (2006) An ABC transporter mediating the membrane detachment of bacterial lipoproteins depending on their sorting signals. *FEBS letters* 580 (4):1164-1170. doi:10.1016/j.febslet.2005.10.038
25. Miyamoto S, Tokuda H (2007) Diverse effects of phospholipids on lipoprotein sorting and ATP hydrolysis by the ABC transporter LolCDE complex. *Biochimica et biophysica acta* 1768 (7):1848-1854. doi:10.1016/j.bbamem.2007.04.005
26. Seydel A, Gounon P, Pugsley AP (1999) Testing the '+2 rule' for lipoprotein sorting in the *Escherichia coli* cell envelope with a new genetic selection. *Molecular microbiology* 34 (4):810-821
27. Okuda S, Tokuda H (2009) Model of mouth-to-mouth transfer of bacterial lipoproteins through inner membrane LolC, periplasmic LolA, and outer membrane LolB. *Proceedings of the National Academy of Sciences of the United States of America* 106 (14):5877-5882. doi:10.1073/pnas.0900896106
28. Ito Y, Kanamaru K, Taniguchi N, Miyamoto S, Tokuda H (2006) A novel ligand bound ABC transporter, LolCDE, provides insights into the molecular mechanisms underlying membrane detachment of bacterial lipoproteins. *Molecular microbiology* 62 (4):1064-1075. doi:10.1111/j.1365-2958.2006.05378.x
29. Matsuyama S, Tajima T, Tokuda H (1995) A novel periplasmic carrier protein involved in the sorting and transport of *Escherichia coli* lipoproteins destined for the outer membrane. *The EMBO journal* 14 (14):3365-3372
30. Tokuda H, Matsuyama S (2004) Sorting of lipoproteins to the outer membrane in *E. coli*. *Biochimica et biophysica acta* 1694 (1-3):IN1-9
31. Harms N, Koningstein G, Dontje W, Muller M, Oudega B, Luirink J, de Cock H (2001) The early interaction of the outer membrane protein phoE with the periplasmic chaperone Skp occurs at the cytoplasmic membrane. *The Journal of biological chemistry* 276 (22):18804-18811. doi:10.1074/jbc.M011194200
32. Ureta AR, Endres RG, Wingreen NS, Silhavy TJ (2007) Kinetic analysis of the assembly of the outer membrane protein LamB in *Escherichia coli* mutants each lacking a secretion or targeting factor in a different cellular compartment. *Journal of bacteriology* 189 (2):446-454. doi:10.1128/JB.01103-06

33. Robert V, Volokhina EB, Senf F, Bos MP, Van Gelder P, Tommassen J (2006) Assembly factor Omp85 recognizes its outer membrane protein substrates by a species-specific C-terminal motif. *PLoS biology* 4 (11):e377. doi:10.1371/journal.pbio.0040377
34. Struyve M, Moons M, Tommassen J (1991) Carboxy-terminal phenylalanine is essential for the correct assembly of a bacterial outer membrane protein. *Journal of molecular biology* 218 (1):141-148
35. 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. *Journal of molecular biology* 269 (4):473-478. doi:10.1006/jmbi.1997.1069
36. Rizzitello AE, Harper JR, Silhavy TJ (2001) Genetic evidence for parallel pathways of chaperone activity in the periplasm of *Escherichia coli*. *Journal of bacteriology* 183 (23):6794-6800. doi:10.1128/JB.183.23.6794-6800.2001
37. Sklar JG, Wu T, Kahne D, Silhavy TJ (2007) Defining the roles of the periplasmic chaperones SurA, Skp, and DegP in *Escherichia coli*. *Genes & development* 21 (19):2473-2484. doi:10.1101/gad.1581007
38. Schafer U, Beck K, Muller M (1999) Skp, a molecular chaperone of gram-negative bacteria, is required for the formation of soluble periplasmic intermediates of outer membrane proteins. *The Journal of biological chemistry* 274 (35):24567-24574
39. Vertommen D, Ruiz N, Leverrier P, Silhavy TJ, Collet JF (2009) Characterization of the role of the *Escherichia coli* periplasmic chaperone SurA using differential proteomics. *Proteomics* 9 (9):2432-2443. doi:10.1002/pmic.200800794
40. Qu J, Mayer C, Behrens S, Holst O, Kleinschmidt JH (2007) The trimeric periplasmic chaperone Skp of *Escherichia coli* forms 1:1 complexes with outer membrane proteins via hydrophobic and electrostatic interactions. *Journal of molecular biology* 374 (1):91-105. doi:10.1016/j.jmb.2007.09.020
41. Ruiz-Perez F, Henderson IR, Nataro JP (2010) Interaction of FkpA, a peptidyl-prolyl cis/trans isomerase with EspP autotransporter protein. *Gut microbes* 1 (5):339-344. doi:10.4161/gmic.1.5.13436
42. CastilloKeller M, Misra R (2003) Protease-deficient DegP suppresses lethal effects of a mutant OmpC protein by its capture. *Journal of bacteriology* 185 (1):148-154
43. Ge X, Wang R, Ma J, Liu Y, Ezemaduka AN, Chen PR, Fu X, Chang Z (2014) DegP primarily functions as a protease for the biogenesis of beta-barrel outer membrane proteins in the Gram-negative bacterium *Escherichia coli*. *The FEBS journal* 281 (4):1226-1240. doi:10.1111/febs.12701
44. Voulhoux R, Bos MP, Geurtsen J, Mols M, Tommassen J (2003) Role of a highly conserved bacterial protein in outer membrane protein assembly. *Science* 299 (5604):262-265. doi:10.1126/science.1078973
45. Heinz E, Lithgow T (2014) A comprehensive analysis of the Omp85/TpsB protein superfamily structural diversity, taxonomic occurrence, and evolution. *Frontiers in microbiology* 5:370. doi:10.3389/fmicb.2014.00370
46. Gentle IE, Burri L, Lithgow T (2005) Molecular architecture and function of the Omp85 family of proteins. *Molecular microbiology* 58 (5):1216-1225. doi:10.1111/j.1365-2958.2005.04906.x
47. Webb CT, Heinz E, Lithgow T (2012) Evolution of the beta-barrel assembly machinery. *Trends in microbiology* 20 (12):612-620. doi:10.1016/j.tim.2012.08.006
48. Sanchez-Pulido L, Devos D, Genevrois S, Vicente M, Valencia A (2003) POTRA: a conserved domain in the FtsQ family and a class of beta-barrel outer membrane proteins. *Trends in biochemical sciences* 28 (10):523-526

49. Kim S, Malinverni JC, Sliz P, Silhavy TJ, Harrison SC, Kahne D (2007) Structure and function of an essential component of the outer membrane protein assembly machine. *Science* 317 (5840):961-964. doi:10.1126/science.1143993
50. Knowles TJ, Jeeves M, Bobat S, Dancea F, McClelland D, Palmer T, Overduin M, Henderson IR (2008) Fold and function of polypeptide transport-associated domains responsible for delivering unfolded proteins to membranes. *Molecular microbiology* 68 (5):1216-1227. doi:10.1111/j.1365-2958.2008.06225.x
51. Hagan CL, Kim S, Kahne D (2010) Reconstitution of outer membrane protein assembly from purified components. *Science* 328 (5980):890-892. doi:10.1126/science.1188919
52. Onufryk C, Crouch ML, Fang FC, Gross CA (2005) Characterization of six lipoproteins in the sigmaE regulon. *Journal of bacteriology* 187 (13):4552-4561. doi:10.1128/JB.187.13.4552-4561.2005
53. Malinverni JC, Werner J, Kim S, Sklar JG, Kahne D, Misra R, Silhavy TJ (2006) YfiO stabilizes the YaeT complex and is essential for outer membrane protein assembly in *Escherichia coli*. *Molecular microbiology* 61 (1):151-164. doi:10.1111/j.1365-2958.2006.05211.x
54. Sklar JG, Wu T, Gronenberg LS, Malinverni JC, Kahne D, Silhavy TJ (2007) Lipoprotein SmpA is a component of the YaeT complex that assembles outer membrane proteins in *Escherichia coli*. *Proceedings of the National Academy of Sciences of the United States of America* 104 (15):6400-6405. doi:10.1073/pnas.0701579104
55. Tellez R, Jr., Misra R (2012) Substitutions in the BamA beta-barrel domain overcome the conditional lethal phenotype of a DeltabamB DeltabamE strain of *Escherichia coli*. *Journal of bacteriology* 194 (2):317-324. doi:10.1128/JB.06192-11
56. Charlson ES, Werner JN, Misra R (2006) Differential effects of yfgL mutation on *Escherichia coli* outer membrane proteins and lipopolysaccharide. *Journal of bacteriology* 188 (20):7186-7194. doi:10.1128/JB.00571-06
57. Rigel NW, Schwalm J, Ricci DP, Silhavy TJ (2012) BamE modulates the *Escherichia coli* beta-barrel assembly machine component BamA. *Journal of bacteriology* 194 (5):1002-1008. doi:10.1128/JB.06426-11
58. Noinaj N, Kuszak AJ, Gumbart JC, Lukacik P, Chang H, Easley NC, Lithgow T, Buchanan SK (2013) Structural insight into the biogenesis of beta-barrel membrane proteins. *Nature* 501 (7467):385-390. doi:10.1038/nature12521
59. Ni D, Wang Y, Yang X, Zhou H, Hou X, Cao B, Lu Z, Zhao X, Yang K, Huang Y (2014) Structural and functional analysis of the beta-barrel domain of BamA from *Escherichia coli*. *FASEB journal : official publication of the Federation of American Societies for Experimental Biology* 28 (6):2677-2685. doi:10.1096/fj.13-248450
60. Noinaj N, Kuszak AJ, Balusek C, Gumbart JC, Buchanan SK (2014) Lateral opening and exit pore formation are required for BamA function. *Structure* 22 (7):1055-1062. doi:10.1016/j.str.2014.05.008
61. Selkrig J, Mosbahi K, Webb CT, Belousoff MJ, Perry AJ, Wells TJ, Morris F, Leyton DL, Totsika M, Phan MD, Celik N, Kelly M, Oates C, Hartland EL, Robins-Browne RM, Ramarathinam SH, Purcell AW, Schembri MA, Strugnell RA, Henderson IR, Walker D, Lithgow T (2012) Discovery of an archetypal protein transport system in bacterial outer membranes. *Nature structural & molecular biology* 19 (5):506-510, S501. doi:10.1038/nsmb.2261
62. Heinz E, Selkrig J, Belousoff MJ, Lithgow T (2015) Evolution of the Translocation and Assembly Module (TAM). *Genome biology and evolution* 7 (6):1628-1643. doi:10.1093/gbe/evv097
63. Shen HH, Leyton DL, Shiota T, Belousoff MJ, Noinaj N, Lu J, Holt SA, Tan K, Selkrig J, Webb CT, Buchanan SK, Martin LL, Lithgow T (2014) Reconstitution of a nanomachine

- driving the assembly of proteins into bacterial outer membranes. *Nature communications* 5:5078. doi:10.1038/ncomms6078
64. Burall LS, Harro JM, Li X, Lockett CV, Himpsl SD, Hebel JR, Johnson DE, Mobley HL (2004) *Proteus mirabilis* genes that contribute to pathogenesis of urinary tract infection: identification of 25 signature-tagged mutants attenuated at least 100-fold. *Infection and immunity* 72 (5):2922-2938
65. Struve C, Forestier C, Krogfelt KA (2003) Application of a novel multi-screening signature-tagged mutagenesis assay for identification of *Klebsiella pneumoniae* genes essential in colonization and infection. *Microbiology* 149 (Pt 1):167-176
66. Selkrig J, Belousoff MJ, Headey SJ, Heinz E, Shiota T, Shen HH, Beckham SA, Bamert RS, Phan MD, Schembri MA, Wilce MC, Scanlon MJ, Strugnell RA, Lithgow T (2015) Conserved features in TamA enable interaction with TamB to drive the activity of the translocation and assembly module. *Scientific reports* 5:12905. doi:10.1038/srep12905
67. Gruss F, Zahringer F, Jakob RP, Burmann BM, Hiller S, Maier T (2013) The structural basis of autotransporter translocation by TamA. *Nature structural & molecular biology* 20 (11):1318-1320. doi:10.1038/nsmb.2689
68. Sato K, Naito M, Yukitake H, Hirakawa H, Shoji M, McBride MJ, Rhodes RG, Nakayama K (2010) A protein secretion system linked to bacteroidete gliding motility and pathogenesis. *Proceedings of the National Academy of Sciences of the United States of America* 107 (1):276-281. doi:10.1073/pnas.0912010107
69. Korotkov KV, Gonen T, Hol WG (2011) Secretins: dynamic channels for protein transport across membranes. *Trends in biochemical sciences* 36 (8):433-443. doi:10.1016/j.tibs.2011.04.002
70. Whitfield C (2006) Biosynthesis and assembly of capsular polysaccharides in *Escherichia coli*. *Annual review of biochemistry* 75:39-68. doi:10.1146/annurev.biochem.75.103004.142545
71. Reichow SL, Korotkov KV, Hol WG, Gonen T (2010) Structure of the cholera toxin secretion channel in its closed state. *Nature structural & molecular biology* 17 (10):1226-1232. doi:10.1038/nsmb.1910
72. Korotkov KV, Pardon E, Steyaert J, Hol WG (2009) Crystal structure of the N-terminal domain of the secretin GspD from ETEC determined with the assistance of a nanobody. *Structure* 17 (2):255-265. doi:10.1016/j.str.2008.11.011
73. Koster M, Bitter W, de Cock H, Allaoui A, Cornelis GR, Tommassen J (1997) The outer membrane component, YscC, of the Yop secretion machinery of *Yersinia enterocolitica* forms a ring-shaped multimeric complex. *Molecular microbiology* 26 (4):789-797
74. Collins RF, Ford RC, Kitmitto A, Olsen RO, Tonjum T, Derrick JP (2003) Three-dimensional structure of the *Neisseria meningitidis* secretin PilQ determined from negative-stain transmission electron microscopy. *Journal of bacteriology* 185 (8):2611-2617
75. Linderoth NA, Model P, Russel M (1996) Essential role of a sodium dodecyl sulfate-resistant protein IV multimer in assembly-export of filamentous phage. *Journal of bacteriology* 178 (7):1962-1970
76. Opalka N, Beckmann R, Boisset N, Simon MN, Russel M, Darst SA (2003) Structure of the filamentous phage pIV multimer by cryo-electron microscopy. *Journal of molecular biology* 325 (3):461-470
77. Chami M, Guilvout I, Gregorini M, Remigy HW, Muller SA, Valerio M, Engel A, Pugsley AP, Bayan N (2005) Structural insights into the secretin PulD and its trypsin-resistant core. *The Journal of biological chemistry* 280 (45):37732-37741. doi:10.1074/jbc.M504463200

78. Schraidt O, Marlovits TC (2011) Three-dimensional model of Salmonella's needle complex at subnanometer resolution. *Science* 331 (6021):1192-1195. doi:10.1126/science.1199358
79. Hardie KR, Lory S, Pugsley AP (1996) Insertion of an outer membrane protein in *Escherichia coli* requires a chaperone-like protein. *The EMBO journal* 15 (5):978-988
80. Weiss MS, Wacker T, Weckesser J, Welte W, Schulz GE (1990) The three-dimensional structure of porin from *Rhodobacter capsulatus* at 3 Å resolution. *FEBS letters* 267 (2):268-272
81. Pautsch A, Schulz GE (1998) Structure of the outer membrane protein A transmembrane domain. *Nature structural biology* 5 (11):1013-1017. doi:10.1038/2983
82. Spagnuolo J, Opalka N, Wen WX, Gagic D, Chabaud E, Bellini P, Bennett MD, Norris GE, Darst SA, Russel M, Rakonjac J (2010) Identification of the gate regions in the primary structure of the secretin pIV. *Molecular microbiology* 76 (1):133-150. doi:10.1111/j.1365-2958.2010.07085.x
83. Disconzi E, Guilvout I, Chami M, Masi M, Huysmans GH, Pugsley AP, Bayan N (2014) Bacterial secretins form constitutively open pores akin to general porins. *Journal of bacteriology* 196 (1):121-128. doi:10.1128/JB.00750-13
84. Van der Meeren R, Wen Y, Van Gelder P, Tommassen J, Devreese B, Savvides SN (2013) New insights into the assembly of bacterial secretins: structural studies of the periplasmic domain of XcpQ from *Pseudomonas aeruginosa*. *The Journal of biological chemistry* 288 (2):1214-1225. doi:10.1074/jbc.M112.432096
85. Brok R, Van Gelder P, Winterhalter M, Ziese U, Koster AJ, de Cock H, Koster M, Tommassen J, Bitter W (1999) The C-terminal domain of the *Pseudomonas* secretin XcpQ forms oligomeric rings with pore activity. *Journal of molecular biology* 294 (5):1169-1179. doi:10.1006/jmbi.1999.3340
86. Guilvout I, Hardie KR, Sauvonnet N, Pugsley AP (1999) Genetic dissection of the outer membrane secretin PulD: are there distinct domains for multimerization and secretion specificity? *Journal of bacteriology* 181 (23):7212-7220
87. Nouwen N, Ranson N, Saibil H, Wolpensinger B, Engel A, Ghazi A, Pugsley AP (1999) Secretin PulD: association with pilot PulS, structure, and ion-conducting channel formation. *Proceedings of the National Academy of Sciences of the United States of America* 96 (14):8173-8177
88. Spreter T, Yip CK, Sanowar S, Andre I, Kimbrough TG, Vuckovic M, Pfuetzner RA, Deng W, Yu AC, Finlay BB, Baker D, Miller SI, Strynadka NC (2009) A conserved structural motif mediates formation of the periplasmic rings in the type III secretion system. *Nature structural & molecular biology* 16 (5):468-476. doi:10.1038/nsmb.1603
89. Shevchik VE, Robert-Baudouy J, Condemine G (1997) Specific interaction between OutD, an *Erwinia chrysanthemi* outer membrane protein of the general secretory pathway, and secreted proteins. *The EMBO journal* 16 (11):3007-3016. doi:10.1093/emboj/16.11.3007
90. Reichow SL, Korotkov KV, Gonen M, Sun J, Delarosa JR, Hol WG, Gonen T (2011) The binding of cholera toxin to the periplasmic vestibule of the type II secretion channel. *Channels (Austin)* 5 (3):215-218
91. Pineau C, Guschinskaya N, Robert X, Gouet P, Ballut L, Shevchik VE (2014) Substrate recognition by the bacterial type II secretion system: more than a simple interaction. *Molecular microbiology* 94 (1):126-140. doi:10.1111/mmi.12744
92. Korotkov KV, Johnson TL, Jobling MG, Pruneda J, Pardon E, Heroux A, Turley S, Steyaert J, Holmes RK, Sandkvist M, Hol WG (2011) Structural and functional studies on the interaction of GspC and GspD in the type II secretion system. *PLoS pathogens* 7 (9):e1002228. doi:10.1371/journal.ppat.1002228

93. Login FH, Fries M, Wang X, Pickersgill RW, Shevchik VE (2010) A 20-residue peptide of the inner membrane protein OutC mediates interaction with two distinct sites of the outer membrane secretin OutD and is essential for the functional type II secretion system in *Erwinia chrysanthemi*. *Molecular microbiology* 76 (4):944-955. doi:10.1111/j.1365-2958.2010.07149.x
94. Hu NT, Hung MN, Chen DC, Tsai RT (1998) Insertion mutagenesis of XpsD, an outer-membrane protein involved in extracellular protein secretion in *Xanthomonas campestris* pv. *campestris*. *Microbiology* 144 ( Pt 6):1479-1486. doi:10.1099/00221287-144-6-1479
95. Helm RA, Barnhart MM, Seifert HS (2007) pilQ Missense mutations have diverse effects on PilQ multimer formation, piliation, and pilus function in *Neisseria gonorrhoeae*. *Journal of bacteriology* 189 (8):3198-3207. doi:10.1128/JB.01833-06
96. Guilvout I, Nickerson NN, Chami M, Pugsley AP (2011) Multimerization-defective variants of dodecameric secretin PulD. *Research in microbiology* 162 (2):180-190. doi:10.1016/j.resmic.2011.01.006
97. Daefler S, Russel M (1998) The *Salmonella typhimurium* InvH protein is an outer membrane lipoprotein required for the proper localization of InvG. *Molecular microbiology* 28 (6):1367-1380
98. Okon M, Moraes TF, Lario PI, Creagh AL, Haynes CA, Strynadka NC, McIntosh LP (2008) Structural characterization of the type-III pilot-secretin complex from *Shigella flexneri*. *Structure* 16 (10):1544-1554. doi:10.1016/j.str.2008.08.006
99. Nickerson NN, Tosi T, Dessen A, Baron B, Raynal B, England P, Pugsley AP (2011) Outer membrane targeting of secretin PulD protein relies on disordered domain recognition by a dedicated chaperone. *The Journal of biological chemistry* 286 (45):38833-38843. doi:10.1074/jbc.M111.279851
100. Gu S, Rehman S, Wang X, Shevchik VE, Pickersgill RW (2012) Structural and functional insights into the pilotin-secretin complex of the type II secretion system. *PLoS pathogens* 8 (2):e1002531. doi:10.1371/journal.ppat.1002531
101. Koo J, Burrows LL, Howell PL (2012) Decoding the roles of pilotins and accessory proteins in secretin escort services. *FEMS microbiology letters* 328 (1):1-12. doi:10.1111/j.1574-6968.2011.02464.x
102. Schuch R, Maurelli AT (2001) MxiM and MxiJ, base elements of the Mxi-Spa type III secretion system of *Shigella*, interact with and stabilize the MxiD secretin in the cell envelope. *Journal of bacteriology* 183 (24):6991-6998. doi:10.1128/JB.183.24.6991-6998.2001
103. Trindade MB, Job V, Contreras-Martel C, Pelicic V, Dessen A (2008) Structure of a widely conserved type IV pilus biogenesis factor that affects the stability of secretin multimers. *Journal of molecular biology* 378 (5):1031-1039. doi:10.1016/j.jmb.2008.03.028
104. Lario PI, Pfuetschner RA, Frey EA, Creagh L, Haynes C, Maurelli AT, Strynadka NC (2005) Structure and biochemical analysis of a secretin pilot protein. *The EMBO journal* 24 (6):1111-1121. doi:10.1038/sj.emboj.7600610
105. Izore T, Perdu C, Job V, Attree I, Faudry E, Dessen A (2011) Structural characterization and membrane localization of ExsB from the type III secretion system (T3SS) of *Pseudomonas aeruginosa*. *Journal of molecular biology* 413 (1):236-246. doi:10.1016/j.jmb.2011.07.043
106. Koo J, Tang T, Harvey H, Tammam S, Sampaleanu L, Burrows LL, Howell PL (2013) Functional mapping of PilF and PilQ in the *Pseudomonas aeruginosa* type IV pilus system. *Biochemistry* 52 (17):2914-2923. doi:10.1021/bi3015345
107. Guilvout I, Chami M, Engel A, Pugsley AP, Bayan N (2006) Bacterial outer membrane secretin PulD assembles and inserts into the inner membrane in the absence of its pilotin. *The EMBO journal* 25 (22):5241-5249. doi:10.1038/sj.emboj.7601402

108. Collin S, Guilvout I, Nickerson NN, Pugsley AP (2011) Sorting of an integral outer membrane protein via the lipoprotein-specific Lol pathway and a dedicated lipoprotein pilotin. *Molecular microbiology* 80 (3):655-665. doi:10.1111/j.1365-2958.2011.07596.x
109. Ast VM, Schoenhofen IC, Langen GR, Stratilo CW, Chamberlain MD, Howard SP (2002) Expression of the ExeAB complex of *Aeromonas hydrophila* is required for the localization and assembly of the ExeD secretion port multimer. *Molecular microbiology* 44 (1):217-231
110. Strozen TG, Stanley H, Gu Y, Boyd J, Bagdasarian M, Sandkvist M, Howard SP (2011) Involvement of the GspAB complex in assembly of the type II secretion system secretin of *Aeromonas* and *Vibrio* species. *Journal of bacteriology* 193 (9):2322-2331. doi:10.1128/JB.01413-10
111. Li G, Howard SP (2010) ExeA binds to peptidoglycan and forms a multimer for assembly of the type II secretion apparatus in *Aeromonas hydrophila*. *Molecular microbiology* 76 (3):772-781. doi:10.1111/j.1365-2958.2010.07138.x
112. Vanderlinde EM, Zhong S, Li G, Martynowski D, Grochulski P, Howard SP (2014) Assembly of the type two secretion system in *Aeromonas hydrophila* involves direct interaction between the periplasmic domains of the assembly factor ExeB and the secretin ExeD. *PloS one* 9 (7):e102038. doi:10.1371/journal.pone.0102038
113. Condemine G, Shevchik VE (2000) Overproduction of the secretin OutD suppresses the secretion defect of an *Erwinia chrysanthemi* outB mutant. *Microbiology* 146 ( Pt 3):639-647. doi:10.1099/00221287-146-3-639
114. Bose N, Taylor RK (2005) Identification of a TcpC-TcpQ outer membrane complex involved in the biogenesis of the toxin-coregulated pilus of *Vibrio cholerae*. *Journal of bacteriology* 187 (7):2225-2232. doi:10.1128/JB.187.7.2225-2232.2005
115. Schmidt SA, Bieber D, Ramer SW, Hwang J, Wu CY, Schoolnik G (2001) Structure-function analysis of BfpB, a secretin-like protein encoded by the bundle-forming-pilus operon of enteropathogenic *Escherichia coli*. *Journal of bacteriology* 183 (16):4848-4859. doi:10.1128/JB.183.16.4848-4859.2001
116. Viarre V, Cascales E, Ball G, Michel GP, Filloux A, Voulhoux R (2009) HxcQ liposecretin is self-piloted to the outer membrane by its N-terminal lipid anchor. *The Journal of biological chemistry* 284 (49):33815-33823. doi:10.1074/jbc.M109.065938
117. Jain S, Moscicka KB, Bos MP, Pachulec E, Stuart MC, Keegstra W, Boekema EJ, van der Does C (2011) Structural characterization of outer membrane components of the type IV pili system in pathogenic *Neisseria*. *PloS one* 6 (1):e16624. doi:10.1371/journal.pone.0016624
118. Collin S, Guilvout I, Chami M, Pugsley AP (2007) YaeT-independent multimerization and outer membrane association of secretin PulD. *Molecular microbiology* 64 (5):1350-1357. doi:10.1111/j.1365-2958.2007.05743.x
119. Hoang HH, Nickerson NN, Lee VT, Kazimirova A, Chami M, Pugsley AP, Lory S (2011) Outer membrane targeting of *Pseudomonas aeruginosa* proteins shows variable dependence on the components of Bam and Lol machineries. *mBio* 2 (6). doi:10.1128/mBio.00246-11
120. Filloux A (2004) The underlying mechanisms of type II protein secretion. *Biochimica et biophysica acta* 1694 (1-3):163-179. doi:10.1016/j.bbamcr.2004.05.003
121. Nivaskumar M, Francetic O (2014) Type II secretion system: a magic beanstalk or a protein escalator. *Biochimica et biophysica acta* 1843 (8):1568-1577. doi:10.1016/j.bbamcr.2013.12.020
122. Pugsley AP, Kornacker MG, Poquet I (1991) The general protein-export pathway is directly required for extracellular pullulanase secretion in *Escherichia coli* K12. *Molecular microbiology* 5 (2):343-352

123. Voulhoux R, Ball G, Ize B, Vasil ML, Lazdunski A, Wu LF, Filloux A (2001) Involvement of the twin-arginine translocation system in protein secretion via the type II pathway. *The EMBO journal* 20 (23):6735-6741. doi:10.1093/emboj/20.23.6735
124. Hirst TR, Holmgren J (1987) Conformation of protein secreted across bacterial outer membranes: a study of enterotoxin translocation from *Vibrio cholerae*. *Proceedings of the National Academy of Sciences of the United States of America* 84 (21):7418-7422
125. Dallas WS, Falkow S (1979) The molecular nature of heat-labile enterotoxin (LT) of *Escherichia coli*. *Nature* 277 (5695):406-407
126. Lathem WW, Grys TE, Witowski SE, Torres AG, Kaper JB, Tarr PI, Welch RA (2002) StcE, a metalloprotease secreted by *Escherichia coli* O157:H7, specifically cleaves C1 esterase inhibitor. *Molecular microbiology* 45 (2):277-288
127. Francetic O, Belin D, Badaut C, Pugsley AP (2000) Expression of the endogenous type II secretion pathway in *Escherichia coli* leads to chitinase secretion. *The EMBO journal* 19 (24):6697-6703. doi:10.1093/emboj/19.24.6697
128. Tauschek M, Gorrell RJ, Strugnell RA, Robins-Browne RM (2002) Identification of a protein secretory pathway for the secretion of heat-labile enterotoxin by an enterotoxigenic strain of *Escherichia coli*. *Proceedings of the National Academy of Sciences of the United States of America* 99 (10):7066-7071. doi:10.1073/pnas.092152899
129. Robins-Browne RM (1987) Traditional enteropathogenic *Escherichia coli* of infantile diarrhea. *Reviews of infectious diseases* 9 (1):28-53
130. Robins-Browne RM, Levine MM, Rowe B, Gabriel EM (1982) Failure to detect conventional enterotoxins in classical enteropathogenic (serotyped) *Escherichia coli* strains of proven pathogenicity. *Infection and immunity* 38 (2):798-801
131. Long-Krug SA, Weikel CS, Tiemens KT, Hewlett EL, Levine MM, Guerrant RL (1984) Does enteropathogenic *Escherichia coli* produce heat-labile enterotoxin, heat-stable enterotoxins a or b, or cholera toxin A subunits? *Infection and immunity* 46 (2):612-614
132. Nesta B, Valeri M, Spagnuolo A, Rosini R, Mora M, Donato P, Alteri CJ, Del Vecchio M, Buccato S, Pezzicoli A, Bertoldi I, Buzzigoli L, Tuscano G, Falduto M, Ripa V, Ashhab Y, Bensi G, Fontana MR, Seib KL, Mobley HL, Pizza M, Soriani M, Serino L (2014) SsIE elicits functional antibodies that impair in vitro mucinase activity and in vivo colonization by both intestinal and extraintestinal *Escherichia coli* strains. *PLoS pathogens* 10 (5):e1004124. doi:10.1371/journal.ppat.1004124
133. Valeri M, Rossi Paccani S, Kasendra M, Nesta B, Serino L, Pizza M, Soriani M (2015) Pathogenic *E. coli* exploits SsIE mucinase activity to translocate through the mucosal barrier and get access to host cells. *PloS one* 10 (3):e0117486. doi:10.1371/journal.pone.0117486
134. Eddy SR (2009) A new generation of homology search tools based on probabilistic inference. *Genome informatics International Conference on Genome Informatics* 23 (1):205-211
135. Juncker AS, Willenbrock H, Von Heijne G, Brunak S, Nielsen H, Krogh A (2003) Prediction of lipoprotein signal peptides in Gram-negative bacteria. *Protein science : a publication of the Protein Society* 12 (8):1652-1662. doi:10.1110/ps.0303703
136. Geer LY, Domrachev M, Lipman DJ, Bryant SH (2002) CDART: protein homology by domain architecture. *Genome research* 12 (10):1619-1623. doi:10.1101/gr.278202
137. Frickey T, Lupas A (2004) CLANS: a Java application for visualizing protein families based on pairwise similarity. *Bioinformatics* 20 (18):3702-3704. doi:10.1093/bioinformatics/bth444
138. Mundell DH, Anselmo CR, Wishnow RM (1976) Factors influencing heat-labile *Escherichia coli* enterotoxin activity. *Infection and immunity* 14 (2):383-388

139. Levine MM, Bergquist EJ, Nalin DR, Waterman DH, Hornick RB, Young CR, 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 (8074):1119-1122
140. Guzman LM, Belin D, Carson MJ, Beckwith J (1995) Tight regulation, modulation, and high-level expression by vectors containing the arabinose PBAD promoter. *Journal of bacteriology* 177 (14):4121-4130
141. Posfai G, Koob MD, Kirkpatrick HA, Blattner FR (1997) Versatile insertion plasmids for targeted genome manipulations in bacteria: isolation, deletion, and rescue of the pathogenicity island LEE of the *Escherichia coli* O157:H7 genome. *Journal of bacteriology* 179 (13):4426-4428
142. Datsenko KA, Wanner BL (2000) One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proceedings of the National Academy of Sciences of the United States of America* 97 (12):6640-6645. doi:10.1073/pnas.120163297
143. Laemmli UK (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* 227 (5259):680-685
144. Chambers MC, Maclean B, Burke R, Amodei D, Ruderman DL, Neumann S, Gatto L, Fischer B, Pratt B, Egertson J, Hoff K, Kessner D, Tasman N, Shulman N, Frewen B, Baker TA, Brusniak MY, Paulse C, Creasy D, Flashner L, Kani K, Moulding C, Seymour SL, Nuwaysir LM, Lefebvre B, Kuhlmann F, Roark J, Rainer P, Detlev S, Hemenway T, Huhmer A, Langridge J, Connolly B, Chadick T, Holly K, Eckels J, Deutsch EW, Moritz RL, Katz JE, Agus DB, MacCoss M, Tabb DL, Mallick P (2012) A cross-platform toolkit for mass spectrometry and proteomics. *Nature biotechnology* 30 (10):918-920. doi:10.1038/nbt.2377
145. Dunstan RA, Heinz E, Wijeyewickrema LC, Pike RN, Purcell AW, Evans TJ, Praszker J, Robins-Browne RM, Strugnell RA, Korotkov KV, Lithgow T (2013) Assembly of the type II secretion system such as found in *Vibrio cholerae* depends on the novel Pilotin AspS. *PLoS pathogens* 9 (1):e1003117. doi:10.1371/journal.ppat.1003117
146. Burghout P, Beckers F, de Wit E, van Boxtel R, Cornelis GR, Tommassen J, Koster M (2004) Role of the pilot protein YscW in the biogenesis of the YscC secretin in *Yersinia enterocolitica*. *Journal of bacteriology* 186 (16):5366-5375. doi:10.1128/JB.186.16.5366-5375.2004
147. Koo J, Tammam S, Ku SY, Sampaleanu LM, Burrows LL, Howell PL (2008) PilF is an outer membrane lipoprotein required for multimerization and localization of the *Pseudomonas aeruginosa* Type IV pilus secretin. *Journal of bacteriology* 190 (21):6961-6969. doi:10.1128/JB.00996-08
148. Strozen TG, Li G, Howard SP (2012) YghG (GspSbeta) is a novel pilot protein required for localization of the GspSbeta type II secretion system secretin of enterotoxigenic *Escherichia coli*. *Infection and immunity* 80 (8):2608-2622. doi:10.1128/IAI.06394-11
149. Tosi T, Nickerson NN, Mollica L, Jensen MR, Blackledge M, Baron B, England P, Pugsley AP, Dessen A (2011) Pilotin-secretin recognition in the type II secretion system of *Klebsiella oxytoca*. *Molecular microbiology* 82 (6):1422-1432. doi:10.1111/j.1365-2958.2011.07896.x
150. Korotkov KV, Hol WG (2013) Crystal structure of the pilotin from the enterohemorrhagic *Escherichia coli* type II secretion system. *Journal of structural biology* 182 (2):186-191. doi:10.1016/j.jsb.2013.02.013
151. Kim K, Oh J, Han D, Kim EE, Lee B, Kim Y (2006) Crystal structure of PilF: functional implication in the type 4 pilus biogenesis in *Pseudomonas aeruginosa*. *Biochemical and biophysical research communications* 340 (4):1028-1038. doi:10.1016/j.bbrc.2005.12.108
152. Dische Z, Shettles LB (1951) A new spectrophotometric test for the detection of methylpentose. *The Journal of biological chemistry* 192 (2):579-582

153. Nenninger AA, Robinson LS, Hultgren SJ (2009) Localized and efficient curli nucleation requires the chaperone-like amyloid assembly protein CsgF. *Proceedings of the National Academy of Sciences of the United States of America* 106 (3):900-905. doi:10.1073/pnas.0812143106
154. Krissinel E, Henrick K (2004) Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. *Acta crystallographica Section D, Biological crystallography* 60 (Pt 12 Pt 1):2256-2268. doi:10.1107/S0907444904026460
155. Das D, Chiu HJ, Farr CL, Grant JC, Jaroszewski L, Knuth MW, Miller MD, Tien HJ, Elsliger MA, Deacon AM, Godzik A, Lesley SA, Wilson IA (2014) Crystal structure of a putative quorum sensing-regulated protein (PA3611) from the *Pseudomonas*-specific DUF4146 family. *Proteins* 82 (6):1086-1092. doi:10.1002/prot.24455
156. Werner J, Misra R (2005) YaeT (Omp85) affects the assembly of lipid-dependent and lipid-independent outer membrane proteins of *Escherichia coli*. *Molecular microbiology* 57 (5):1450-1459. doi:10.1111/j.1365-2958.2005.04775.x
157. Palomino C, Marin E, Fernandez LA (2011) The fimbrial usher FimD follows the SurA-BamB pathway for its assembly in the outer membrane of *Escherichia coli*. *Journal of bacteriology* 193 (19):5222-5230. doi:10.1128/JB.05585-11
158. Jain S, Goldberg MB (2007) Requirement for YaeT in the outer membrane assembly of autotransporter proteins. *Journal of bacteriology* 189 (14):5393-5398. doi:10.1128/JB.00228-07
159. Lehr U, Schutz M, Oberhettinger P, Ruiz-Perez F, Donald JW, Palmer T, Linke D, Henderson IR, Autenrieth IB (2010) C-terminal amino acid residues of the trimeric autotransporter adhesin YadA of *Yersinia enterocolitica* are decisive for its recognition and assembly by BamA. *Molecular microbiology* 78 (4):932-946. doi:10.1111/j.1365-2958.2010.07377.x
160. Bodelon G, Marin E, Fernandez LA (2009) Role of periplasmic chaperones and BamA (YaeT/Omp85) in folding and secretion of intimin from enteropathogenic *Escherichia coli* strains. *Journal of bacteriology* 191 (16):5169-5179. doi:10.1128/JB.00458-09
161. Tashiro Y, Nomura N, Nakao R, Senpuku H, Kariyama R, Kumon H, Kosono S, Watanabe H, Nakajima T, Uchiyama H (2008) Opr86 is essential for viability and is a potential candidate for a protective antigen against biofilm formation by *Pseudomonas aeruginosa*. *Journal of bacteriology* 190 (11):3969-3978. doi:10.1128/JB.02004-07
162. Lenhart TR, Akins DR (2010) *Borrelia burgdorferi* locus BB0795 encodes a BamA orthologue required for growth and efficient localization of outer membrane proteins. *Molecular microbiology* 75 (3):692-709. doi:10.1111/j.1365-2958.2009.07015.x
163. Acosta F, Ferreras E, Berenguer J (2012) The beta-barrel assembly machinery (BAM) is required for the assembly of a primitive S-layer protein in the ancient outer membrane of *Thermus thermophilus*. *Extremophiles : life under extreme conditions* 16 (6):853-861. doi:10.1007/s00792-012-0480-x
164. Reynolds ES (1963) The use of lead citrate at high pH as an electron-opaque stain in electron microscopy. *The Journal of cell biology* 17:208-212
165. Webb CT, Selkrig J, Perry AJ, Noinaj N, Buchanan SK, Lithgow T (2012) Dynamic association of BAM complex modules includes surface exposure of the lipoprotein BamC. *Journal of molecular biology* 422 (4):545-555. doi:10.1016/j.jmb.2012.05.035
166. Clements A, Bursac D, Gatsos X, Perry AJ, Civecristov S, Celik N, Likic VA, Poggio S, Jacobs-Wagner C, Strugnell RA, Lithgow T (2009) The reducible complexity of a mitochondrial molecular machine. *Proceedings of the National Academy of Sciences of the United States of America* 106 (37):15791-15795. doi:10.1073/pnas.0908264106

167. Bos MP, Robert V, Tommassen J (2007) Biogenesis of the gram-negative bacterial outer membrane. *Annual review of microbiology* 61:191-214.  
doi:10.1146/annurev.micro.61.080706.093245
168. Knowles TJ, Scott-Tucker A, Overduin M, Henderson IR (2009) Membrane protein architects: the role of the BAM complex in outer membrane protein assembly. *Nature reviews Microbiology* 7 (3):206-214. doi:10.1038/nrmicro2069
169. Hagan CL, Silhavy TJ, Kahne D (2011) beta-Barrel membrane protein assembly by the Bam complex. *Annual review of biochemistry* 80:189-210. doi:10.1146/annurev-biochem-061408-144611
170. Merdanovic M, Clausen T, Kaiser M, Huber R, Ehrmann M (2011) Protein quality control in the bacterial periplasm. *Annual review of microbiology* 65:149-168.  
doi:10.1146/annurev-micro-090110-102925
171. Hagan CL, Kahne D (2011) The reconstituted *Escherichia coli* Bam complex catalyzes multiple rounds of beta-barrel assembly. *Biochemistry* 50 (35):7444-7446.  
doi:10.1021/bi2010784
172. Rossiter AE, Leyton DL, Tveen-Jensen K, Browning DF, Sevastyanovich Y, Knowles TJ, Nichols KB, Cunningham AF, Overduin M, Schembri MA, Henderson IR (2011) The essential beta-barrel assembly machinery complex components BamD and BamA are required for autotransporter biogenesis. *Journal of bacteriology* 193 (16):4250-4253.  
doi:10.1128/JB.00192-11
173. Ieva R, Tian P, Peterson JH, Bernstein HD (2011) Sequential and spatially restricted interactions of assembly factors with an autotransporter beta domain. *Proceedings of the National Academy of Sciences of the United States of America* 108 (31):E383-391.  
doi:10.1073/pnas.1103827108
174. Roman-Hernandez G, Peterson JH, Bernstein HD (2014) Reconstitution of bacterial autotransporter assembly using purified components. *eLife* 3:e04234.  
doi:10.7554/eLife.04234
175. Hagan CL, Westwood DB, Kahne D (2013) bam Lipoproteins Assemble BamA in vitro. *Biochemistry* 52 (35):6108-6113. doi:10.1021/bi400865z
176. Misra R, Stikeleather R, Gabriele R (2015) In vivo roles of BamA, BamB and BamD in the biogenesis of BamA, a core protein of the beta-barrel assembly machine of *Escherichia coli*. *Journal of molecular biology* 427 (5):1061-1074. doi:10.1016/j.jmb.2014.04.021
177. Rassam P, Copeland NA, Birkholz O, Toth C, Chavent M, Duncan AL, Cross SJ, Housden NG, Kaminska R, Seger U, Quinn DM, Garrod TJ, Sansom MS, Piehler J, Baumann CG, Kleantous C (2015) Supramolecular assemblies underpin turnover of outer membrane proteins in bacteria. *Nature* 523 (7560):333-336. doi:10.1038/nature14461
178. Konovalova A, Perlman DH, Cowles CE, Silhavy TJ (2014) Transmembrane domain of surface-exposed outer membrane lipoprotein RcsF is threaded through the lumen of beta-barrel proteins. *Proceedings of the National Academy of Sciences of the United States of America* 111 (41):E4350-4358. doi:10.1073/pnas.1417138111
179. Casadaban MJ (1976) Transposition and fusion of the lac genes to selected promoters in *Escherichia coli* using bacteriophage lambda and Mu. *Journal of molecular biology* 104 (3):541-555
180. Baba T, Ara T, Hasegawa M, Takai Y, Okumura Y, Baba M, Datsenko KA, Tomita M, Wanner BL, Mori H (2006) Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology* 2:2006 0008.  
doi:10.1038/msb4100050
181. Kunkel TA (1985) Rapid and efficient site-specific mutagenesis without phenotypic selection. *Proceedings of the National Academy of Sciences of the United States of America* 82 (2):488-492

182. Obadia B, Lacour S, Doublet P, Baubichon-Cortay H, Cozzone AJ, Grangeasse C (2007) Influence of tyrosine-kinase Wzc activity on colanic acid production in *Escherichia coli* K12 cells. *Journal of molecular biology* 367 (1):42-53. doi:10.1016/j.jmb.2006.12.048
183. Kelley LA, Mezulis S, Yates CM, Wass MN, Sternberg MJ (2015) The Phyre2 web portal for protein modeling, prediction and analysis. *Nature protocols* 10 (6):845-858. doi:10.1038/nprot.2015.053
184. Hammar M, Arnqvist A, Bian Z, Olsen A, Normark S (1995) Expression of two csg operons is required for production of fibronectin- and congo red-binding curli polymers in *Escherichia coli* K-12. *Molecular microbiology* 18 (4):661-670
185. Epstein EA, Reizian MA, Chapman MR (2009) Spatial clustering of the curlin secretion lipoprotein requires curli fiber assembly. *Journal of bacteriology* 191 (2):608-615. doi:10.1128/JB.01244-08
186. Nennering AA, Robinson LS, Hammer ND, Epstein EA, Badtke MP, Hultgren SJ, Chapman MR (2011) CsgE is a curli secretion specificity factor that prevents amyloid fibre aggregation. *Molecular microbiology* 81 (2):486-499. doi:10.1111/j.1365-2958.2011.07706.x
187. Buddelmeijer N (2015) The molecular mechanism of bacterial lipoprotein modification-how, when and why? *FEMS microbiology reviews* 39 (2):246-261. doi:10.1093/femsre/fuu006
188. Collins RF, Derrick JP (2007) Wza: a new structural paradigm for outer membrane secretory proteins? *Trends in microbiology* 15 (3):96-100. doi:10.1016/j.tim.2007.01.002
189. Collin S, Krehenbrink M, Guilvout I, Pugsley AP (2013) The targeting, docking and anti-proteolysis functions of the secretin chaperone PulS. *Research in microbiology* 164 (5):390-396. doi:10.1016/j.resmic.2013.03.023
190. Nesper J, Hill CM, Paiment A, Harauz G, Beis K, Naismith JH, Whitfield C (2003) Translocation of group 1 capsular polysaccharide in *Escherichia coli* serotype K30. Structural and functional analysis of the outer membrane lipoprotein Wza. *The Journal of biological chemistry* 278 (50):49763-49772. doi:10.1074/jbc.M308775200
191. Vasseur P, Soscia C, Voulhoux R, Filloux A (2007) PelC is a *Pseudomonas aeruginosa* outer membrane lipoprotein of the OMA family of proteins involved in exopolysaccharide transport. *Biochimie* 89 (8):903-915. doi:10.1016/j.biochi.2007.04.002
192. Balasingham SV, Collins RF, Assalkhou R, Homberset H, Frye SA, Derrick JP, Tonjum T (2007) Interactions between the lipoprotein PilP and the secretin PilQ in *Neisseria meningitidis*. *Journal of bacteriology* 189 (15):5716-5727. doi:10.1128/JB.00060-07
193. Sathiyamoorthy K, Mills E, Franzmann TM, Rosenshine I, Saper MA (2011) The crystal structure of *Escherichia coli* group 4 capsule protein GfcC reveals a domain organization resembling that of Wza. *Biochemistry* 50 (24):5465-5476. doi:10.1021/bi101869h
194. Huysmans GH, Guilvout I, Pugsley AP (2013) Sequential steps in the assembly of the multimeric outer membrane secretin PulD. *The Journal of biological chemistry* 288 (42):30700-30707. doi:10.1074/jbc.M113.489112
195. Guilvout I, Chami M, Disconzi E, Bayan N, Pugsley AP, Huysmans GH (2014) Independent domain assembly in a trapped folding intermediate of multimeric outer membrane secretins. *Structure* 22 (4):582-589. doi:10.1016/j.str.2014.02.009
196. Ford RC, Brunkan-LaMontagne AL, Collins RF, Clarke BR, Harris R, Naismith JH, Whitfield C (2009) Structure-function relationships of the outer membrane translocon Wza investigated by cryo-electron microscopy and mutagenesis. *Journal of structural biology* 166 (2):172-182. doi:10.1016/j.jsb.2009.02.005

## ***APPENDIX A1 – Protein sequences used to build the PulS\_OutS like pilotin HMM***

### **PulS sequences**

>gi|131604|sp|P20440.1|PULS\_KLEPN RecName: Full=Pullulanase secretion protein pulS;  
Flags: Precursor  
>gi|50122037|ref|YP\_051204.1| general secretion pathway lipoprotein [Pectobacterium  
atrosepticum SCRI1043]  
>gi|27529247|emb|CAD13130.1| Yts1S protein [Yersinia enterocolitica (type 0:8)]  
>gi|45442618|ref|NP\_994157.1| putative lipoprotein [Yersinia pestis biovar Microtus str.  
91001]  
>gi|3152954|emb|CAA46372.1| outS [Erwinia chrysanthemi]  
>gi|157371346|ref|YP\_001479335.1| putative secretion protein [Serratia proteamaculans 568]  
>gi|157371346|ref|YP\_001479335.1| putative secretion protein [Serratia proteamaculans 568]  
>gi|208811287|ref|ZP\_03253047.1| type II secretion protein [Escherichia coli O157:H7 str.  
EC4206]  
>gi|16759163|ref|NP\_454780.1| hypothetical protein STY0189 [Salmonella enterica subsp.  
enterica serovar Typhi str. CT18]  
>gi|284009154|emb|CBA76186.1| conserved hypothetical protein [Arsenophonus nasoniae]  
>gi|227358058|ref|ZP\_03842400.1| conserved hypothetical protein [Proteus mirabilis ATCC  
29906]  
>gi|27382096|ref|NP\_773625.1| hypothetical protein blr6985 [Bradyrhizobium japonicum  
USDA 110]  
>gi|152968711|ref|YP\_001333820.1| hypothetical protein KPN\_00130 [Klebsiella  
pneumoniae subsp. pneumoniae MGH 78578]  
>gi|110804184|ref|YP\_687703.1| hypothetical protein SFV\_0114 [Shigella flexneri 5 str.  
8401]

### **OutS Sequences**

>gi|251788917|ref|YP\_003003638.1| lipoprotein, PulS/OutS family [Dickeya zeae Ech1591]  
>gi|271501358|ref|YP\_003334383.1| lipoprotein, PulS/OutS family [Dickeya dadantii  
Ech586]  
>gi|242238908|ref|YP\_002987089.1| lipoprotein, PulS/OutS family [Dickeya dadantii  
Ech703]  
>gi|261820708|ref|YP\_003258814.1| lipoprotein, PulS/OutS family [Pectobacterium  
wasabiae WPP163]  
>gi|50122037|ref|YP\_051204.1| general secretion pathway lipoprotein [Pectobacterium  
atrosepticum SCRI1043]  
>gi|253689240|ref|YP\_003018430.1| lipoprotein, PulS/OutS family [Pectobacterium  
carotovorum subsp. carotovorum PC1]  
>gi|227114138|ref|ZP\_03827794.1| general secretion pathway lipoprotein [Pectobacterium  
carotovorum subsp. brasiliensis PBR1692]  
>gi|227327888|ref|ZP\_03831912.1| general secretion pathway lipoprotein [Pectobacterium  
carotovorum subsp. carotovorum WPP14]

>gi|206579648|ref|YP\_002240362.1| pullulanase secretion protein PulS [Klebsiella pneumoniae 342]  
 >gi|288937068|ref|YP\_003441127.1| lipoprotein, PulS/OutS family [Klebsiella variicola At-22]  
 >gi|238893145|ref|YP\_002917879.1| pullulanase-specific type II secretion system outer membrane lipoprotein [Klebsiella pneumoniae NTUH-K2044]  
 >gi|328537618|gb|EGF63837.1| lipoprotein, PulS/OutS family [Klebsiella sp. MS 92-3]  
 >gi|262044784|ref|ZP\_06017829.1| pullulanase secretion protein PulS [Klebsiella pneumoniae subsp. rhinoscleromatis ATCC 13884]  
 >gi|131604|sp|P20440.1|PULS\_KLEPN RecName: Full=Pullulanase secretion protein pulS; Flags: Precursor  
 >gi|10955279|ref|NP\_052620.1| EtpO [Escherichia coli O157:H7 str. Sakai]  
 >gi|208811287|ref|ZP\_03253047.1| type II secretion protein [Escherichia coli O157:H7 str. EC4206]  
 >gi|260718944|ref|YP\_003225085.1| putative type II secretion protein etpO [Escherichia coli O103:H2 str. 12009]  
 >gi|238760739|ref|ZP\_04621856.1| hypothetical protein yaldo0001\_37900 [Yersinia aldovae ATCC 35236]  
 >gi|238756272|ref|ZP\_04617588.1| hypothetical protein yruck0001\_32620 [Yersinia ruckeri ATCC 29473]  
 >gi|238785141|ref|ZP\_04629135.1| hypothetical protein yberc0001\_7400 [Yersinia bercovieri ATCC 43970]  
 >gi|238796147|ref|ZP\_04639658.1| hypothetical protein ymoll0001\_4880 [Yersinia mollaretii ATCC 43969]  
 >gi|157371346|ref|YP\_001479335.1| putative secretion protein [Serratia proteamaculans 568]  
 >gi|293394399|ref|ZP\_06638698.1| general secretion pathway lipoprotein [Serratia odorifera DSM 4582]  
 >gi|292489345|ref|YP\_003532232.1| hypothetical protein EAMY\_2877 [Erwinia amylovora CFBP1430]  
 >gi|292898431|ref|YP\_003537800.1| type II secretion system lipoprotein [Erwinia amylovora ATCC 49946]  
 >gi|123443764|ref|YP\_001007735.1| putative secretion protein [Yersinia enterocolitica subsp. enterocolitica 8081]  
 >gi|259907432|ref|YP\_002647788.1| General secretion pathway lipoprotein [Erwinia pyrifoliae Ep1/96]  
 >gi|283477264|emb|CAY73177.1| Uncharacterized protein yacC precursor [Erwinia pyrifoliae DSM 12163]  
 >gi|188532928|ref|YP\_001906725.1| General secretion pathway lipoprotein [Erwinia tasmaniensis Et1/99]

### **EtpO sequences**

>gi|10955279|ref|NP\_052620.1| EtpO [Escherichia coli O157:H7 str. Sakai]  
 >gi|260718944|ref|YP\_003225085.1| putative type II secretion protein etpO [Escherichia coli O103:H2 str. 12009]  
 >gi|253689240|ref|YP\_003018430.1| lipoprotein, PulS/OutS family [Pectobacterium carotovorum subsp. carotovorum PC1]  
 >gi|227327888|ref|ZP\_03831912.1| general secretion pathway lipoprotein [Pectobacterium carotovorum subsp. carotovorum WPP14]

>gi|291285885|ref|YP\_003502702.1| EptO [Escherichia coli O55:H7 str. CB9615]  
 >gi|227114138|ref|ZP\_03827794.1| general secretion pathway lipoprotein [Pectobacterium carotovorum subsp. brasiliensis PBR1692]  
 >gi|261820708|ref|YP\_003258814.1| lipoprotein, PulS/OutS family [Pectobacterium wasabiae WPP163]  
 >gi|50122037|ref|YP\_051204.1| general secretion pathway lipoprotein [Pectobacterium atrosepticum SCRI1043]  
 >gi|288937068|ref|YP\_003441127.1| lipoprotein, PulS/OutS family [Klebsiella variicola At-22]  
 >gi|206579648|ref|YP\_002240362.1| pullulanase secretion protein PulS [Klebsiella pneumoniae 342]  
 >gi|238893145|ref|YP\_002917879.1| pullulanase-specific type II secretion system outer membrane lipoprotein [Klebsiella pneumoniae NTUH-K2044]  
 >gi|328537618|gb|EGF63837.1| lipoprotein, PulS/OutS family [Klebsiella sp. MS 92-3]  
 >gi|262044784|ref|ZP\_06017829.1| pullulanase secretion protein PulS [Klebsiella pneumoniae subsp. rhinoscleromatis ATCC 13884]  
 >gi|131604|sp|P20440.1|PULS\_KLEPN RecName: Full=Pullulanase secretion protein pulS; Flags: Precursor  
 >gi|242238908|ref|YP\_002987089.1| lipoprotein, PulS/OutS family [Dickeya dadantii Ech703]  
 >gi|307131921|ref|YP\_003883937.1| OutS lipoprotein [Dickeya dadantii 3937]  
 >gi|271501358|ref|YP\_003334383.1| lipoprotein, PulS/OutS family [Dickeya dadantii Ech586]  
 >gi|251788917|ref|YP\_003003638.1| lipoprotein, PulS/OutS family [Dickeya zeae Ech1591]  
 >gi|188532928|ref|YP\_001906725.1| General secretion pathway lipoprotein [Erwinia tasmaniensis Et1/99]  
 >gi|238760739|ref|ZP\_04621856.1| hypothetical protein yaldo0001\_37900 [Yersinia aldovae ATCC 35236]  
 >gi|259907432|ref|YP\_002647788.1| General secretion pathway lipoprotein [Erwinia pyrifoliae Ep1/96]  
 >gi|283477264|emb|CAY73177.1| Uncharacterized protein yacC precursor [Erwinia pyrifoliae DSM 12163]  
 >gi|123443764|ref|YP\_001007735.1| putative secretion protein [Yersinia enterocolitica subsp. enterocolitica 8081]  
 >gi|292489345|ref|YP\_003532232.1| hypothetical protein EAMY\_2877 [Erwinia amylovora CFBP1430]

**APPENDIX A2 – Protein sequence scores from the Pilotin HMM in EPEC**

<b>GI No.</b>	<b>Accension No.</b>	<b>E Value</b>	<b>Description</b>
<b>215485286</b>	<b>YP_002327717.1</b>	<b>1.10e-41</b>	<b>hypothetical protein E2348C_0125 (YacC)</b>
215487727	YP_002330158.1	2.30e-04	hypothetical protein E2348C_2659
215488402	YP_002330833.1	4.30e-04	fused signal transducer for aerotaxis sensory component/methyl accepting chemotaxis component
215486229	YP_002328660.1	4.90e-04	hypothetical protein E2348C_1110
215485614	YP_002328045.1	6.00e-04	copper/silver efflux system outer membrane protein CusC
215488147	YP_002330578.1	7.10e-04	fused phosphoenolpyruvate-protein phosphotransferase PtsP/GAF domain
215489045	YP_002331476.1	7.80e-04	putative inner membrane protein translocase component YidC
215489129	YP_002331560.1	8.60e-04	putative uroporphyrinogen III C-methyltransferase
215488466	YP_002330897.1	9.30e-04	hypothetical protein E2348C_3429
215487730	YP_002330161.1	1.10e-03	hypothetical protein E2348C_2662
215489626	YP_002332057.1	1.20e-03	predicted type I restriction-modification enzyme R subunit
215485323	YP_002327754.1	1.30e-03	deoxyguanosinetriphosphate triphosphohydrolase
215489459	YP_002331890.1	1.50e-03	hypothetical protein E2348C_4437
215487198	YP_002329629.1	1.70e-03	hypothetical protein E2348C_2117
215486877	YP_002329308.1	1.80e-03	phosphoenolpyruvate synthase
215486291	YP_002328722.1	2.00e-03	flagellar hook-associated protein FlgK
215489699	YP_002332130.1	2.00e-03	hypothetical protein E2348C_4685
215488233	YP_002330664.1	2.30e-03	predicted peptidase
215486438	YP_002328869.1	2.90e-03	putative sulfate transporter YchM
215488970	YP_002331401.1	3.10e-03	intimin EaeA
215488191	YP_002330622.1	3.10e-03	peptide chain release factor 2
215488893	YP_002331324.1	3.30e-03	hypothetical protein E2348C_3862
215487283	YP_002329714.1	3.40e-03	tyrosine kinase
215485263	YP_002327694.1	3.50e-03	SecA regulator SecM
215488678	YP_002331109.1	3.60e-03	predicted inner membrane protein
215489510	YP_002331941.1	3.60e-03	predicted Fe-S electron transport protein
215489072	YP_002331503.1	3.60e-03	F0F1 ATP synthase subunit alpha
215485630	YP_002328061.1	3.90e-03	enterobactin synthase subunit F
215487798	YP_002330229.1	4.00e-03	hypothetical protein E2348C_2731
215487164	YP_002329595.1	4.00e-03	predicted sensory kinase in two-component regulatory system with YedW
<b>215488290</b>	<b>YP_002330721.1</b>	<b>4.20e-03</b>	<b>hypothetical protein E2348C_3251 (YghG/AspS)</b>
215485724	YP_002328155.1	4.70e-03	sensor protein KdpD
215485726	YP_002328157.1	4.80e-03	potassium-transporting ATPase subunit B
215486069	YP_002328500.1	4.80e-03	paraquat-inducible protein B
215485953	YP_002328384.1	4.90e-03	predicted capsid scaffolding protein
215486431	YP_002328862.1	4.90e-03	dihydroxyacetone kinase subunit M
215486722	YP_002329153.1	4.90e-03	predicted fimbrial protein-like protein
215486034	YP_002328465.1	5.00e-03	predicted peptidase with chaperone function

215488853	YP_002331284.1	5.00e-03	putative xylose transport
215488633	YP_002331064.1	5.10e-03	FKBP-type peptidyl-prolyl cis-trans isomerase
215487754	YP_002330185.1	5.10e-03	predicted alcohol dehydrogenase in ethanolamineutilization
215487293	YP_002329724.1	5.10e-03	hypothetical protein E2348C_2213
215485768	YP_002328199.1	5.20e-03	hypothetical protein E2348C_0630
215486455	YP_002328886.1	5.20e-03	nitrate/nitrite sensor protein NarX
215487430	YP_002329861.1	5.30e-03	multidrug transporter membrane component/ATP-binding component
215488482	YP_002330913.1	5.40e-03	polynucleotide phosphorylase/polyadenylase
215489659	YP_002332090.1	5.80e-03	conserved inner membrane protein
215486650	YP_002329081.1	5.90e-03	azoreductase
215486103	YP_002328534.1	5.90e-03	predicted outer membrane lipoprotein
215487900	YP_002330331.1	6.00e-03	predicted sensory kinase in two-component system with QseF
215487472	YP_002329903.1	6.00e-03	bifunctional UDP-glucuronic acid decarboxylase/UDP-4-amino-4-deoxy-L-arabinose formyltransferase
215486590	YP_002329021.1	6.10e-03	predicted multidrug-efflux transport protein
215485453	YP_002327884.1	6.20e-03	taurine transporter substrate binding subunit
215486917	YP_002329348.1	6.20e-03	predicted tail fiber protein
215485581	YP_002328012.1	6.30e-03	tRNA 2-selenouridine synthase
215489013	YP_002331444.1	6.30e-03	sensory histidine kinase UhpB
215485978	YP_002328409.1	6.50e-03	hypothetical protein E2348C_0844
215486670	YP_002329101.1	6.70e-03	predicted benzoate transporter
215489334	YP_002331765.1	6.80e-03	zinc resistance protein
215485221	YP_002327652.1	6.80e-03	ATP-dependent helicase HepA
215485612	YP_002328043.1	6.90e-03	sensor kinase CusS
215486206	YP_002328637.1	6.90e-03	Efa1/LifA-like protein
215485367	YP_002327798.1	6.90e-03	membrane-bound lytic murein transglycosylase D
215489673	YP_002332104.1	7.00e-03	hypothetical protein E2348C_4659
215487143	YP_002329574.1	7.10e-03	flagellar hook-length control protein
215488090	YP_002330521.1	7.20e-03	sucrose porin precursor
215485500	YP_002327931.1	7.30e-03	2-carboxybenzaldehyde reductase
215487647	YP_002330078.1	7.40e-03	fused predicted PTS enzymes
215489108	YP_002331539.1	7.40e-03	ATP-dependent RNA helicase RhlB
215485456	YP_002327887.1	7.60e-03	taurine dioxygenase
215485610	YP_002328041.1	7.80e-03	bacteriophage N4 receptor
215485478	YP_002327909.1	8.00e-03	exonuclease subunit SbcC
215488481	YP_002330912.1	8.10e-03	lipoprotein NlpI
215485339	YP_002327770.1	8.10e-03	periplasmic chaperone
215487531	YP_002329962.1	8.40e-03	tRNA pseudouridine synthase A
215489173	YP_002331604.1	8.90e-03	DNA recombination protein RmuC
215485686	YP_002328117.1	9.00e-03	glutamate/aspartate transport system permease GltK
215486246	YP_002328677.1	9.10e-03	outer membrane lipoprotein
215488672	YP_002331103.1	9.10e-03	outer membrane porin HofQ
215488215	YP_002330646.1	9.80e-03	hypothetical protein E2348C_3169

**APPENDIX A3 – Representative Coomassie stained SDS-PAGE of membrane fractions after BamA depletion and expression of GspD, Wza and CsgG.**

