MONASH University

Investigating members of the Omp85 protein superfamily in *Klebsiella pneumoniae*

Von Vergel Ligutum Torres
BSc

A thesis submitted for the degree of Doctor of Philosophy at
Monash University in 2018
Faculty of Medicine, Nursing and Health Sciences
Department of Microbiology

Copyright notice

Abstract

Klebsiella infections are fast-emerging as a huge burden to public health being acquired both in community and nosocomial environments. The bacterium is the aetiological agent linked to severe ailments including pneumonia, septicaemia and meningitis. In *Klebsiella* spp. and other Gram-negative bacteria, members of the Omp85 protein superfamily are universally distributed in the genomes. The Omp85 protein superfamily have been categorised based on shared sequence and structural characteristics with essential functions linked to protein translocation and assembly into the outer membrane (OM)

In laboratory *Escherichia coli* K12 strains, there are only two Omp85 proteins present – BamA & TamA. BamA is a core protein subunit of the β-barrel Assembly Machinery (BAM) complex essential in correct assembly and integration of β-barrel proteins into the OM, whilst TamA is a core protein subunit of the Translocation and Assembly Module (TAM) that acts synergistically with the BAM complex in the assembly of at least some virulence factors. Interestingly, in *Klebsiella pneumoniae* and other species in the genus, additional Omp85 paralogues which I term BamL and BamK (divergent from BamA and TamA), are conserved in the *Klebsiella* genome which are yet to be characterised. In this thesis, I aimed to characterise the function of these *K. pneumoniae* specific Omp85s through different approaches that included: genetics, biochemistry, phylogenomics and molecular modeling.

In Chapter 3, I investigated the genetic regulation of the Omp85 family in *K. pneumoniae* through gene expression studies and also attempted to identify genetic regulators controlling transcription of these genes. In this chapter, we report progress on the expression levels of Omp85s gene under a range of growth conditions in a *K. pneumoniae* model but report cryptic regulation for *bamL* and *bamK*. In Chapter 4, I studied Omp85 member BamL and found that it is almost exclusively found in a two-member operon as the downstream partner. The function of this operon is yes to be determined but both members encode for OM-localised proteins and could represent a novel nanomachine. In Chapter 5, I studied BamK that shared high sequence similarity to BamA and could represent a conserved gene duplication. In this study, I was able to show that *bamK* is a highly conserved element in the core genome of *Klebsiella* spp., and its expression rescues a loss-of-function Δ*bamA* mutant in *K. pneumoniae* and *E. coli* models.

Together, the results of this study sets the groundwork for elucidating the functional roles of these paralogous Omp85 members in *K. pneumoniae.* Understanding how these additional paralogues contribute to outer membrane biogenesis could prove valuable in better treatment of *K. pneumoniae* infections but other species which also harbour similar divergent Omp85 members.

Declaration

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

Signature: ___████████
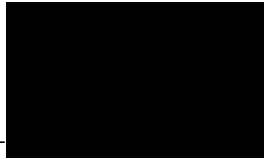
Print Name: Von Torres

Date: 14th August 2018

Preface

Publications during enrolment

-Bi, W., Liu, H., Dunstan, R.A., Li , B., **Torres, V.V.L.**, Cao, J., Chen, L., Wilksch, J.J., Strugnell R.A., Lithgow, T. & Zhou, T. (2017) Extensively Drug-Resistant *Klebsiella pneumoniae* Causing Nosocomial Bloodstream Infections in China: Molecular Investigation of Antibiotic Resistance Determinants, Informing Therapy, and Clinical Outcomes. *Front. Microbiol.* **8**.

-**Torres,V.V. L.**, Heinz, E., Stubenrauch, C. J., Wilksch, J. J., Cao, H., Yang, J., Clements, A., Dunstan, R. A., Felicity, A. , Webb, C. T., Dougan, Gordon, Strugnell, R. A., Hay, I. D. & Litgow, T. (2018) An investigation into the Omp85 protein BamK in hypervirulent *Klebsiella pneumoniae*, and its role in outer membrane biogenesis. *Mol. Microbiol.* **0**: 1-35.

Experiments conducted by others for this work:

- Dr. Ji Yang (The University of Melbourne, Australia) analysed the upstream region of the *bamK* coding region of *Klebsiella pneumoniae* B5055 for Gram-negative promoter elements – Chapter 3.0, Figure 3.2.1

-Dr. David Steer and Dr. Ralf Schittenhelm (Monash University, Australia) performed mass-spectrometry and initial analysis to identify DNA binding protein candidates – Chapter 3.0, Section 3.4 (Appendix 3.0)

-Mr. Liam McIntyre (The University of Melbourne, Australia) performed a high-throughput BLAST search for *bamL* and *lupV* sequences from a manual curation of *Klebsiella pneumoniae* genomes using an in-house developed Python script, output data was analysed by me – Chapter 4.0, Figure 4.1.2

 Dr. Seong Chow (Monash University, Australia) performed the macrophage uptake assay and parallel live cell imaging assays – Chapter 4.0, Figure 4.2.4-5

-Dr. Eva Heinz (The Wellcome Sanger Institute, United Kingdom) performed the bioinformatics analysis illustrating the distribution of *bamA* and *bamK* genes from a globally diverse set of *Klebsiella pneumoniae* genomes – Chapter 5.0, Figure 5.1.4.

-Dr. Hanwei Cao (The University of Melbourne, Australia) performed the mouse infection assays using *Klebsiella pneumoniae* B5055 strains – Chapter 5.0, Figure 5.2.2

-Dr. Chris Stubenrauch constructed *Escherichia coli* mutant strains and performed immunoblot assays – Chapter 5.0, Figure 5.5.4

All other experimentation comprises my original work.

Acknowledgements

I have been lucky to cross paths with many people whilst undertaking my Ph.D. In this section, I'd like to write few words to all of them as I am forever grateful and indebted as they have moulded, encouraged and mentored me both as a young researcher and person through this chapter in my life.

I would first like to thank my main supervisor - **Professor Trevor Lithgow**. Trevor provided me an opportunity to undertake a Ph.D. in his laboratory. He gave me a chance! A supervisor I respect for his approach and breadth of knowledge in science. Throughout this research project, his honesty and unwavering support had saved me from falling to pieces or losing hope when I hit the my personal "lows", but his encouragement and support made me run during the "highs". His pursuit and passion in science is inspiring, and with all that I am eternally thankful for being such a heroic mentor (scientific and personally) and fostering a great environment to be a part of.

To my associate supervisors who provided me unconditional support and advice throughout: **Dr. Iain Hay** and **Dr. Eva Heinz**. Iain has been instrumental in the work presented and was always supportive, willing to teach and patient with my dismal learning curve and "austere" antics. A scientific dossier personified, he had sound advice for any query I had. Eva likewise, was instrumental to this project as she provided the foundations for the thesis questions posed. Constantly checking in to see how I was travelling throughout my project and giving me critical advice/support even when relocating to another hemisphere is something I never expected. Her resolute approach and confidence is something I really admire and hope to develop one day. I have been so lucky to be supervised by this pair.

To my Ph.D. committee: **Professor Julian Rood, Professor Richard Strugnell, Dr John Boyce and Dr. Ralf Schittenhelm**. Their critical advice and helpful scientific discussions and support through the years I am grateful for. To my collaborators who have provided scientific support for this study: **Dr. Ji Yang, Dr. David Steer, Mr. Liam McIntyre, Dr. Seong Chow and Dr. Hanwei Cao** for reasons listed in the previous section.

A shout-out goes towards all my colleagues and friends in the **Lithgow laboratory**, past and present. Special thanks to: **Dr. Dilshan Gunasinghe (office neighbour), Eric Mandela, Dr.**

Victoria Hewitt, Dr. Denisse Leyton, Ms. Rebecca Bamert (lab neighbour/colleague), Dr. Takuya Shiota, Dr. Jonathan Wilksch, Dr. Chaille Webb, Dr. Grishma Vadlamani, Dr. Andrea Rocker and **Ms. Kher-Shing Tan**. All providing friendship, emotional and academic support through the years!

A more personal thank you goes towards: **Dr. Abigail Clements** and **Dr. Felicity Alcock**. A dynamic duo that provided additional groundwork for my Ph.D. work. To **Dr. Chris Stubenrauch** and **Dr. Rhys Dunstan**, a more contemporary dynamic duo of the Lithgow laboratory at the time of writing this section. Chris and Rhys have been saving graces throughout my Ph.D. by being great sources of banter and helping me through the initial drafting of this thesis. Along with the previously mentioned lab members and alumni, I am lucky to call them friends and colleagues.

More thanks go towards our science neighbours: **Traven, Naderer** and **Kwok laboratories**. Special thanks to: **Ms. Tricia Lo, Ms. Julie Nguyen, Dr. Barbara Koch, Dr. Jiyoti Verma-Gaur, Dr Seong Chow, Dr. Pankaj Deo, Dr. Mary Speir, Ms Gilu Abraham** and **Dr. Sam Palframan**. Members of the **Microbiology department** of Monash University. Special thanks to: **Dr. Jackie Cheung** and **Dr. Jessica Wisniewski.**

Finally, to all my **family** and **friends**. For keeping me grounded and sane through this whole journey I sincerely thank you for just being here.

Table of contents

List of Figures

List of Tables

Abbreviations

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

× g	multiplied by gravity
°C	degree Celsius
APS	ammonium persulfate
ATP	adenosine triphosphate
BAM	β-barrel Assemby Machinery
BN	Blue Native
cDNA	complementary DNA
CPS	Capsular polysaccharide
Cryo-EM	cryo-electron microscopy
Ct	cycle threshold
DDM	n-dodecyl -D-maltoside
DMEM	Dulbecco's Modified Eagle's Medium
DNA	Deoxyribonucleic acid
DTT	Dithiothreitol
ECL	enhanced chemiluminescence
EDTA	Ethylenediaminetetraacetic acid
EMSA	Electrophoretic mobility shift assay
g	Gram
GFP	green fluorescent protein
h / hr	Hour
HRP	horseradish peroxidase
IM	Inner Membrane
IMP	Inner Membrane Protein
IPTG	sopropyl β-D-1-thiogalactopyranoside
kDa	kilo Dalton
KpBamA	BamA derived from *K. pneumoniae* B5055
KpBamK	BamK derived from K. pneumoniae B5055
L	Litre
L4	Loop 4 (extracellular)
L6	Loop 6 (extracellular)
L7	Loop 7 (extracellular)
LB	Lysogeny Broth
LB	Lysogeny Broth (Miller)
LC-*MS*	Liquid chromatography–mass spectrometry
LOL	Lolcalisation of lipoprotein
Lpp	Braun's lipoprotein
LPS	Lipopolysaccharide
M	Molar
M9	Minimal medium (M9)
MCS	Multiple Cloning Site
min	Minute
mL	Millilitre
mm	Millimetre
mM	Millimolar
mRNA	Messenger RNA

MS	Mass Spectrometry
MS	mass spectrometry
ng	Nanogram
NMR	Nuclear magnetic resonance
ns	non-significant
OD600	Optical Density 600
OM	Outer Membrane
OMP	Outer Membrane Protein
PAGE	Polyacrylamide gel electrophoresis
PBS	Phosphate-buffered *saline*
PCR	polymerase chain reaction
PMSF	phenylmethane sulfonyl fluoride
POTRA	Polypeptide transport-associated
PVDF	Polyvinylidene fluoride
qPCR	Quantitative reverse transcription PCR
RNA	ribonucleic acid
rpm	Revolutions per minute
s	Second
SAXS	Small-angle X-ray scattering
SD	Standard deviation
SDS	Sodium Dodecyl Sulfate
SEM	Standard error of the mean
T5SS (a-f)	Type 5 Secretion System (a-f)
TAM	Translocation and Assembly Module
TAT	twin-arginine translocon
TB	Terrific Broth
TCA	2,2,2-trichloroacetic acid
TEMED	N,N,N',N'-tetramethylethane-1,2-diamine
TEV	Tobacco Etch Virus
TPS	Two-partner Secretion
wt	Wildtype
β(1-16)	β-strand (1-16)
μg	Microgram
μL	Microlitre
σ^E	Sigma E (Sigma 24)

1

Introduction

<u>1.0 The cell envelope of Gram-negative bacteria</u>

The cell envelope is a multilayered feature and plays a variety of important roles for Gram-negative bacterial survival (Costerton *et al.*, 1974, Silhavy *et al.*, 2010). In Gram-negative bacteria such as *Escherichia coli* and *Klebsiella pneumoniae*, the cell envelope is comprised of two fluid lipid bilayers that differ in composition and function, termed the inner membrane (IM) and the outer membrane (OM) (Figure 1.0.1). The IM and OM are estimated to be ~200 angstroms apart, separated by an aqueous space termed the periplasm. Situated within this periplasmic space is a major structural component of the cell wall, a layer of peptidoglycan that protects against sudden changes in osmolality (Matias *et al.*, 2003, Cohen *et al.*, 2017). The IM is the site of many metabolic and biochemical processes (Hendler & Burgess, 1974, Ingledew & Poole, 1984, Price & Driessen, 2010), the periplasm provides an isolated compartment for many oxidising and degradative reactions (Miller & Salama, 2018), and the OM acts as a semi-permeable barrier permitting diffusion of nutrients and necessary ions while preventing entry of noxious agents such as toxins and antibiotics (Koebnik *et al.*, 2000, Decad & Nikaido, 1976). The IM and peptidoglycan, followed by the OM, together surround the cytoplasm but only the OM is in direct contact with the extracellular environment making it the first line of defence against adverse physical and chemical environments. Together, these layers of the cell define cellular compartments in Gram-negative bacteria but also govern the selective permeability of molecules to and from these compartments to maintain conditions for optimal growth and homeostasis.

The IM is a symmetrical lipid bilayer, with both leaflets composed of phospholipids. This membrane features outer-leaflet embedded lipoproteins and integral membrane proteins with α-helical transmembrane segments (Wallin & Heijne, 2008, Facey & Kuhn, 2004). The activity of various membrane proteins contributes to build a proton gradient across the IM that generates a proton-motive force, providing for energy-dependent processes such as oxidative phosphorylation and active transport to occur at the IM-cytoplasm interface (Erhardt *et al.*, 2014, Du *et al.*, 2015). The anchorage

of lipoproteins in the inner leaflet of the IM is by virtue of lipidation of an N-terminal cysteine residue. These lipoproteins play important roles in periplasmic activities, including assembly of cell envelope components and trans-wall signalling pathways (Tokuda & Narita, 2010, Asmar *et al.*, 2017).

The periplasm is as an aqueous compartment that contains a variety of important soluble proteins as well as the structurally-important peptidoglycan layer (Cohen *et al.*, 2017, Miller & Salama, 2018). The periplasm is devoid of nucleotides; therefore, a common question is how the myriad of energy-dependent processes can occur in the absence of ATP and other obvious energy carriers. Periplasmic proteins have roles in bi-directional transport of substrates across the periplasm, such as small molecules and nutrients imported across the OM that are processed in the periplasm before delivery to the cytoplasm (Beacham, 1979, Liu *et al.*, 2004), the sensing domains of the histidine kinase members of two component systems are frequently located in the periplasm to relay external stress to the cytoplasm (Stock *et al.*, 2000, Dutta & Inouye, 2000), and components of the OM are modified and chaperoned to the OM itself (Lycklama a Nijeholt & Driessen, 2012, Sklar *et al.*, 2007b). The periplasm also contains intrinsic protein quality control capabilities, as several proteases are sequestered in the periplasm where they are responsible for degradation of a variety of substrates, including potentially toxic misfolded cell envelope proteins (Soltes *et al.*, 2017, Miot & Betton, 2004).

The peptidoglycan layer provides shape and rigidity to the bacterial cell wall and is a common target for antibiotics (Vollmer *et al.*, 2008, Katayama *et al.*, 2000). The peptidoglycan can be thought of as a mesh network composed of disaccharide polymers linked by pentapeptides, and this meshwork layer is anchored to the OM by the abundant "Braun's lipoprotein" Lpp (Dramsi *et al.*, 2008). The peptidoglycan layer does not appear to present a major physical barrier for proteins traversing the periplasm, as proteins in the 50-100 kDa range have been shown to pass through without hindrance (Vollmer *et al.*, 2008, Demchick & Koch, 1996).

The OM lipid bilayer composition differs from the IM. The OM is asymmetric in that the inner leaflet is composed of phospholipids, but the outer leaflet is instead built of lipopolysaccharide (LPS) (Koebnik *et al.*, 2000). LPS is crucial for stability of Gram-negative cell walls but also provides protection against the extracellular environment (Zhang *et al.*, 2013). Located alongside LPS in many pathogenic Gram-negative species, especially *K. pneumoniae*, a capsular polysaccharide (CPS) or capsule layer is also present (Yoshida *et al.*, 2000, Willis & Whitfield, 2013). The CPS layer is comprised of high molecular weight polysaccharides that are linear or branched structures of repeating two to seven monosaccharides (Zamze *et al.*, 2002). The CPS layer also confers protection

against environmental pressures (e.g. desiccation, antimicrobial peptides) but also functions as a "cloaking mechanism" that evades the host immune system by physically preventing receptor binding or by active modification of capsular composition mimicking the host thereby preventing recognition (Cress *et al.*, 2014, Doorduijn *et al.*, 2016).

In *E. coli* and *K. pneumoniae*, the vast majority of OM proteins belong to one of two major classes: lipoproteins, peripherally tethered to the inner leaflet of the OM via a lipid moiety attached to a N-terminal cysteine residue (Nakayama *et al.*, 2012) and β-barrel proteins, which are integrally embedded within the OM (Koebnik *et al.*, 2000, Schulz, 2002). It should be noted, that there are a few OMPs that are embedded into the outer membrane through transmembrane α-helices, such as: Wza, a polysaccharide translocon pore component; PorB, an aquaporin found in *Corynebacterium glutamicum* for nutrient exchange and VirB10, a transenvelope spanning component of the Type IV secretion systems found throughout the Gram-negative phyla (Dong *et al.*, 2006, Ziegler *et al.*, 2008, Chandran *et al.*, 2009). For the purpose of this thesis, OMPs will refer solely to outer membrane proteins that adopt a β-barrel topology which are integrally embedded into the OM.

OM localised lipoproteins interact with OMPs, soluble periplasmic proteins, the peptidoglycan layer, and other lipoproteins to accomplish a diverse array of functions. Similarly, several of these β-barrel OMPs are tightly regulated in respect to synthesis, assembly and activity as they also play key roles in many essential cellular processes. These cellular processes include, but are not limited to, transport of essential nutrients and cofactors (e.g. OmpF) (Nikaido & Vaara, 1985), proteolysis (e.g. OmpT) (McCarter *et al.*, 2004), protein secretion (e.g. TolC) (Zgurskaya *et al.*, 2011), adhesion (e.g. Ag43) (Klemm *et al.*, 2004), iron acquisition (e.g. FhuA) (Bonhivers *et al.*, 1998), OM biogenesis (e.g. BamA and LptD) and mechanical support to the overall cell wall (e.g. OmpA) (Voulhoux *et al.*, 2003, Haarmann *et al.*, 2010, Park *et al.*, 2012).
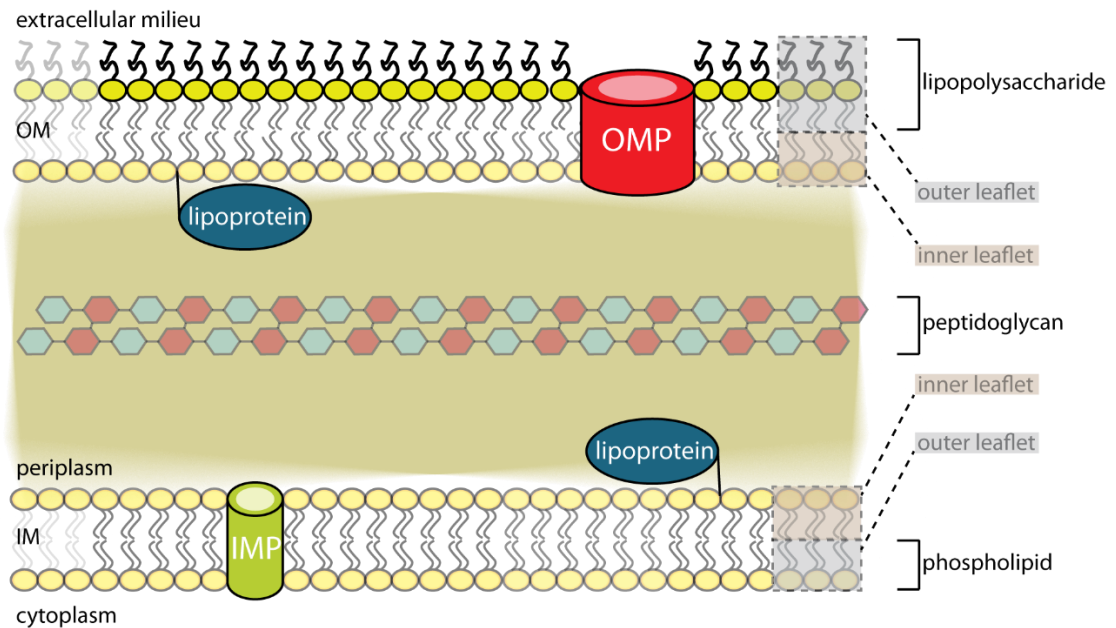
Figure 1.0.1: Schematic of the Gram-negative cell envelope. The Gram-negative cell envelope is composed of two membranes: the inner membrane (IM) and the outer membrane (OM). The two membranes are separated by an aqueous space, termed the periplasm, where the peptidoglycan layer is also situated. The OM is asymmetric at its outer leaflet (cell surface facing) is comprised of lipopolysaccharide and its inner leaflet is made of phospholipids. The IM is symmetrical as both leaflets (inner and outer) are made of phospholipids. Transmembrane proteins embedded within the OM and IM are termed outer membrane proteins (OMPs) and inner membrane proteins (IMPs), respectively. Typically, lipoproteins are anchored through acylation to either membrane (in the periplasmic facing leaflets).

1.1 Outer membrane proteins

Outer membrane proteins (OMPs) are integral membrane proteins that, while having widely dissimilar functions, harbour common structural features, as revealed by a number of physical characteristics that typify this family of proteins. The membrane integrated domain of OMPs is a β-barrel, where a series of anti-parallel β-strands come together such that hydrogen bond donors and acceptors are compensated, ultimately forming a closed cylindrical structure with a lumen (Figure 1.1.1 A). The β-strands are amphipathic in nature, as the residues alternate between hydrophilic and hydrophobic (Wimley, 2002). The side chains of the hydrophilic residues face towards the interior lumen of the β-barrel structure, while the side chains of the hydrophobic residues point outwards interacting with the hydrocarbon tails of the lipid membrane (Buchanan, 1999, Koebnik *et al.*, 2000). To date, all solved crystal structures of OMPs from Gram-negative bacteria contain an even number of antiparallel β-strands, and for monomeric β-barrels the strand number ranges from 8 β-strands (eg. OmpX) to 26 β-strands (eg. LptD) (Vogt & Schulz, 1999, Botos *et al.*, 2016).

The β-strands of OMPs are connected by what are defined as "turns" (facing the periplasm) and "loops" (displayed on the extracellular face of the outer membrane) (Koebnik *et al.*, 2000). While inter-strand turns tend to be very short segments of polypeptide (often 3-4 residues), extracellular loops can be extensive in length and are intrinsically flexible along much of their length. In many cases features in the loops contribute specialised functions that include ligand binding, control of permeability, or protein folding and stability (Buchanan, 1999, Wimley, 2003, Fairman *et al.*, 2011). Another characteristic of β-barrel proteins is a feature referred to as an "aromatic girdle": a series of membrane-facing aromatic side chains that typically reside at the top and bottom boundaries of the barrel at the membrane-water interface (Figure 1.1.1 B). These aromatic residues are important for the assembly and stability of β-barrel proteins into the OM (Deol *et al.*, 2004, Hong *et al.*, 2007).

Some OMPs contain N-terminal and/or C-terminal extramembrane domains resulting in multidomain architectures. Examples include the chaperone domains of fimbrial usher proteins (Waksman & Hultgren, 2009), the peptidoglycan-binding C-terminal domain of OmpA (Samsudin *et al.*, 2016) and virulence-associated autotransporters with N-terminal extensions (Henderson & Nataro, 2001, Henderson *et al.*, 2004). An alternative means to build complex architectures is seen in the case of OMPs that adopt higher order oligomeric states, either homo-oligomeric (e.g. porins) or hetero-oligomeric (e.g. LptD/E, TonB-dependent receptors) (Botos *et al.*, 2016, Noinaj *et al.*, 2010). These higher order oligomeric states are proposed to facilitate the regulatory or substrate induced cooperativity between active sites and binding interfaces or provide a structural purpose such as

scaffold/filamentous and transenvelope architectures (Perica *et al.*, 2014, Hashimoto & Panchenko, 2010). Finally, there are β-barrels where the cylindrical shape is formed from more than one copy of the protein, where each subunit contributes only a small number of β-strands. This category includes the efflux channel protein TolC (a trimer where each subunit contributes 4 β-strands to a 12-stranded β-barrel) (Koronakis *et al.*, 2004), trimeric autotransporters (a trimer where each subunit contributes 4 β-strands to a 12-stranded β-barrel) (Linke *et al.*, 2006, Bassler *et al.*, 2015), and the Type 2 Secretion System secretin complex (a 15-mer where each subunit contributes 4 β-strands to a 60-stranded β-barrel) (Yan *et al.*, 2017, Hay *et al.*, 2017).
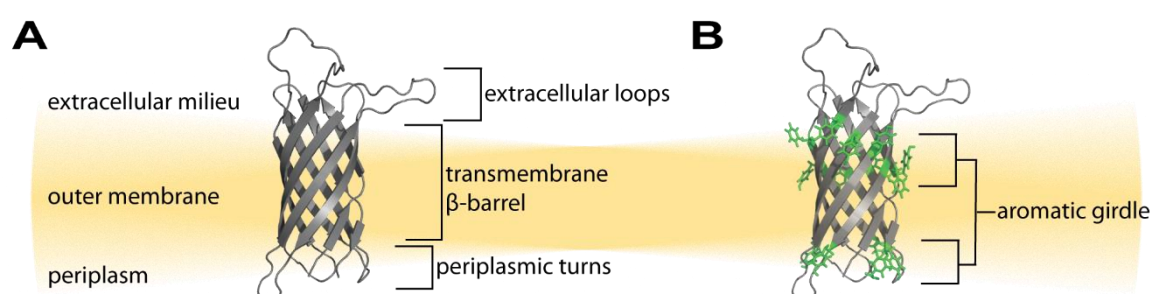
Figure 1.1.1: Anatomy of an outer membrane protein (OMP). (A) Structure of OmpX (PDB: 2MNH) from *E. coli*. OmpX is a β-barrel OMP which is integrally embedded within the Gram-negative OM. The β-strands are anti-parallel and form the barrel-like domain. Connecting each β-strand to the next, the extracellular facing loops are typically longer than the periplasmic turns. (B) The OmpX structure with sidechains indicated for aromatic residues (green) forming the "aromatic girdle" which localise to the membrane-water interface of the asymmetric bilayer, contributing to OMP membrane insertion and structural stability.

As mentioned previously, OMPs perform a variety of cellular functions in Gram-negative bacteria (Wimley, 2003). Based on their function, OMPs can be classified into one of four broad categories: porins (i.e. channels), translocons, enzymes, and structural OMPs (Figure 1.1.2). It should be noted that some OMPs may fall into more than one of these categories. Porins are a class of protein that acts as molecular sieve for small and hydrophilic compounds (<600 kDa), forming a diffusion channel or, in some cases such as LamB, a facilitated diffusion channel (Nikaido & Vaara, 1985). They include some of the most abundant proteins in the Gram-negative OM (e.g. the *E. coli* protein OmpC) and are often found to be trimers where each protein forms a complete β-barrel, but their preferred state in the membrane seems to be as three stably-associated barrels (Rocque & McGroarty, 1989, Cowan *et al.*, 1992). In addition to passive and facilitated diffusion, there are examples of β-barrel proteins that provide for active transport (e.g. BtuB imports Vitamin B12 against a concentration gradient) (Hufton *et al.*, 1995), where these transport channels typically require cooperation with energy transducing complexes (eg. TonB, AcrA/B) in the IM (Brinkman & Larsen, 2008, Zhang *et al.*, 2017). Translocons include OMPs that are involved in the export of proteins or polysaccharides

across the OM with examples including autotransporters and two-partner secretion (TPS) systems. The virulence associated autotransporter and TPS secretion systems of Gram-negative bacteria will be discussed in further detail in this chapter (Section 1.3). Enzymatic OMPs discovered to date have functions that include proteases (OmpT) and phospholipases (e.g. OmpLA) (Vandeputte-Rutten *et al.*, 2001, Snijder & Dijkstra, 2000). The final category includes OMPs that serve structural purposes, contributing to the formation and maintenance of the cell envelope, such as peptidoglycan-OM linkage (e.g. OmpA), capsule-OM linkage (e.g. Wzi) (Bushell *et al.*, 2013). Together, OMPs enable the OM to fulfil its role as a protective physical barrier that enables the selective traffic across the hydrophiid lipid bilayer. Furthermore, some OMPs play direct roles in virulence, extending the importance of studying bacterial OMPs and their biogenesis to medical interest.
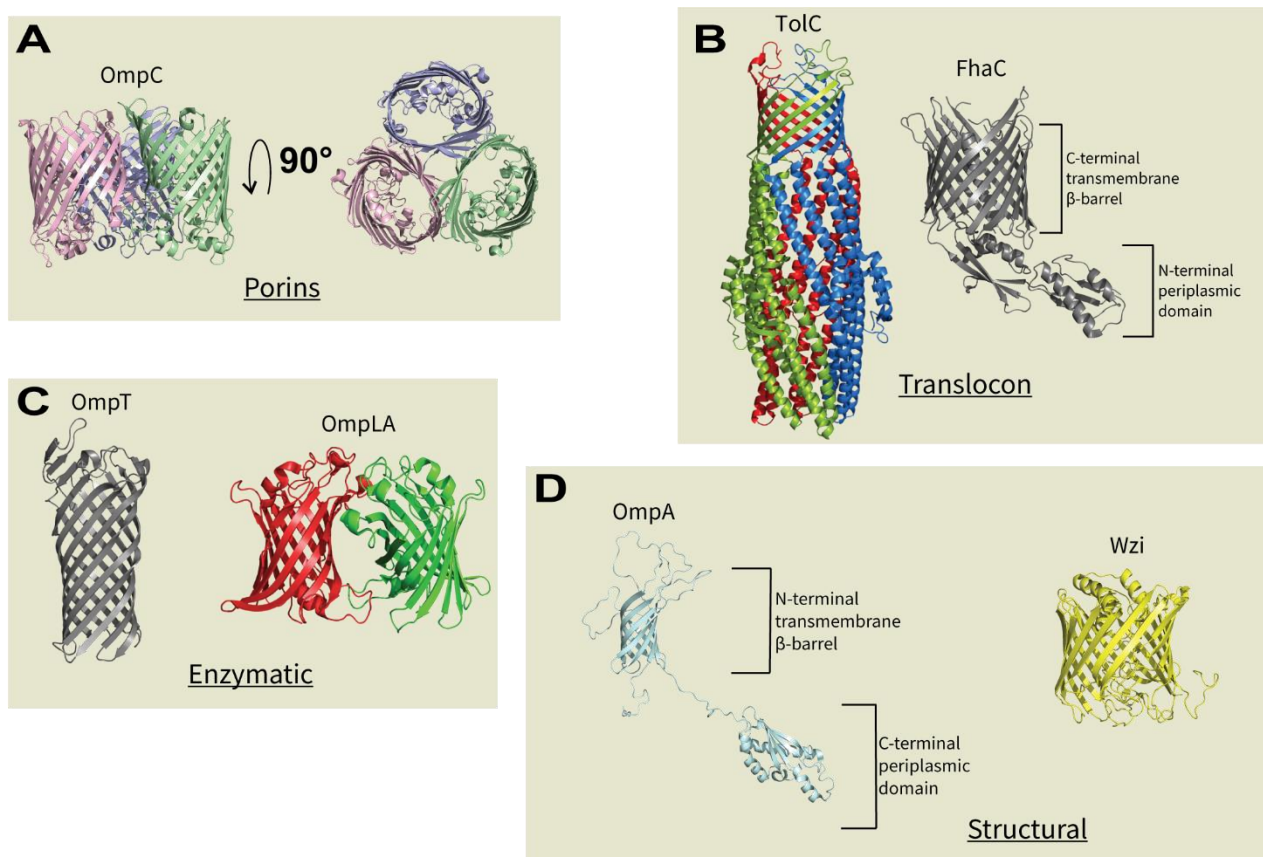
Figure 1.1.2: Diversity of OMPs from Gram-negative bacteria. (A) Porins – Structure of the OmpC trimer (PDB: 2XE5) from *E. coli*. Major porins from Gram-negative bacteria form higher older multimers that are usually homo-oligomeric. Shown as a side view and top-down view, the monomers trimerise into a triangular formation. The lumen of each monomer is relatively small, but it is not obstructed, thereby allowing passage of small solutes into, and out of, the cell (Nikaido, 1994). (B) Translocon Structures of FhaC (PDB: 3NJT) and TolC trimer (PDB: 1EK9) from *E. coli*. Translocons – facilitate the transfer of client substrates across the OM (in or out) but the mechanism between OMPs can differ. FhaC contains soluble N-terminal extensions, which is believed to be involved in the binding of substrates for translocation through the lumen for secretion (Delattre *et al.*, 2011). TolC differs, as the OMP is formed by three monomers each making a third of the channel domain. Not shown in this image is TolC in complex with an IM localised motor system such as AcrAB complex. The AcrAB-TolC multidrug efflux pump spans the entire Gram-negative envelope for active transport of a diverse array of compounds. (C) Enzymatic - Structures of OmpT (PDB: 1I78) and OmpLA dimer (PDB: 1FW2) from *E. coli*. The OmpT monomer contains a cell-surface facing protease domain that cleaves antimicrobial peptides produced during a host immune response to enable infection and is thought to contribute to removing misfolded OMPs (Hui *et al.*, 2010, Lyu & Zhao, 2015). OmpLA represents another enzymatic OMP wherein it dimerises into its active form to degrade phospholipids that have mistargeted flipped to the outer leaflet of the OM, which typically happens during cell stress (e.g. addition of EDTA) (Vaara, 1992, Snijder & Dijkstra, 2000). (D) Structural - Structure of *K. pneumoniae* OmpA (PDB: 2K01 N-terminal and 5NHX C-terminal) and *E. coli* Wzi (PDB: 2YNK). OmpA is a structural OMP involved in anchoring the peptidoglycan cell wall to the outer membrane in Gram-negative bacteria. Wzi is a structural OMP involved in Group 1 capsule assembly in a range of Gram-negative bacterial species. Mutagenesis studies of the extracellular loops of this OMP have been shown to be vital for capsular polysaccharide binding and assembly of a CPS layer the bacterial cell surface (Bushell *et al.*, 2013).

1.2 Outer membrane protein assembly

In Prokaryotes, all proteins are synthesised in the cytoplasm with an estimated 20% being targeted for assembly into membranes (Krogh *et al.*, 2001). The targeting of OMPs from the cytoplasm to the OM requires the action of translocons in the IM and OM and help from a suite of molecular chaperones in the cytoplasm and periplasm (Figure 1.2.1). OMP biogenesis, assembly and insertion of a nascent β-barrel into the OM, thus requires coordinated steps of: (i) OMP translation, targeting to and translocation across the IM, (ii) traversing the periplasm to the OM and (iii) folding and insertion into the OM. In some cases, OMPs can fall off these pathways and would be potentially toxic if accumulated in the periplasm (Goemans *et al.*, 2014, Costello *et al.*, 2016); thus, (iv) quality control and genetic regulatory systems are also in place to protect against these issues.
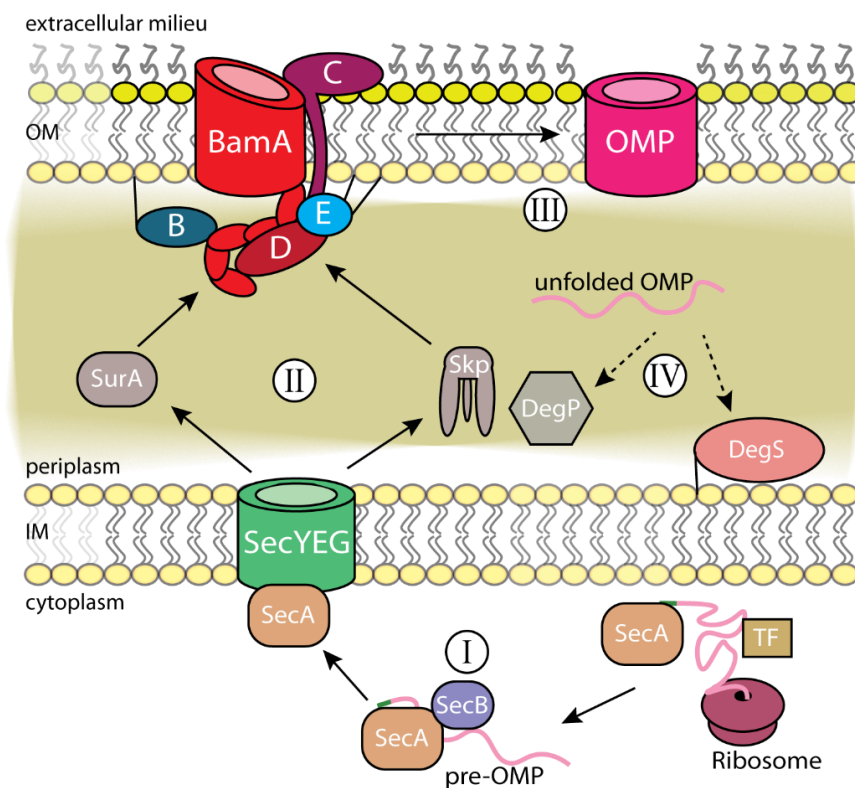
Figure 1.2.1: Overview of outer membrane protein assembly. (I) Newly translated unprocessed OMP precursors (pre-OMPs) emerging from a ribosome are bound by SecA and trigger factor (TF) to keep the pre-OMP in a translocation-competent state. The pre-OMP is subsequently transferred to cytoplasmic chaperone SecB which targets the cargo to the IM localised SecYEG translocon. The SecYEG translcon facilitates the translocation of the pre-OMP across the IM using energy from SecA bound ATPase hydrolysis. At the IM, still in an unfolded form, the pre-OMP is processed where its N-terminal signal sequence (green) is cleaved then released into the periplasm. (II) The nascent OMP then takes either SurA or the Skp/DegP chaperone pathway for transit to the BAM complex of the OM. (III) Through an unknown mechanism, the BAM complex mediates the folding and membrane insertion of OMPs. The energy required for OMPs folding and inserting into the OM is mysterious, as all subcellular compartments extending radially from the IM are devoid of ATP or other high-energy carriers (Wülfing & Plückthun, 1994). (IV) Misfolded or aggregated OMPs in the periplasm are either recognised by DegP and degraded or bound by DegS, which subsequently cleaves the anti-sigma factor, RseB to release σ^E, thereby causing a protease cascade triggering the envelope σ^E-mediated envelope stress response (De Las Peñas *et al.*, 1997, Alba *et al.*, 2002).

OMP translation and Sec pathway translocation across the IM

The necessary information for delivery of an unprocessed OMP (pre-OMP) to the IM is found in intrinsic sequence determinants usually found in the N-terminal region of pre-OMPs, a stretch of amino acids termed the "signal sequence" or "signal peptide" (Driessen & Nouwen, 2008). A typical OMP signal peptide has a tripartite structure with the following organisation: a positively charged N-terminal region, a hydrophobic stretch (typically 5-16 amino acids) followed by a hydrophilic segment which contains a protease cleavage site (Kendall *et al.*, 1986, von Heijne, 1990). In the first instance, newly translated pre-OMPs emerging from the ribosome are bound by ATPase motor SecA and general chaperone trigger factor (TF). SecA has been shown to have affinity to hydrophobic regions of the signal sequence (Gelis *et al.*, 2007, Singh *et al.*, 2014), and TF is believed to bind to "buried" regions of the natively folded proteins to prevent degradation and aggregation of the nascent pre-OMP (Bechtluft *et al.*, 2010, Saio *et al.*, 2014). The bound nascent pre-OMP then binds to chaperone SecB in the cytoplasm, to maintain a soluble/unfolded translocation-competent state, which are then targeted to the SecYEG translocon (Huang *et al.*, 2016). The SecYEG translocon is an IM localised protein complex that catalyses the translocation of unfolded OMPs across the IM from the N- to C- terminus, in an ATP-dependent process powered by the ATPase SecA motor (Kudva *et al.*, 2013). During this IM translocation process from cytoplasm to periplasm, the signal sequence is cleaved from the nascent pre-OMP by Signal Peptidase I (SPI) (Paetzel *et al.*, 2002).

Traversing the periplasm to the OM

Upon reaching the periplasm, nascent OMPs still need to navigate the periplasm for their destined OM localisation. As nascent OMPs are both unfolded and amphipathic in nature, they are considered prone to aggregation if not protected from the aqueous environment (Bulieris *et al.*, 2003, Walton & Sousa, 2004, Walton *et al.*, 2009); thus, periplasmic chaperones bind and escort nascent OMPs until their eventual folding and insertion into the OM (Costello *et al.*, 2016). Three well characterised proteins have been described in the literature that have been shown to play major roles in nascent OMP trafficking from IM to OM: SurA (Lazar & Kolter, 1996), DegP (Krojer *et al.*, 2002) and Skp (Walton *et al.*, 2009). In *E. coli*, single deletions of any three chaperone members are viable, but double deletions of *surA* and *skp*, or *surA* and *degP*, are synthetically lethal (Rizzitello *et al.*, 2001). These previous studies have been interpreted to highlight two distinct pathways for nascent OMP passage across the periplasm, (i) a SurA-mediated pathway and (ii) a Skp/DegP-mediated pathway, where at least one pathway must be functional for bacterial cells to be viable. SurA sequesters client nascent OMPs by preferentially binding exposed aromatic residues (Xu *et al.*, 2007). Aromatic residues are commonly featured in OMP sequences due to "aromatic girdles" (Deol

et al., 2004, Hong *et al.*, 2007). Studies using OmpA, LptD and FhuA as client OMPs have suggested that SurA undergoes oligomerisation and/or conformational switching, enabling the chaperone to bind a multitude of topologically different client OMPs (Ricci *et al.*, 2013, Vertommen *et al.*, 2009, Bitto & McKay, 2002).

The trimeric Skp and hexameric DegP both function as molecular chaperones with defined internal binding cavities to accommodate nascent OMPs for protection against aggregation or degradation, respectively (Krojer *et al.*, 2002, Walton *et al.*, 2009). The Skp homotrimer has a structure that has been compared to a "jellyfish", where its three tentacle-like extended domains have been shown to be important for initial contact with client nascent OMPs (e.g. OmpA, LamB and BtuB) (Bulieris *et al.*, 2003, Lazar & Kolter, 1996, Jarchow *et al.*, 2008). On the other hand, DegP sequesters nascent OMPs in a cage-like cavity but also functions as a protease for degradation of unfolded and mislocalised OMPs as shown through numerous biochemical assays and through cryo-electron microscopy studies (Krojer *et al.*, 2008, Jiang *et al.*, 2008, Merdanovic *et al.*, 2011). The degradative capabilities of DegP is essential for quality control of OM homeostasis as the periplasmic accumulation of unfolded and mislocalised OMPs is toxic to cells (Bulieris *et al.*, 2003, Walton *et al.*, 2009, Walton & Sousa, 2004).

Folding and insertion into the OM

Nascent OMPs which have successfully completed their journey across the periplasm need to fold into the OM. Given their β-barrel structure, the strand-by-strand folding process to create the barrel has to be co-ordinated with the insertion process that would result in the barrel being imbedded within the plane of the OM. This folding and integration process of an OMP barrel is referred to as "β-barrel assembly" or "β-barrel biogenesis". Several *in vitro* studies have found that OMPs are able to spontaneously assemble into artificial lipid environments (Kleinschmidt, 2003, Tamm *et al.*, 2004, Kleinschmidt, 2015, Kleinschmidt, 2006, Gessmann *et al.*, 2014), but a paradox exists in the *in vivo* bacterial cell envelope landscape. The IM and inner leaflet of the OM of Gram-negative bacteria share lipid compositions: ~75% phosphatidylethanolamine (PE), ~20% phosphatidylglycerol (PG) and ~5% cardiolipin (CL) (Diedrich & Cota-Robles, 1974, Sohlenkamp & Geiger, 2015), therefore the spontaneous insertion and folding of OMPs would be expected to mislocalise a large proportion of OMP molecules to the IM (Carlson & Silhavy, 1993, Grabowicz & Silhavy, 2017b). This does not occur. At least part of the answer to this dilemma comes from the presence of chaperones that slow the assembly process, and the presence in the OM of a modular β-barrel assembly machinery that catalyses the efficient folding and insertion of nascent OMPs into the OM (Ricci & Silhavy, 2012,

Selkrig *et al.*, 2014, Noinaj *et al.*, 2015, Albenne & Ieva, 2017). The essential protein BamA (previously known as YaeT or Omp85; (Voulhoux *et al.*, 2003)) is a member of the Omp85/TpsB protein superfamily, which are found in all bacteria with outer membranes (Heinz & Lithgow, 2014). In *E. coli*, BamA has been shown to associate with four lipoproteins BamB, BamC, BamD and BamE (previously known as YfgL, NpB, YfiO and SmpA) to form what is now known as the BAM complex (the core complex of the β-barrel assembly machinery) (Malinverni *et al.*, 2006, Kim *et al.*, 2007, Hagan *et al.*, 2011, Ricci & Silhavy, 2012). How this heteropentameric complex assembles and inserts OMPs is not fully understood given the physical and energetic barriers that must be overcome, and models of BAM complex activity will be discussed later (Section 1.4)

Quality control and genetic regulation of OMPs

Complex biological processes such as OMP biogenesis require quality control, and bacteria have genetic regulation systems in place to monitor for defects and address these issues to maintain OM biogenesis and homeostasis. The σ^E envelope stress responses prevents the accumulation of OMP species that have become misfolded, aggregated or become localised for too long in the periplasm (Chaba *et al.*, 2007, Grabowicz & Silhavy, 2017a). The accumulation of these aberrant species can be toxic if they interfere with the normal homeostatic mechanisms in the periplasm (Bulieris *et al.*, 2003, Walton *et al.*, 2009, Walton & Sousa, 2004). By the current models for quality control, conserved C-terminal aromatic and hydrophobic residues, which are typically inaccessible in correctly folded OMPs, serve as beacons in unfolded OMPs to attract elements of the quality control machinery, particularly the serine protease DegS in the IM (Hasselblatt *et al.*, 2007, Chaba *et al.*, 2011). Sensing accumulation of unfolded OMPs in the periplasm, DegS can activate a signal cascade culminating in the activation of the σ^E envelope stress response regulon. In *E. coli*, the σ^E envelope stress response decreases the production of highly-abundant OMPs such as OmpC, thereby greatly reducing protein flux into the periplasm. Concomitantly, the σ^E envelope stress response promotes the transcription of periplasmic OMP chaperones/proteases, components of the BAM complex, and other factors required for OM homeostasis (Dartigalongue *et al.*, 2001, Onufryk *et al.*, 2005).

While the σ^E envelope stress response is essential in *E. coli* even under non-stress conditions (Papenfort *et al.*, 2006, Bossi & Figueroa-Bossi, 2007), this is not the case in other Gram-negative species, such as *Salmonella enterica* serovar Typhmirium and *Yersinia pseudotuberculosis* (Palonen *et al.*, 2013, McMeechan *et al.*, 2007). It remains unclear whether species such as these have redundant quality control features that diminishes the importance of the σ^E envelope stress response,

or whether alternative regulatory features serve to prevent the crisis situations that species like *E. coli* face by resorting to an essential stress response.

Biogenesis of OM lipoproteins

The biosynthesis and localisation of lipoproteins shares some of the machinery with OMP biogenesis, but there are also some distinct differences (Figure 1.2.2). Firstly, there are some lipoproteins that are translocated across the IM by the twin-arginine translocon (TAT) pathway (Ize *et al.*, 2003, Shruthi *et al.*, 2010, Palmer & Berks, 2012), a pathway never used by OMPs. However, most pro-lipoproteins are targeted to and translocated across the IM by the SEC translocon, using an OMP-like N-terminal signal peptide, but one which features a distinguishing lipobox motif (Andersson & Von Heijne, 1993, Tokuda & Matsuyama, 2004). Upon exiting the SEC translocon, the lipobox motif is recognised by a membrane protein called Lgt that transfers a diacylglyceryl moiety to a cysteine residue within the lipobox motif, thereby anchoring it to the inner leaflet of the IM (Tokunaga *et al.*, 1982, Sankaran & Wu, 1994). Following this modification, Signal Peptidase II (SPII) cleaves the N-terminal signal peptide from the lipoprotein, and Lnt adds stabilising acyl chains to the newly cleaved N-terminus (Yamagata *et al.*, 1983, Gupta & Wu, 1991). The now acylated nascent lipoprotein is then recognised as a substrate by the LOL complex, initiating transfer of the lipoprotein from the IM by the chaperone LolA, to the assembly protein LolB located in the OM (Tokuda & Matsuyama, 2004).
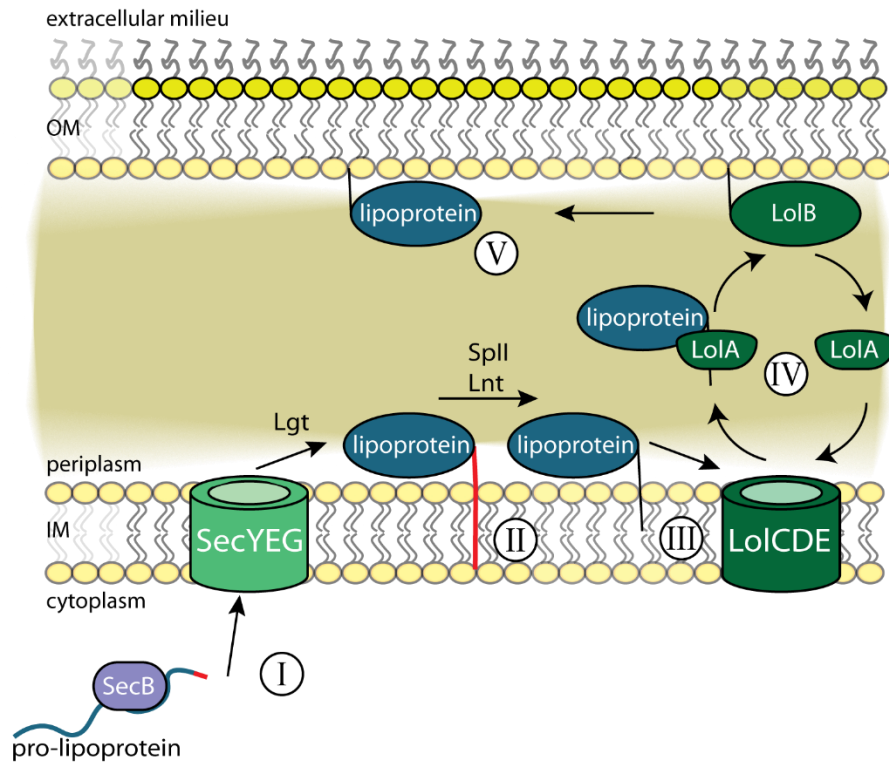
Figure 1.2.2: The OM lipoprotein biogenesis pathway. (I) SEC translocated pro-lipoproteins are acylated by the diacylglyceryltransferase (Lgt) at cysteine found in the lipobox motif. (II) This modification allows type II signal peptidase (SPII) to release the pro-lipoprotein from its signal peptide (shown in red) The newly formed NH_2 group of the +1 cysteine is then acylated by Lnt and the processed lipoprotein is recognised (III) by the LolCDE complex. (IV) The periplasmic chaperone LolA, of the localisation of lipoprotein (LOL) pathway, delivers the lipoprotein from the IM to the acceptor lipoprotein LolB. LolA is continuously recycled through this transfer process. (V) LolB then inserts the processed lipoprotein into the inner leaflet of the OM.

1.3 Omp85/TpsB superfamily

The Omp85/TpsB protein superfamily represents an important group of OMPs in Gram-negative bacteria involved in the biogenesis of OMP substrates (Omp85) or translocate substrates across the OM (TpsB) (Heinz & Lithgow, 2014). This protein superfamily is typified through shared sequence motifs and structural similarities; however, the two groups are distinct in their taxonomic distribution and comprised protein sequences. The taxonomic distribution of Omp85 proteins are found in all Gram-negative bacteria but also in eukaryotes (in mitochondria and plastids, the organelles derived from bacterial endosymbionts), whilst the TpsB members are found exclusively in bacterial lineages (Heinz & Lithgow, 2014). At the sequence level, common sequence motifs dictate a modular structure where at the N-terminal is usually comprised of at least one polypeptide transport-associated (POTRA) domain(s) followed by a C-terminal transmembrane β-barrel domain (Gentle *et al.*, 2005, Clantin *et al.*, 2007). Between the two groups, common sequences which typify their respective β-barrel domains can be delineated into distinct PFAM domains: PF01103 (Bac_Surface_Ag) for Omp85 proteins and PF03865 (ShlB) for TpsB proteins, suggesting separate evolutionary trajectories. In this section I will give a brief overview of the TpsB and Omp85 protein sub-families by discussing their structure, functional role and proposed activities.

TpsB proteins

The anatomy of the TpsB proteins share similarities to the Omp85 proteins (discussed in a later section) but have a few nuanced differences. Comprehensive *in silico* analyses revealed that TpsB proteins can be classified into 2 distinct subsets based on defining characteristics of their N-terminal extensions and/or C-terminal β-barrel domains (Heinz & Lithgow, 2014). The only available structure of a TpsB protein is FhaC of *Bordetella pertussis* (Clantin *et al.*, 2007, Maier *et al.*, 2015). This archetypal protein follows a two-domain architecture: an N-terminal domain of two POTRA repeats, and a 16-stranded C-terminal β-barrel domain (Clantin *et al.*, 2007) (Figure 1.3.1 A). The POTRA domains of TpsB proteins are situated in the periplasm and are required for binding and folding of partner TpsA proteins for translocation (Yang & Braun, 2000, Delattre *et al.*, 2011). Topological studies of a handful of β-barrel domains suggest that TpsB proteins may vary widely in their structures as members are predicted to have 10-16 β-strands (Könninger *et al.*, 1999, Guédin *et al.*, 2000, Jacob-Dubuisson *et al.*, 2004, Surana *et al.*, 2006).

The FhaC β-barrel consists of 16 anti-parallel β-strands where its lumen is occluded by an N-terminal α-helix and its extracellular loop 6 (Clantin *et al.*, 2007, Maier *et al.*, 2015) (Figure 1.3.1 B-C). The exact function of this α-helix is unknown as its deletion from FhaC does not affect its ability to secrete

FHA or adopt a β-barrel fold (Guédin *et al.*, 2000). The α-helix is followed by a 30- residue periplasmic linker in an extended conformation that joins to the N terminus of the membrane-distal POTRA 1 domain (Maier *et al.*, 2015). Unlike the α-helix, the linker is indispensable as deletion abolishes its translocase function (Guédin *et al.*, 2000, Clantin *et al.*, 2007). The L6 of FhaC contains a characteristic VRGY motif of the overall Omp85/TpsB superfamily, where this loop is positioned in the lumen through a salt bridge interaction (Maier *et al.*, 2015, Clantin *et al.*, 2007). Deletion of the FhaC L6 abolishes secretion but does not affect FhaC levels or localisation to the OM (Clantin *et al.*, 2007, Leonard-Rivera & Misra, 2012).
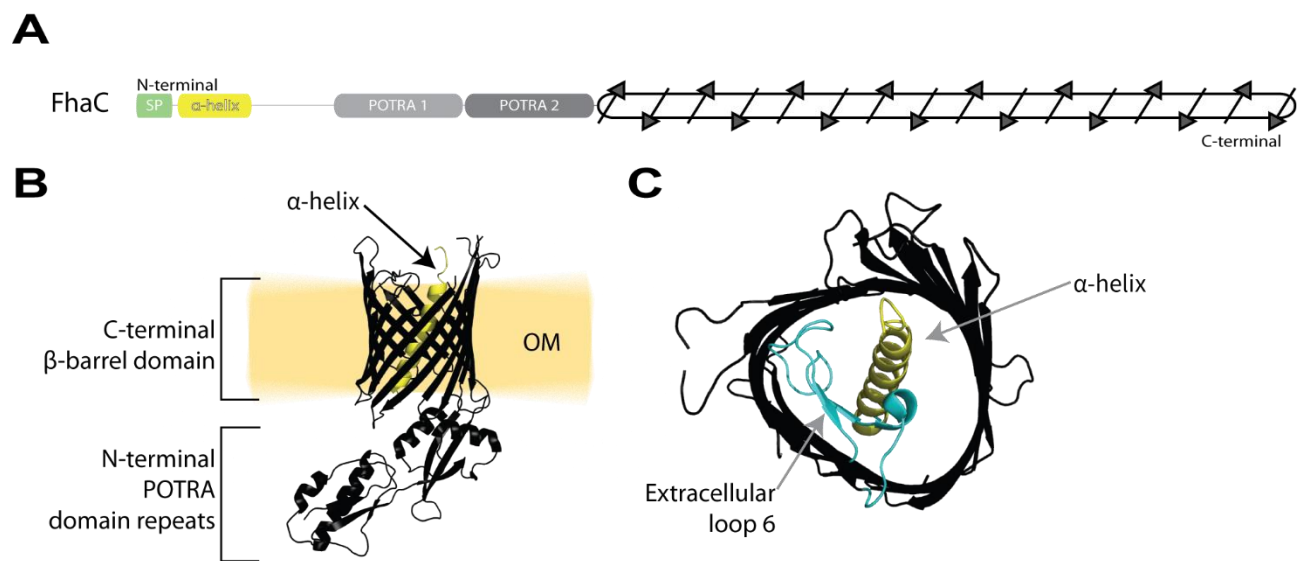
Figure 1.3.1: Architecture and domains of a TpsB. (A) Topological features of the FhaC protein based on its crystal structure (PDB: 3NJT). The signal peptide (SP) of FhaC is shown in green which is cleaved off prior to being assembled and inserted into the OM. The N-terminal domain is composed of α-helix is connected to two POTRA domain repeats by an unstructured linker. The C-terminal domain is transmembrane barrel made of 16 anti-parallel β-strands. (B) The topology of FhaC in an OM is near identical to BamA of *E. coli,* but has some alterations reflecting different functional mechanisms. (C) Extracellular loop 6 sits in the lumen of the transmembrane β-barrel forming a conserved "lid lock" structure (Maier *et al.*, 2015, Clantin *et al.*, 2007). The "lid lock" structure is thought to act as a gate, controlling the translocation of substrate molecules across the OM and is modulated by the movement of POTRA domains displacing the α-helix/extracellular loop 6 "lid lock".

Type V Secretion Systems in bacteria:

Bacterial proteins are synthesised in the cytoplasm and proteins that function on the OM cell surface or within the extracellular milieu must be translocated across IM and OM lipid bilayers to reach their destination. To efficiently and selectively translocate proteins across the bacterial membranes, bacteria have evolved numerous secretion systems. To date, nine families of secretion systems (Type I through Type IX) have been described in bacteria. The Type V Secretion System (T5SS) is widespread in Gram-negative bacteria allowing traffic of proteins across the OM usually for secretion but also in surface-associated structures (Van Ulsen *et al.*, 2014, Fan *et al.*, 2016). There are several sub-classes of the T5SS, across which functional roles include host cell adhesion,

pathogenesis and more recently contact dependent growth inhibition. The subclasses are: T5SSa (classical autotransporters), T5SSb (two-partner secretion), T5SSc (trimeric autotransporters), T5SSd (secreted phospholipases), T5SSe (inverse autotransporters) and T5SSf (two-partner inverse autotransporters). Of these, the T5SSs discussed here are because it encompasses: (i) the T5SSb with a TpsB translocase and a dedicated substrate (Jacob-Dubuisson *et al.*, 2013, Guérin *et al.*, 2017), (ii) the T5SSd with an Omp85 translocase with a translocated substrate domain (Salacha *et al.*, 2010, Casasanta *et al.*, 2017).

T5SSb: Two-partner secretion system

The two-partner secretion (TPS) pathway is capable of secreting very large substrates (~100-650 kDa), usually virulence factors, and some examples include cytolysins, hemolysins, host adhesins and proteases (St Geme *et al.*, 1993, Cotter *et al.*, 1998, Aoki, 2005, Hertle, 2005, Fan *et al.*, 2016). A minimal TPS system is composed of two distinct proteins (Figure 1.3.2 A), a secreted exoprotein and a translocon OMP, generally termed TpsA and TpsB respectively (Jacob-Dubuisson *et al.*, 2004, Newman & Stathopoulos, 2004, Mazar & Cotter, 2006). The prototypical TPS system is the secretion of filamentous haemagglutinin (FHA) (TpsA) by translocon FhaC (TpsB) of *B. pertussis*. The genes encoding TpsA and TpsB are usually found in the same operon (Jacob-Dubuisson *et al.*, 2013, Guérin *et al.*, 2017). Recent work has shown that some TpsB translocons can translocate more than one TpsA effector (Julio & Cotter, 2005) or appear to be promiscuous for other exoprotein substrates (Van Ulsen *et al.*, 2008, ur Rahman & van Ulsen, 2013).

The overall structure of TpsA proteins are thought to adopt a solenoid-like β-helical stem architecture (Figure 1.3.2 B). TpsA proteins harbour long stretches of repeats which are thought to form β-strands that organise into repeating fibrous β-helix folds (Clantin *et al.*, 2004, Kajava & Steven, 2006a, Kajava & Steven, 2006b, Alsteens *et al.*, 2013). These elongated β-helical structures are thought to be structurally important for function as it could provide, structural integrity, protease resistance and binding sites for aggregation and receptor binding (Emsley *et al.*, 1996, Girard *et al.*, 2010). However, full length structural studies pertaining to TpsA subunits have been difficult due to their large size, complicated topologies, and poor solubility. The T5SSb mechanism for protein secretion can be summarised as follows (Figure 1.3.2 C): (i) TpsA substrate proteins are initially translated in the cytoplasm and then translocated across the IM via the Sec machinery. (ii) nascent TpsA protein interacts with the cognate TpsB partner, (iii) the TpsA protein begins folding into its β-helical structure and is translocated across the OM, driven by the energetic gain that comes from protein folding, where it is secreted or remains associated to the cell surface.

T5SSd: hybrid secretion of phospholipases

The exoprotein PlpD from *Pseudomonas aeruginosa* is the prototype of T5SSd (Salacha *et al.*, 2010). PlpD represents a single translocon OMP that translocates a secreted lipolytic N-terminal phospholipase A1 (PLA1) domain across the OM. This system is described as being hybrid, as the effector PLA1 domain involves two other domains for its translocation mechanism: (a) a single POTRA domain and a (b) C-terminal Omp85-like β-barrel (Figure 1.3.2 D). Omp85 β-barrels are discussed later in this chapter. The functional role and mechanism for translocation of the proteolytic PLA1 domain is still unknown but is thought to have a similar repeating solenoid-like structure like other virulence-associated TpsA effectors. In a study by Casasanta *et al.* (2017), the authors found that the T5SSd PLA1 domain of *Fusobacterium nucleatum* binds with affinity to host phosphoinositide signalling lipids, potentially highlighting an intracellular niche during host infection. Furthermore, the authors also found that a conserved motif found between the PLA1 domain and POTRA may dictate if a PLA1 domain of a T5SSd protein is either secreted or remains membrane bound. The exact role of the single POTRA domain is unclear in T5SSd. POTRA domains are thought to act as a docking platform for client substrates and partner proteins, but as the lipolytic N-terminal secreted domain is fused to the C-terminal translocon β-barrel OMP its role could be more to do with substrate sensing or involved in overall conformational regulation for translocation.
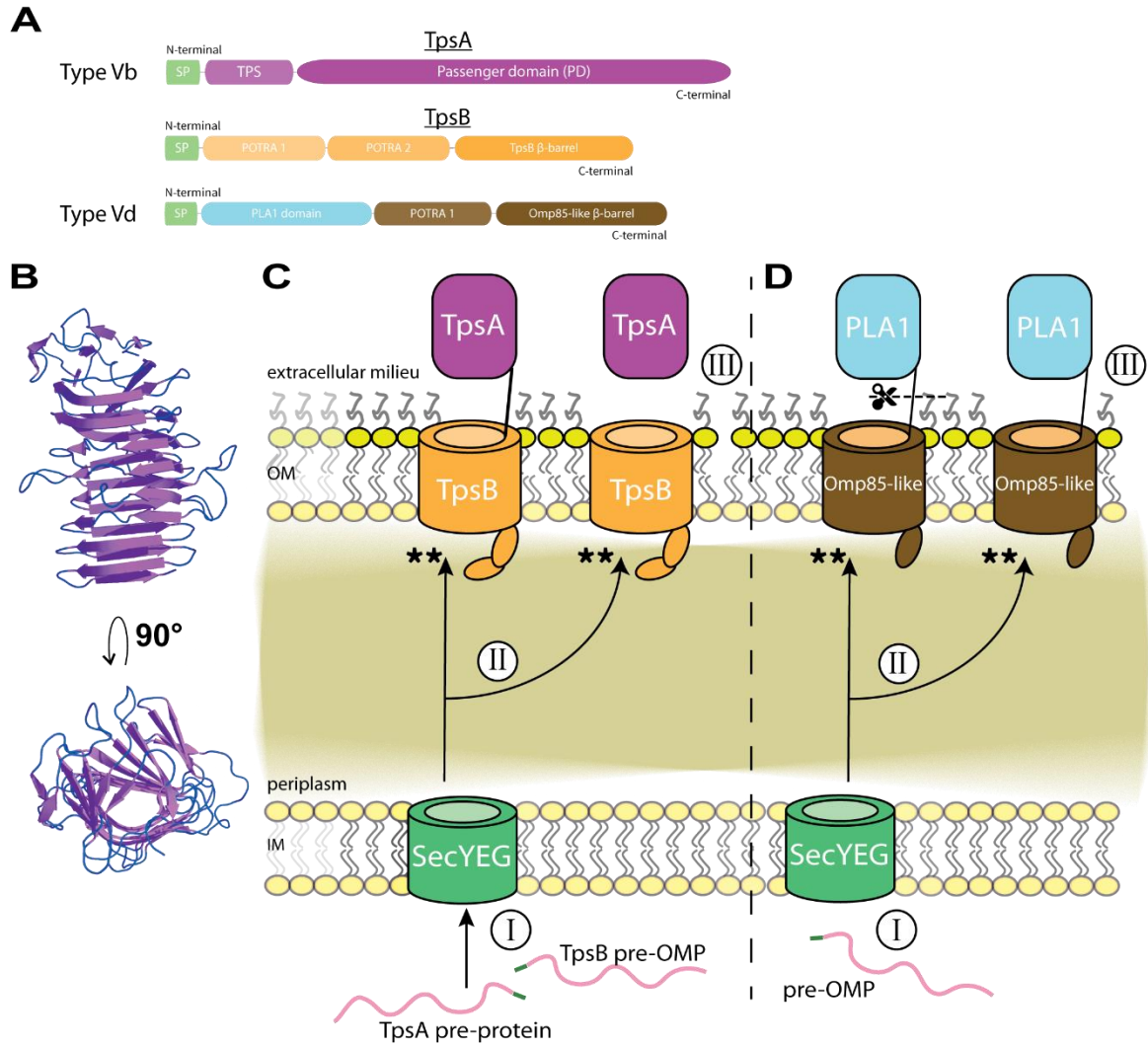
Figure 1.3.2: Comparison of T5SSb and T5SSd secretion. (A) Topological features for proteins of the T5SSb (two-partner secretion) and the T5SSd (hybrid secretion of phospholipases) secretion pathways. The signal peptides (SP) are shown in green for all proteins. T5SSb proteins are a two-partner system made of two proteins: TpsA and TpsB. TpsA proteins contain two domains, a TPS domain and a passenger domain. TpsB proteins contain two N-terminal POTRA domain repeats and a C-terminal 16-stranded β-barrel domain. T5SSd proteins are comprised of three domains, an N-terminal passenger domain phospholipase A1 (PLA1) and a single POTRA domain fused to a C-terminal 16-stranded β-barrel domain. (B) Partial crystal structure of the N-terminal domain of TpsA effector FHA of *B. pertussis* (PDB: 1RWR). TpsA proteins are predicted to adopt an elongated stalk-like structure made of β-helical folds based on repeating sequence motifs. (C) Schematic showing the T5SSb pathway. (I) The newly translated TpsA pre-protein and TpsB pre-OMP are translocated across the IM by the SecYEG translocase. (II) The TpsB pre-OMPs and TpsA pre-protein are then processed and chaperoned across the periplasm where TpsB is inserted into the OM (Figure 1.2.1), and TpsA binds to the POTRA domains of its partner TpsB. (III) The TpsA nascent protein presumably is threaded through the translocon TpsB OMP and remains cell-surface associated or secreted. (D) Schematic showing the T5SSd pathway. (I) The newly translated pre-OMP is translocated across the IM by the SecYEG translocase. (II) The pre-OMP is then processed, chaperoned across the periplasm and assembled into the OM (Figure 1.2.1). (III) The C-terminal β-barrel domain presumably acts as a pore/translocon facilitating passage of the PLA1 domain across the OM. The N-terminal PLA1 domain is then proteolytically cleaved allowing its secretion into the extracellular space or remains associated to the transmembrane domain.
**: omission of the BAM complex and the TAM catalysing the insertion and folding of the β-barrel domains discussed in the later sections.

Omp85 proteins

The Omp85 group of the superfamily are widely distributed and homologues are present in mitochondria (Sam50), chloroplasts (Toc75) and all Gram-negative bacteria phyla (Yen *et al.*, 2002, Heinz & Lithgow, 2014). Recent *in silico* analysis showed that Omp85 proteins can be classified into 8 distinct prokaryotic subfamilies, based on defining characteristics of their N-terminal extensions and/or C-terminal β-barrel domains (Heinz & Lithgow, 2014). The archetypal Omp85 protein is BamA and follows a two-domain architecture: (i) the C-terminal transmembrane β-barrel domain characteristic of the Omp85 protein superfamily, and (ii) 5 POTRA repeats that form an N-terminal domain located in the periplasm (Figure 1.3.3).

Figure 1.3.3: Domain architecture map of BamA. Topological features of the BamA protein based on its crystal structure (PDB: 5D0O). The signal peptide (SP) of BamA is shown in green and is cleaved off prior to being assembled and inserted into the OM. The N-terminal domain is composed of 5 POTRA domain repeats and the C-terminus forms a transmembrane β-barrel domain made up of 16 anti-parallel β-strands.

Each POTRA repeat consists of a conserved N'-β1-α1-α2-β2-β3-C' secondary structure, and have been observed by NMR, SAXS and crystallography under various conditions to sample a variety of conformations regulated by binding of substrates and co-factors (Knowles *et al.*, 2008, Doerner & Sousa, 2015, Zhang *et al.*, 2011, Jansen *et al.*, 2015) (Figure 1.3.4). Phylogenetic analysis of POTRA sequences has shown a remarkable level of diversity (Sánchez-Pulido *et al.*, 2003, Kim *et al.*, 2007), which could be attributed to an under-appreciated diversity in specialised functional roles or species specificity (Heinz & Lithgow, 2014, Browning *et al.*, 2015).
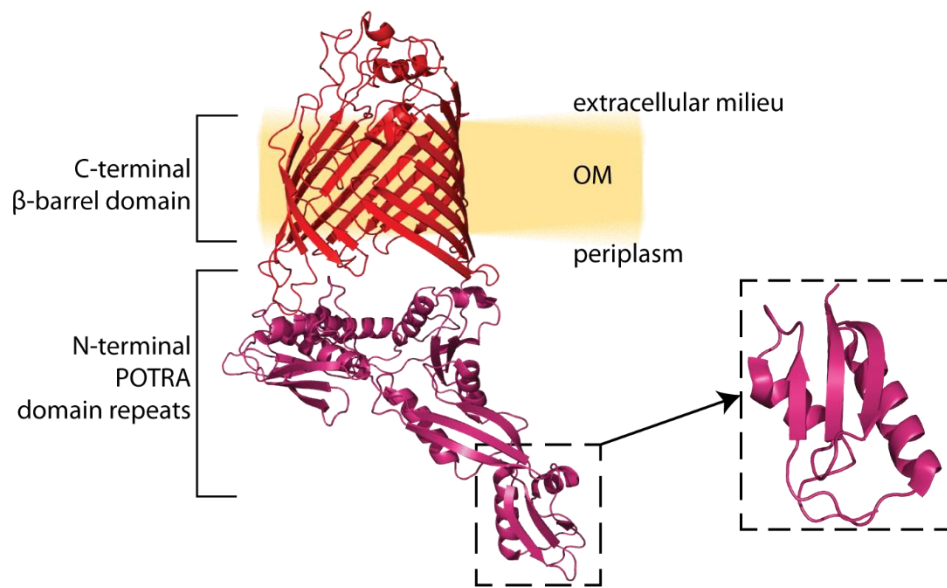
Figure 1.3.4: POTRA domains of an Omp85. Schematic showing the topology of POTRA domain in the archetypal Omp85 BamA (PDB: 5DO0). POTRA domains are a common N-terminal feature of Omp85s. They are soluble and extend from the transmembrane β-barrel domain into the periplasm. In Omp85s, the number of POTRA repeats vary between 0 and 7, but BamA proteins typically contain 5 (Heinz & Lithgow, 2014). Inset, POTRA motifs have a conserved β1-α1-α2-β2-β3 motif and are connected by short linkers allowing flexible conformations. In most Omp85s, the most proximal POTRA domain to the β-barrel is critical, as deletion impact the catalytic activity of the Omp85 resulting in severe OMP assembly defects and consequently compromised OM integrity (Kim *et al.*, 2007).

The C-terminal β-barrel domain of BamA consists of 16 anti-parallel β-strands with both the N-terminal end and C-terminus of the protein situated at the periplasm. Typical of other OMPs, the β-strands are connected by a series of short periplasmic turns and longer extracellular loops (Figure 1.3.5). In available crystal structures of BamA, the barrel lumen is occluded by the extracellular loops which form a "capping dome" (Noinaj *et al.*, 2013, Ni *et al.*, 2014a). Of particular note, extracellular loop 6 (L6) that forms part of the "capping dome" is also buried into the lumen and sits nestled against the β-barrel wall (Noinaj *et al.*, 2013, Ni *et al.*, 2014a). The tip of L6 contains a characteristic VRGY residue motif that is conserved throughout the Omp85 superfamily, where mutation of this motif strongly reduces the catalytic activity of the protein (Gentle *et al.*, 2005, Moslavac *et al.*, 2005, Delattre *et al.*, 2010). BamA is the only integral membrane protein of the BAM complex in *E. coli*; the other four subunits are lipoproteins.
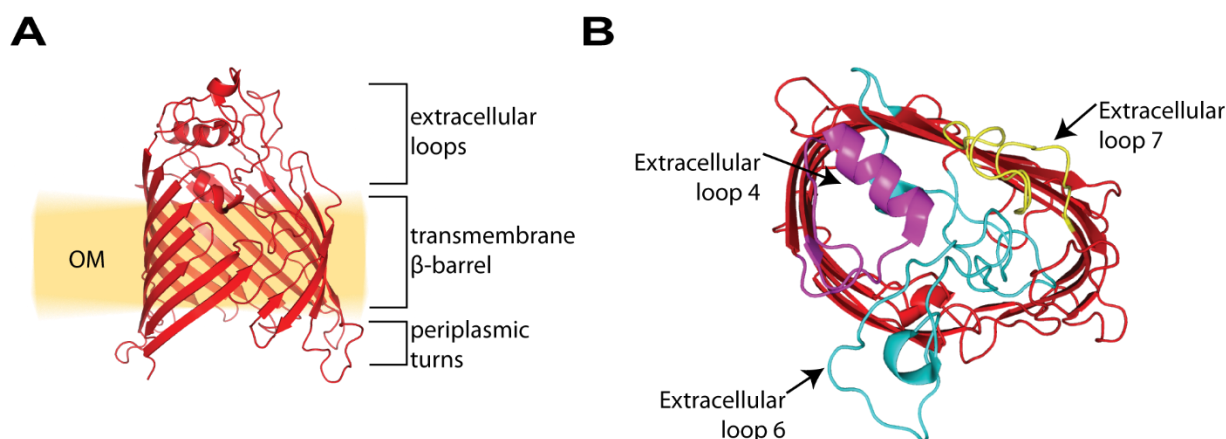
Figure 1.3.5: β-barrel domain of an Omp85. Shown here is a closer look at the C-terminal β-barrel domain of *E. coli* BamA (PDB: 5DO0).(A) The C-terminal β-barrel domain of BamA is typical of other OMPs and contains 16 stranded β-barrel connected by short periplasmic loops and longer extracellular loops.(B) is a top down view of the β-barrel highlighting major extracellular loops 4 (purple), 6 (cyan) and 7 (yellow), which contribute to the "capping dome". The capping dome is a network of electrostatic interactions between BamA extracellular loops that is believed to occlude the lumen preventing diffusion of substrates into and out of the barrel lumen.

1.4 The BAM Complex

The β-barrel assembly machinery (BAM) complex serves to catalyse the proper folding and insertion of OMPs into the OM on a biologically relevant timescale (Plummer & Fleming, 2016, Hussain & Bernstein, 2018). The BAM complex consists of multiple components and can vary in species (Anwari *et al.*, 2012, Webb *et al.*, 2012a). In *E. coli*, the BAM complex is comprised of the five proteins: BamA, BamB, BamC, BamD, and BamE, where alphabetical letters were assigned on the basis of decreasing molecular weight (Table 1.4.1).

First identified in the early 2000s as Omp85 in *Neisseria* and *Synechocystis*, and as YaeT in *E. coli*, based on high sequence conservation to chloroplast homologue Toc75 (Bölter *et al.*, 1998, Voulhoux *et al.*, 2003). A subsequent genetic screen in *E. coli* then identified the lipoprotein YfgL (now BamB) as a suppressor able to overcome OM permeabilities defects, and thereafter as a partner to BamA through protein-protein interaction studies (Charlson *et al.*, 2006, Sklar *et al.*, 2007b, Sklar *et al.*, 2007a, Wu *et al.*, 2005). The other partner lipoproteins: BamC (NlpB), BamD (YfiO) and BamE (SmpA), were identified through protein-protein interaction studies using BamA and BamB as bait (Sklar *et al.*, 2007a, Wu *et al.*, 2005).

Table 1.4.1: Components of the *E. coli* BAM complex. In *E. coli*, the BAM complex is a hetero-pentameric complex localised to the OM that catalyses the folding and insertion of OMPs into the OM. BamA is the only OMP component of the complex and is the largest, whilst BamB through BamE are lipoproteins named in order of molecular weight from highest to lowest. Single knockout studies of the BAM complex have revealed that BamA and BamD are the only essential components in *E. coli* (Malinverni *et al.*, 2006). However, single deletion of the other lipoproteins negatively impacts the function of the BAM complex leading to OM integrity defects (Wu *et al.*, 2005, Sklar *et al.*, 2007a, Ruiz *et al.*, 2005).

Protein, previous alias(es)	Molecular weight OMP or lipoprotein	Taxonomic distribution**	Knockout phenotype
BamA previously YaeT, Omp85	88.4 kDa, OMP	All Proteobacteria	Essential: depletion results in decreased OMP levels and accumulation of unfolded OMPs
BamB previously YfgL	39.9 kDa, lipoprotein	α, β, and γ Proteobacteria	Non-essential: compromised OM integrity, defects in OMP assembly
BamC previously NlpB	34.4 kDa, lipoprotein	β and γ Proteobacteria	Non-essential: mildly compromised OM integrity, mild defects in OMP assembly
BamD previously YfiO	25.8 kDa, lipoprotein	All Proteobacteria	Essential: depletion results in decreased OMP levels and accumulation of unfolded OMPs
BamE previously SmpA	10.4 kDa, lipoprotein	α, β, and γ Proteobacteria	Non-essential: mildly compromised OM integrity, mild defects in OMP assembly

** Based on studies and reviews from: (Anwari *et al.*, 2012, Webb *et al.*, 2012a, Heinz & Lithgow, 2014)

Of these five components, only BamA and BamD are essential for cell viability (Wu *et al.*, 2005, Malinverni *et al.*, 2006). The deletions of BamB, BamC or BamE are non-lethal, but have been reported to negatively impact OM integrity and efficacy of OMP assembly due to an overall decreased catalytic activity (Onufryk *et al.*, 2005, Charlson *et al.*, 2006, Sklar *et al.*, 2007a, Hagan *et al.*, 2010). Stoichiometric estimates of purified Bam complex by mass spectrometry are suggested at 1:1:1:1:1; consistent with the structures of the BAM complex by crystallography and single-particle cryo-electron microscopy (Hagan *et al.*, 2010, Gu *et al.*, 2016, Bakelar *et al.*, 2016, Han *et al.*, 2016, Iadanza *et al.*, 2016). The BAM complex is modular in architecture, with detergent-solubilisation of the native complex from the outer membrane of *E. coli* yielding a BAM$_{AB}$ module and a BAM$_{CDE}$ module (Sklar *et al.*, 2007a, Kim *et al.*, 2007, Vuong *et al.*, 2008, Hagan *et al.*, 2010, Noinaj *et al.*, 2011, Webb *et al.*, 2012b). Several structural studies have shown that select interactions between the POTRA repeats of BamA with BAM lipoproteins stabilises the BAM complex architecture by providing specific attachment/anchor points for the BAM$_{AB}$ and BAM$_{CDE}$ modules of the overall molecular machine (Bakelar *et al.*, 2016, Gu *et al.*, 2016). The POTRA to lipoprotein interactions are also hypothesised as: (i) a regulatory mechanism for dynamic conformational changes of the BamA monomer of the BAM complex (Bakelar *et al.*, 2016, Gu *et al.*, 2016), (ii) recognition and binding of nascent OMPs, (iii) a docking point for periplasmic chaperones, and (iv) direct interactions with the OM lipid membrane (Knowles *et al.*, 2008, Bennion *et al.*, 2010, Fleming *et al.*, 2016, Lee *et al.*, 2018). Despite the bevy of structural analysis on the BAM complex and its individual components

(Figure 1.4.2), the mechanism for OMP assembly and insertion remains largely unknown. In this section I will briefly discuss the five components of the BAM complex from A through to E. It should however be noted, that while these features in the architecture of the BAM complex in *E. coli* may be found to be similar in *K. pneumoniae* and other closely related bacteria, there is known diversity in the subunit composition of the BAM complex outside the Gamma-Proteobacteria. While BamA and BamD appear to be ubiquitously distributed in all bacteria with an outer membrane, BamB, BamC or BamE evolved much more recently (Webb *et al.*, 2012a, Anwari *et al.*, 2012).
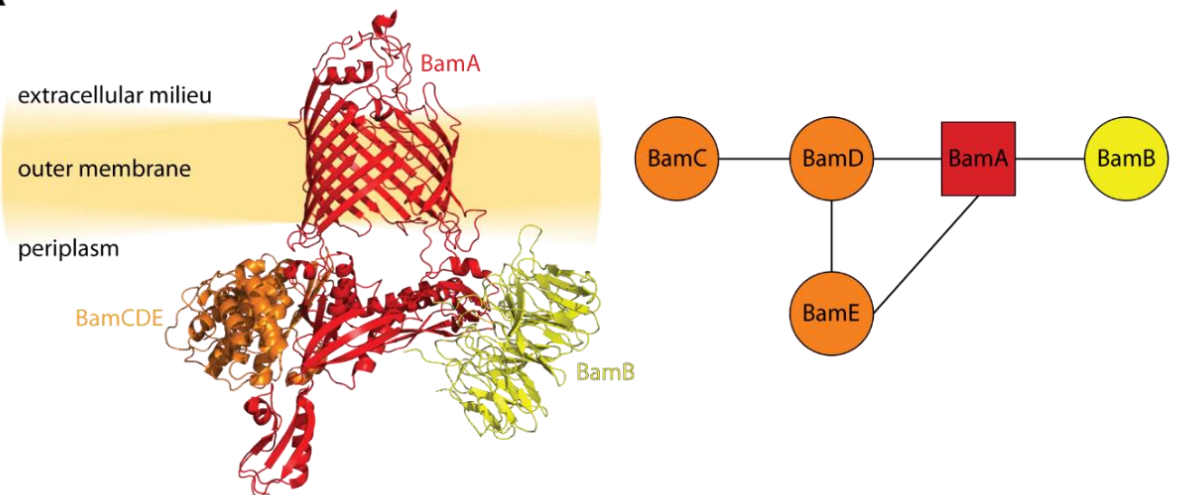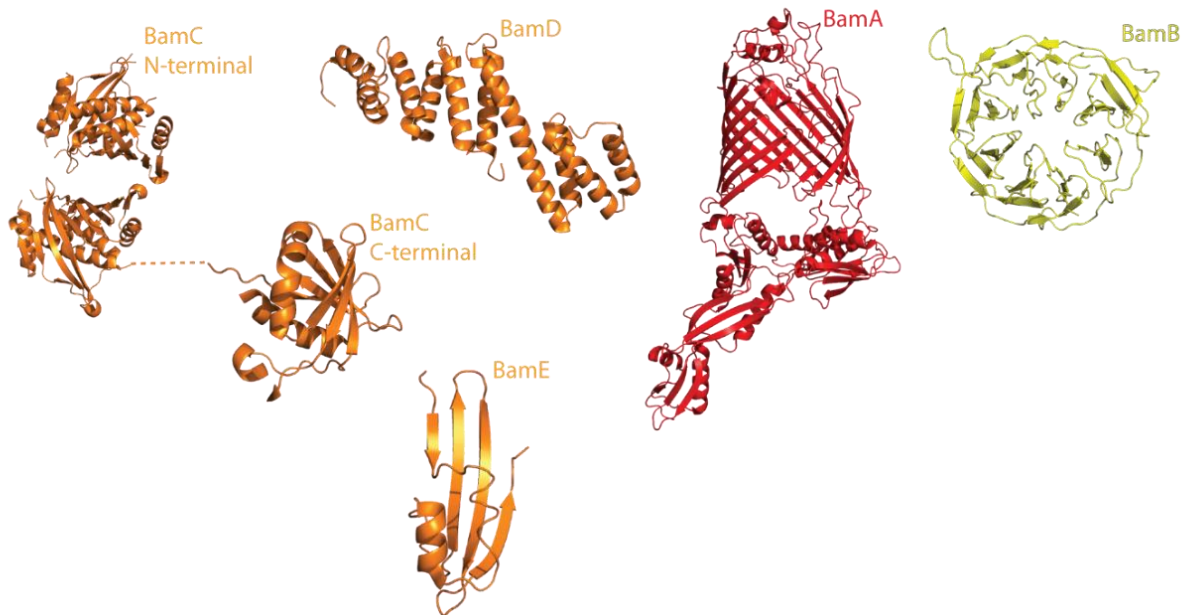
Figure 1.4.2: Crystal structure of the *E. coli* BAM complex, interaction map and individual components. (A) Left, side view of the *E. coli* BAM complex (PDB: 5D0O) with all five components (BamA through BamE) show in relation to the Gram-negative OM layer. Right, interaction map showing the binding between BAM subunits observed through structural studies and experimental studies. (B) Individual BAM subunits from the previously mentioned crystal structure (PDB: 5D0O), except BamC (PDB: 2YH6 and 3SNS for its N- and C-terminal globular domains, respectively), as it was only partly resolved.

BamA

The central component of the BAM complex is the β-barrel OMP BamA (Figure 1.4.2), with transcriptional control coordinated within the σ^E regulon (Dartigalongue *et al.*, 2001). The reported association of BamA with its partners has been shown by monitoring the behaviour of BamA through size-exclusion chromatography, BN-PAGE, and co-immunoprecipitation analyses (Wu *et al.*, 2005, Malinverni *et al.*, 2006, Sklar *et al.*, 2007a, Hagan *et al.*, 2010, Webb *et al.*, 2012b). The earliest experimental evidence that BamA plays an essential role in OMP biogenesis was through a gene depletion study that demonstrated the periplasmic accumulation of unfolded OMPs and also showed an *in vivo* interaction occurs between unfolded OMPs and BamA (Voulhoux *et al.*, 2003). In later *in vitro* studies, BamA in isolation has been shown to accelerate the assembly of nascent OMPs into the OM (Plummer & Fleming, 2015), and that addition of the cognate lipoproteins increases the activity of BamA in these assays (Hagan *et al.*, 2010, Gessmann *et al.*, 2014). This would suggest BamA is performing enzymatically: by reducing the activation barrier for nascent OMPs to spontaneously fold and insert into the OM, whilst the BAM lipoproteins are involved in more supporting roles potentially regulating BamA conformation or providing a binding platform for chaperones and nascent OMPs (Ricci *et al.*, 2012, Bakelar *et al.*, 2017, Bakelar *et al.*, 2016).

The C-terminal β-barrel domain of BamA has three noticeable features which may contribute to its activity in catalysing OMP assembly. Firstly, the β-barrel domain BamA is comprised of 16 anti-parallel β-strands and crystal structures of BamA have revealed that the inter-strand extracellular loops of BamA are involved both in the stabilisation of the β-barrel domain (Browning *et al.*, 2015, Thoma *et al.*, 2018), and in formation of a "capping dome" that occludes the β-barrel lumen, thereby preventing substrates to exit longitudinally (Noinaj *et al.*, 2014) (See Figure 1.3.5).

Secondly, whereas the hydrogen bonds between the first and last β-strands of a transmembrane β-barrel are usually observed to stabilise the overall closed barrel structure, molecular dynamics simulations and crystal structures for BamA have captured conformations that show an open junction where the first (β1) and last (β16) β-strands are bent away from each other resulting in an "unzipped" or "open" conformation, where there is a reduced hydrogen bonding network (Bakelar *et al.*, 2016, Gu *et al.*, 2016, Han *et al.*, 2016, Iadanza *et al.*, 2016, Doerner & Sousa, 2017) (Figure 1.4.3 A). These "unzipped" or "open" conformations have been hypothesised to contribute to a lateral gating mechanism where this suspected conformational cycling enables alternation between an open and closed state for controlled lateral exit of OMP substrates (Rigel *et al.*, 2013) (Figure 1.4.3 B). Other studies have further investigated this β-barrel seam of BamA by introducing engineered disulphide

cross-links locking this region resulting in a lethal *E. coli* phenotype (Noinaj et al., 2014). In a study by Doerner and Sousa (2017), the authors were able to show that despite β-barrel locking, the BAM complex was able to accelerate folding and insertion of OMP substrate OmpX *in vitro*. However, *in vivo* experiments demonstrated that barrel locking resulted in nonviable cells and postulated the dynamic nature of the BamA β-barrel seam is important for destabilisation of the OM.
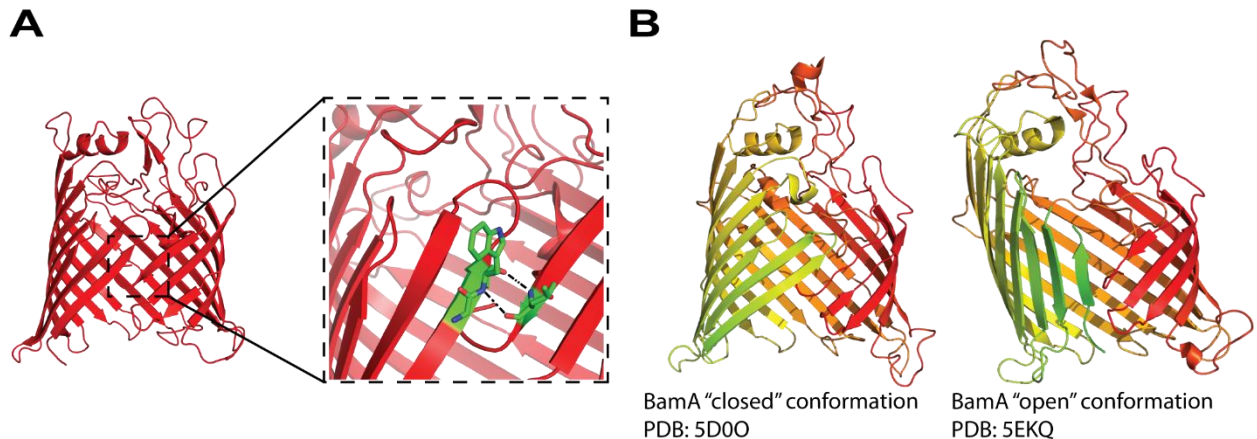
Figure 1.4.3: The β-barrel seam of BamA and "conformational cycling". (A) Weak hydrogen bonding (dashed black lines) seen between strands β1 and β16 (highlighted in green) forms the basis of the "lateral gate" feature of BamA. The "lateral gate" is hypothesised to permit controlled lateral release of OMP substrates into the OM (PDB ID: 4K3B). (B) Two crystal structures of the BamA monomer demonstrating different conformations described as "closed" or "open" states. BamA is thought to cycle through these conformations, which changes upon binding of partner BAM lipoproteins, to control the lateral release of OMP substrates into the OM.

Finally, the BamA β-barrel domain also shows some distinct differences in terms of β-strand lengths that form the β-barrel domain. The first (β1) and last (β16) β-strands that form the lateral gate are distinctly shorter than the other 14 β-strands (Noinaj *et al.*, 2013). These shorter β-strands place the residues which constitute aromatic girdles of the OMP closer, potentially generating a lopsided OM where the hydrophobic thickness is suspected to be much smaller when compared to the directly opposite side of the BamA β-barrel (Figure 1.4.4). This potentially disrupted OM layer is suspected to provide a more energetically favourable environment for folding and insertion of OMPs (Plummer & Fleming, 2016).
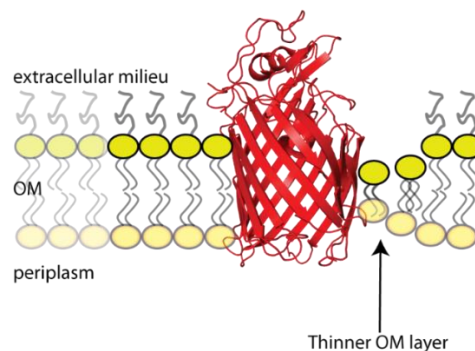
Figure 1.4.4: Proposed localised membrane destabilisation caused by BamA. The hydrophobic thickness and density of the two sides of BamA is thought to be different due to the shorter β-strand lengths at one end of its β-barrel domain. This thinned membrane is believed to lower the kinetic energy required for folding and insertion of OMPs.

BamB

Expression of *bamB* is controlled by the σ^E regulon (Dartigalongue *et al.*, 2001, Onufryk *et al.*, 2005). BamB is a WD40 protein and adopts an eight-bladed β-propeller structure (see Figure 1.4.2 B), with four β-strands per blade. WD40 proteins often play a role in ligand-binding or protein-protein interactions, and both of these ideas have been proposed for BamB function. Structural and biochemical studies have shown that BamB directly interacts with the POTRA domain (specifically, POTRA 3) of BamA (Wu *et al.*, 2005, Vuong *et al.*, 2008, Heuck *et al.*, 2011) (see Figure 1.4.2 A). In a recent study using super-resolution imaging and in situ cross-linking, the authors observed that the function of BamB is to mediate interactions between neighbouring BAM complexes, to form focused OMP assembly precincts across the bacterial cell surface (Gunasinghe *et al.*, 2018).

BamC

There is substantial structural information on the various domains of BamC (Knowles *et al.*, 2009, Albrecht & Zeth, 2010, Kim *et al.*, 2011) (see Figure 1.4.2 B), but a full-length structure in the context of the BAM complex is missing. Recent crystal structures and single-particle cryo-EM structures of the BAM complex required to model in residues of the BamC sequence due to inconsistencies in the corresponding density (Gu *et al.*, 2016, Iadanza *et al.*, 2016). An intriguing aspect of BamC, one that probably holds the key to why it is so flexible in the detergent-solubilized BAM complexes, comes from studies that show BamC is exposed on the surface of the bacterial cell (Webb *et al.*, 2012b, Gunasinghe *et al.*, 2018). As a transmembrane protein, removal of the membrane by detergent may destabilise the interactions that BamC makes with the other subunits of the BAM complex. Unlike *bamB*, deletion of *bamC* does not drastically affect OMP levels but still results in noticeable OM permeability defects (Onufryk *et al.*, 2005, Wu *et al.*, 2005). However, deletions of *bamC* in combination of *bamB* or *bamE* results in drastically lower levels of BamD (Malinverni *et al.*, 2006, Wu *et al.*, 2005, Rigel *et al.*, 2012). Like *bamB* and *bamA*, transcriptional control of *bamC* is coordinated by the σ^E regulon (Onufryk *et al.*, 2005, Dartigalongue *et al.*, 2001).

BamD

Transcriptional control of *bamD* is regulated by the σ^E regulon (Dartigalongue *et al.*, 2001, Onufryk *et al.*, 2005). Sequence analysis and structural studies showed that BamD is comprised of five tetratricopeptide (TPR) repeats (see Figure 1.4.2 B). In *E. coli*, the TPR repeats are arranged into two domains connected by an extended α-helix: 3 TPR repeats in the N terminus and 2 TPR repeats in the C-terminus (Dong *et al.*, 2012). It has been suggested that the N-terminal TPR domains might interact with incoming OMP substrates (Hagan *et al.*, 2015, Hagan *et al.*, 2013), while the C-terminal

domain interacts with BamA, BamC and BamE (Gu *et al.*, 2016, Han *et al.*, 2016). In addition to the structural data, genetic evidence and affinity purification experiments confirm that BamD is responsible for the stable interaction of BamA with the BAM$_{CDE}$ sub-complex (Malinverni *et al.*, 2006, Sklar *et al.*, 2007a). As with BamA, depletion of BamD also results in lower steady-state levels of OMPs as judged by immunoblots of bacterial membrane extracts (Malinverni *et al.*, 2006, Wu *et al.*, 2005).

BamE

BamE is a monomeric lipoprotein component that consists of two N-terminal α-helices and a C-terminal β-sheet containingthree3 β-strands (see Figure 1.4.2 B) (Endo *et al.*, 2011). BamE is thought to play an important role in stabilising the interaction of BamD to BamA (Sklar *et al.*, 2007a, Rigel *et al.*, 2012). Additionally, BamE is thought to play a regulatory role in BamA monomer conformation as its deletion results in proteolytic lability of BamA (Rigel *et al.*, 2012). Physiologically, *bamE* is regulated by the σ^{E} regulon (Rezuchova *et al.*, 2003) and while shut-down of *bamE* does not drastically affect OMP levels, steady state levels of the stress-sensing RcsF/OMP complexes are reduced in a *bamE* deletion strain (Konovalova *et al.*, 2016).

Proposed mechanisms by which the BAM complex mediates OMP assembly

A holistic understanding of how the BAM complex functions in OMP biogenesis is still unknown. Several models have proposed how OMP assembly and insertion into the OM membrane occurs and, in this section, I will discuss two prevailing mechanisms: (i) the "assisted model", and (ii) the "budding model". The BAM complex 'assisted model' posits that nascent OMP folding occurs substantially in the periplasm, with integration of the folded barrel primed via membrane defects created by BamA. The 'budding model' proposes a direct role of BamA in generating protein-protein contacts with its substrate in order that it builds the substrate β-barrel that ultimately buds away from the BAM complex. As discussed below, both models have experimental evidence in support of their respective claims. The two models presented may actually represent two sides of a puzzle where certain OMP topologies are favoured by one mode of action more so than the other.

In the BAM complex assisted model, OMP folding begins in the periplasm and is thereby primed for integration into the OM, with the exposure of the hydrophobic surface necessitating integration into the membrane. This integration step would be facilitated at membrane sites disturbed by the action of the lateral gate in BamA (Figure 1.4.5). The spontaneity of β-barrel proteins to readily fold and insert into artificial lipid bilayers or detergent environments has been measured and is a multistep process (Fleming, 2015). Briefly summarised as follows: A nascent OMP sequestered by SurA or Skp is escorted to the BAM complex in a "partially folded" state. The degree of folding is not currently known but is hypothesised to be substantial. The delivered OMP then initiates penetration into the membrane for eventual insertion as a folded cylinder. Consistent with this hypothesis were the observations of an introduced membrane defect found in proteoliposomes containing BamA (Sinnige *et al.*, 2014). A concern with this model is extrapolation from the data gathered with purified components and non-model/artificial membranes, and the extent to which this remains reasonable in an *in vivo* setting, as native membranes may not be as penetrable due to other structural features such as LPS and other OMPs. Recent studies making use of reconstituted membrane extracts, replete with LPS and other protein components to track partially folded OMP intermediates (LptD; Chimalakonda *et al.* (2011), Lee *et al.* (2016) and EspP; Ieva *et al.* (2011), Pavlova *et al.* (2013)) even before OM insertion lend critical support to the BAM complex assisted model.
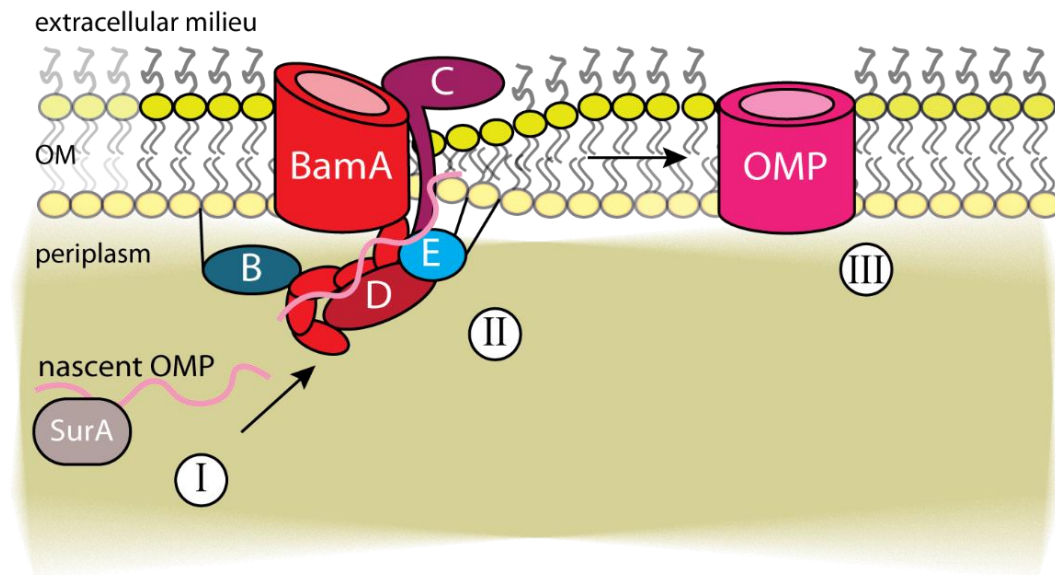
Figure 1.4.5: The BAM complex assisted model. This model proposes that OMP folding is driven by intrinsic thermodynamic and sequence properties of the nascent OMP and the BAM complex only supports thinning of the membrane thereby allowing nascent OMPs to assemble and insert utilising less energy. (I) Nascent OMPs are chaperoned across the periplasm and targeted to the BAM complex. (II) The BAM complex is thought to catalyse local membrane defects by virtue of select β-strands of shorter length coupled with additional disruption due to the dynamic nature of the β-barrel seam/lateral gate. The BAM complex POTRA domains potentially guides the nascent OMP to the destabilised area. (III) The nascent OMP can now assemble and insert into the OM as it is now in an energetically favourable environment.

The BAM complex budding model theorises that BamA opens laterally at its seam, allowing the uncoupled β-strands to template nascent OMPs to form their β-strand structure progressively, into a β-sheet that eventually buds off the BAM complex and into the OM, as a fully-folded and membrane-integrated OMP (Figure 1.4.6). The model was inspired by observations in several structural studies of BamA that revealed hydrogen bonding between the first (β1) and last (β16) β-strands is relatively weak (Noinaj *et al.*, 2014, Noinaj *et al.*, 2013). This weakness is also evident from a comparatively low thermodynamic stability for BamA with a melting temperature (T_m) determined to be at the physiologically relevant 37 °C (Burgess *et al.*, 2008). It is therefore hypothesised that this unorthodox opening enables nascent OMPs to anneal to the exposed β-strands of the BamA barrel, thereby forming their own β-strands in a process called β-augmentation (Heuck *et al.*, 2011). New β-strands would be added in a sequential manner, using the previously made strand as a new template, thereby enlarging the BamA-OMP hybrid complex. The substrate OMP being sequentially folded would eventually "bud off" and exit adjacently from the lateral gate into the OM, as its first and last β-strand presumably come into contact thus forming a completely complemented β-barrel domain.
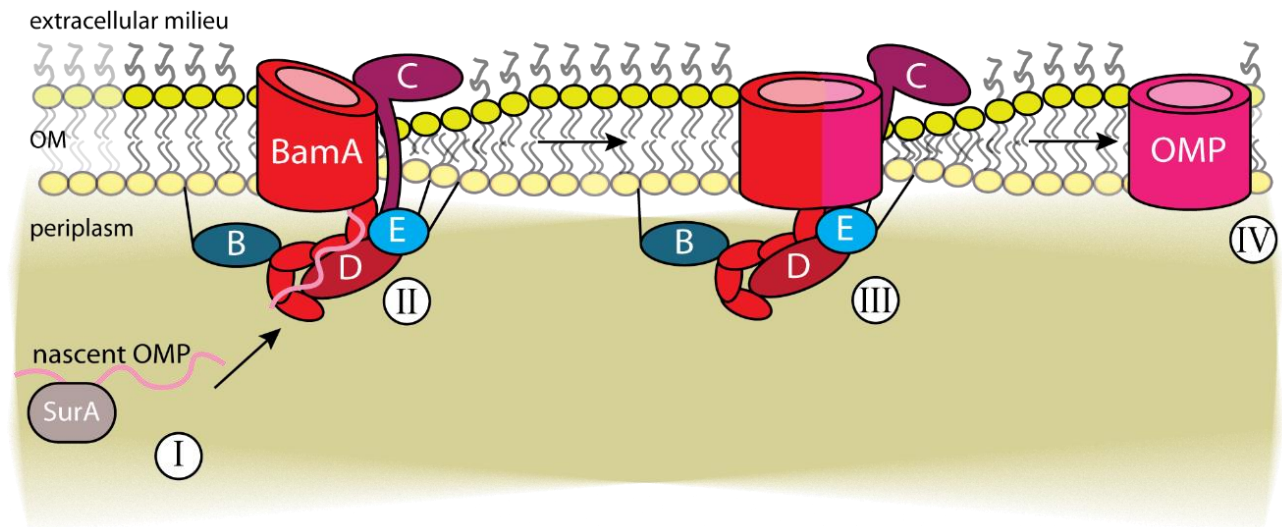
Figure 1.4.6: The BAM complex budding model. This model proposes that nascent OMP substrates BamA acts as a structural template for nascent OMPs and laterally inserts these substrates into the OM. (I) Nascent OMPs are chaperoned across the periplasm and targeted to the BAM complex. (II) The nascent OMP threads through the lumen of the β-barrel domain and forms a hybrid BamA-OMP barrel. (III) The first strand of the nascent OMP uses one of the exposed strands of the β-strands as a template and forms β-strands in a sequential manner. This mechanism is sometimes referred to as β-augmentation as the barrel domain enlarges over time for assembly (Heuck *et al.*, 2011). A similar mechanism has been reported experimentally for the analogous SAM complex in mitochondria (Höhr *et al.*, 2018). (IV) Upon completion of all β-strands of the OMP architecture, the assembled OMP is then thought to "bud" off laterally from BamA to insert into the OM.

In many Gram-negative bacteria BamA-related sequences exist and shown to function in OMP biogenesis. In a study by Selkrig *et al.* (2012), the authors discovered that an Omp85 protein called TamA (previously YtfM) is a component of the <u>T</u>ranslocation and <u>A</u>ssembly <u>M</u>odule (TAM) of the β-barrel assembly machinery. The TAM spans the periplasm, being composed of TamA in the outer membrane and TamB (previously YtfN) which is anchored in the IM (Selkrig *et al.*, 2012, Selkrig *et al.*, 2015, Shen *et al.*, 2014). The TAM plays a role in the assembly and insertion of at least some virulence associated OMPs (Stubenrauch *et al.*, 2016, Heinz *et al.*, 2016) (Figure 1.4.7). The structure of TamA is that of a characteristic Omp85 protein with a 16-stranded β-barrel and N-terminal POTRA repeats which in this case sums to three which are located in the periplasm (Selkrig *et al.*, 2012, Selkrig *et al.*, 2015, Gruss *et al.*, 2013). SAXS and NMR analyses showed that the POTRA domain of TamA is structured as a rigid body (Selkrig *et al.*, 2015), and the latter domains have been shown to function as a lever that pushes against TamB in response to the presence of a nascent OMP (Shen *et al.*, 2014). TamA is presumed to have a a lateral gate, which can bind nascent OMP segments through the gate (Bamert *et al.*, 2017). Despite this substantial knowledge on the mechanism by which the TAM functions as a module, it remains far from clear how the TAM complex cooperates with the BAM complex to drive the folding and insertion of substrates like FimD in which they both participate (Stubenrauch *et al.*, 2016).
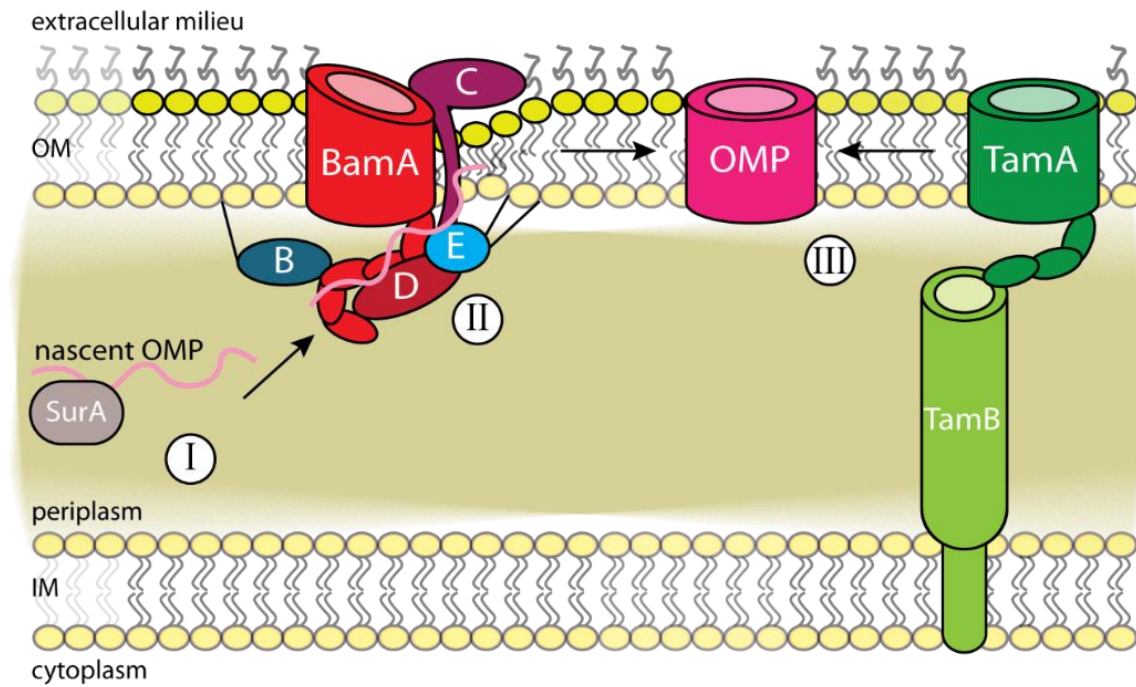
Figure 1.4.7: A further module of the β-barrel assembly machinery in Gram-negative bacteria. The TAM is comprised of two components, TamA and TamB. TamA is an Omp85 member with a similar architecture to BamA. Its N-terminal domain is made of three POTRA repeat domains and binds to IM localised partner protein TamB. Studies have shown that TAM complex works cooperatively with the BAM complex to efficiently assemble select OMP substrates (Stubenrauch *et al.*, 2016). The exact mechanism of how these two complexes cooperate for efficient assembly of client substrates is unknown. (I) Nascent OMPs are chaperoned across the periplasm and targeted to the BAM complex. (II) The nascent OMP is prepared for folding and insertion through the BAM complex, which could require the aid of the TAM complex (III) The OMP assembles and inserts into the OM.

1.5 Studies on outer membrane function and biogenesis in *Klebsiella pneumoniae* an important pathogen of humans

Klebsiella pneumoniae is a Gram-negative opportunistic pathogen and is the causative agent of human infections ranging from pneumonia, to septicaemia, to urinary tract infections and hepatic abscesses (Podschun & Ullmann, 1998, Bagley, 1985, Dhingra, 2008, Siu *et al.*, 2012). The ecological distribution of *K. pneumoniae* extends beyond hospital environments and is found as a commensal gastrointestinal flora of humans but also frequently colonises: other mammals (agricultural animals), plants, birds, fish and invertebrates (Brisse & Van Duijkeren, 2005, Bowring *et al.*, 2017, He *et al.*, 2017, Davis & Price, 2016, Stenkat *et al.*, 2014, Brahmi *et al.*, 2018, Parthasarathi *et al.*, 2007). Moreover, ecological surveys have identified the ubiquitous distribution of *K. pneumoniae* isolates in diverse environments particularly soil and water bodies (Podschun & Ullmann, 1998, Singh *et al.*, 2017b), where it often resorts to biofilm modes of growth on abiotic or biotic surfaces. This growth strategy is also used in the context of human infection. Biofilms of *K. pneumoniae* have been reported on in-dwelling catheters as well as other hospital equipment, and in the context of liver and kidney tissues (Stahlhut *et al.*, 2012, Snitkin *et al.*, 2012, Safdar & Maki, 2005, Tambyah *et al.*, 2002). Treatment of *K. pneumoniae* infections is becoming more challenging due to the global emergence of multi-drug resistant and hypervirulent pathotypes (Holt *et al.*, 2015a, Shon *et al.*, 2013).

The wide ecological range of *K. pneumoniae* has been suggested to be driven by the highly diverse genome compositions among species and its intrinsic propensity for horizontal gene transfer events (Holt *et al.*, 2015a, Blin *et al.*, 2017, Comandatore *et al.*, 2018). *K. pneumoniae* are highly diverse in respect to genome composition and have been reported to harbour ~2000 core genes that are shared among virtually all strains, and up to ~3500 accessory genes that may be present or absent (Holt *et al.*, 2015a). The disproportionate ratio of 'accessory genome' to 'core genome' genes are factors thought to strongly influence the metabolic capacities and success of *K. pneumoniae* in a wide range of environmental niches (Bohlin *et al.*, 2017, Holt *et al.*, 2015a).

K. pneumoniae is a member of the family *Enterobacteriaceae*, which includes other significant pathogens such as *E. coli, Yersinia* species, *Salmonella* species, and *Shigella* species. In a previous study by Hsieh *et al.* (2016), the authors demonstrated that deletion of *bamB* from the core genome of *K. pneumoniae* ultimately resulted in reduced adherence and host cell invasion, increased susceptibility to antibiotics and host neutrophil phagocytosis. This study reinforces the fact that the BAM complex is essential for the assembly and insertion of several pertinent virulence factors. The starting point to the research described in this thesis was the observation that, in *K. pneumoniae*, the

core genome contains four members of the Omp85 family of proteins: BamA and TamA, but also two additional Omp85 proteins: an Omp85 without POTRA domains (subfamily: noNterm, example locus identifier: BN49_0007), and a seemingly paralogous BamA gene duplicate (subfamily: BamA, example locus identifier: BN49_4981). These additional *K. pneumoniae* Omp85s proteins BN49_0007 and BN49_4981, will be referred to as BamL and BamK hereafter. Addressing the question of why *K. pneumoniae* has conserved four Omp85 proteins became the foundation of my work (Figure 1.5.1). The BAM complex and the TAM are known modules of the β-barrel assembly machinery; therefore, a conservative hypothesis for the function of these novel *K. pneumoniae* Omp85 would be that they also function in an aspect of OMP biogenesis. None the less, why *K. pneumoniae* would require additional modules beyond those found in other species of Enterobacteriaceae remains puzzling. Could these Omp85 proteins represent functionally redundant gene duplications providing specificity towards certain substrates? Or part of a transcriptionally-controlled program of gene expression that no longer included *bamA* and *tamA*?

I reasoned that investigating the functional roles of these novel Omp85 proteins would provide fundamental knowledge that could enable better understanding of OM biogenesis in *K. pneumoniae.* Previous efforts have aimed to utilise BamA/Omp85 proteins for vaccines and antibacterial therapies as they are highly conserved among species and usually play essential functions linked to OMP biogenesis (Wedege *et al.*, 2013, Singh *et al.*, 2017a, Vij *et al.*, 2018, Storek *et al.*, 2018), and so understanding the these further Omp85s might provide developmental insights for targeted therapeutics.
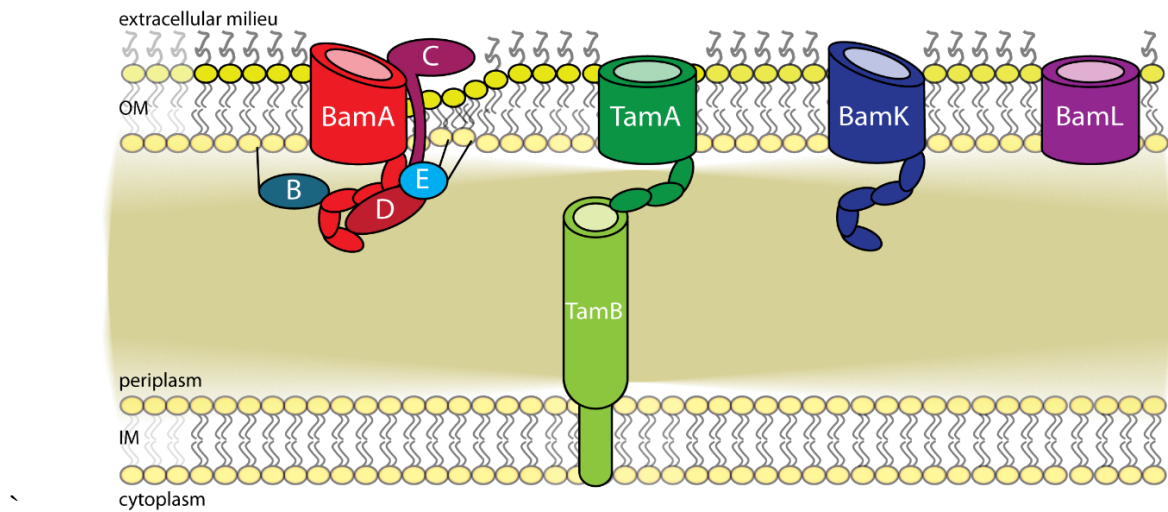
Figure 1.5.1: The Omp85 family members of *K. pneumoniae*. In *K. pneumoniae*, four distinct Omp85 members are present within their genomes. Studies in closely phylogenetically related species have found that the BamA and TamA play roles in OMP biogenesis; however, two more Omp85 family members are conserved within the core genome of *K. pneumoniae*. The pervasiveness of *K. pneumoniae* to harbour a second copy of BamA is curious (BamK). Does this gene duplicate encode for a functional gene? Does it play a role in OMP biogenesis, as its sequence is highly similar to the Omp85 blueprint – BamA? The second *K. pneumoniae* specific Omp85 is part of the noNterm subfamily of Omp85s (BamL). The assembly and insertion of OMPs in Gram-negative bacteria have all featured foldase and insertase machineries where the Omp85 components feature at least one POTRA domain. Does this family member also play a role in assembly and insertion of OMPs or protein secretion? The aim of this thesis is to understand the role of these divergent *K. pneumoniae* Omp85s.

1.6 Thesis rationale and aims

Despite the impressive advances in our knowledge of OMP biogenesis as described and detailed in this chapter, we understand surprisingly little about the molecular mechanisms and overlapping functions of Omp85 biogenesis factors, especially in organisms other than *E. coli*. Work presented here describes efforts to address these gaps in our knowledge using a *K. pneumoniae* model. In particular, I sought to address:

1. What are the comparative gene expression levels of Omp85s in *K. pneumoniae*? What genetic regulation controls the expression of Omp85s in *K. pneumoniae*? (**Chapter 3**)

2. What is the functional role of Omp85 without POTRA domains? Do these members contribute to OMP biogenesis through folding and inserting or by providing a translocon pore? (**Chapter 4**)

3. Why does *K. pneumoniae* retain two copies of the essential gene *bamA*? Are there two BAM complex folding pathways? Are there any differences in function or structure between the two proteins? (**Chapter 5**)

2

Materials and Methods

2.0 Bacterial growth media and conditions

E. coli and *K. pneumoniae* strains used in this study are listed in Table 2.0.1. Bacterial strains were cultured in sterile lysogeny broth (LB) liquid medium (CSH Protocols, 2016), or M9 minimal medium (standard) (CSH Protocols, 2010) supplemented with yeast nitrogen broth, or pre-sterilised Dulbecco's Modified Eagle Medium (DMEM) (Gibco).

For pulse chase assays, M9+S media contained: 1 mM $MgSO_4$, 0.1 mM $CaCl_2$, 1.12 mM thiamine, 0.2% (wt/vol) glucose and M9 salts, as a nitrogen source - an amino-acid drop-out mixture was used: 45.4 $pg.mL^{-1}$ of each standard amino acid, except leucine (227 $pg.mL^{-1}$) instead of ammonium chloride. A non-sulphur minimal media (M9-S) was also used, which was M9+S prepared without methionine and cysteine and 1 mM $MgCl_2$ in place of $MgSO_4$.

Culture media supplemented with carbohydrates were filter sterilised using Stericups® with 0.22 μm Express™ Plus membranes (Merck Millipore), but where media was required at only low volumes (less than 50 mL) it was filter sterilised using an Acrodisc® syringe filter with a 0.2 μm Supor® membrane (PALL Life Sciences). In all other cases, media was sterilised by autoclaving at 121 °C for 20 minutes at 15 psi.

Plasmids used in this study are listed in Table 2.0.2. When required for maintaining plasmid carriage, growth media was supplemented with antibiotics at the following concentrations: 100 $μg.mL^{-1}$ ampicillin, 30 $μg.mL^{-1}$ kanamycin, 34 $μg.mL^{-1}$ chloramphenicol or 200 $μg.mL^{-1}$ rifampicin . Where solid media was required, 15 $g.L^{-1}$ agar (Merck Millipore, 12177) was added to LB growth media prior to autoclaving.

Unless otherwise indicated, liquid cultures of *K. pneumoniae* and *E. coli* were routinely incubated at 37 °C on an orbital platform rotating 200 strokes per minute (25 mm orbit). For growth on solid

medium, *K. pneumoniae* and *E. coli* were incubated on LB agar for 16-24 hours at 37 °C. In typical experiments, single colonies were then selected and used to inoculate fresh LB (5-10 mL liquid) media where cells were subsequently incubated for 12 - 16 hours. This saturated culture was typically used to prepare glycerol stock solutions, and/or to inoculate fresh media. To estimate the growth phase of *E. coli* and *K. pneumoniae*, the optical density at 600 nm (OD_{600}) was measured with an Ultrospec 10 cell density meter (GE Healthcare Life Sciences) or using the Spark™ 10M multimode plate reader (Tecan).

Storage conditions and bacterial recovery

To create frozen glycerol stocks, an aliquot of saturated culture (from routine culturing) was transferred to a cryogenic vial and mixed with an equal volume of 80 % v/v glycerol (40 % v/v final concentration) then placed in to a -80 °C freezer for long-term storage. Plasmid stocks were stored in *E. coli* DH5α Δ*recA* mutant. For bacterial recovery, glycerol stocks were removed from the -80 °C storage freezer, preventing the full thawing of the glycerol stock, and a small portion of the frozen the glycerol culture was scratched away and used for streak plating on solid LB media for subsequent inoculation of starter cultures.

Preparation of competent bacterial cells and transformation

To prepare chemically competent cells with $CaCl_2$, saturated overnight cultures were diluted 1:100 into fresh 200 mL LB and incubated until an OD_{600} of ~0.6. Cells were chilled for 30 minutes on ice, subjected to centrifugation (*E. coli* - 5000 × g, 10 minutes, 4 °C), and the cell pellet was resuspended in 28 mL ice-cold 0.1 M $CaCl_2$. After chilling and following centrifugation to isolate the cells, the cell pellet was resuspended in 1 mL ice-cold 0.1 M $CaCl_2$ and incubated on ice for 2 hours. Ice-cold LB media containing 30% v/v glycerol (500 μL) was then added to the suspension and 100 μL aliquots were snap frozen in liquid nitrogen. The cell aliquots snap frozen with liquid nitrogen and were stored in a -80 °C freezer.

To transform $CaCl_2$-competent *E. coli* strains, cells were first thawed on ice. Purified plasmid (50 ng) was added and, after gentle mixing, the sample was incubated on ice for 10 minutes. The sample was then heat-shocked for 90 seconds at 42 °C, incubated on ice for 5 minutes, and transferred to 950 μL of non-selective LB media. Selective solid media was then spread plated with 500 or 1000 μL of transformation mixture. Following overnight incubation (typically 12-16 hours) at 37 °C, a well-isolated single colony of transformants was used to inoculate fresh LB (liquid) media for overnight incubation.

For high efficiency transformations, electrocompetent *E. coli* or *K. pneumoniae* were used. To prepare electrocompetent cells, saturated overnight cultures were diluted 1:100 into fresh 200 mL LB and incubated until an OD_{600} of ~0.4, cells were then centrifuged four times (*E. coli* $5000 \times g$ or *K. pneumoniae* $15000 \times g$, 10 minutes, 4 °C) where the cell pellet was resuspended using increasingly smaller volumes of ice-cold glycerol (10 % v/v) (e.g. 25 mL, then 12.5 mL, then 6.25 mL, then 1 mL. Aliquots (100 μL) were either used for transformation immediately, or snap frozen in liquid nitrogen and stored at -80 °C. When required, frozen aliquots were thawed on ice before incubation with plasmid.

For transformation, electrocompetent cells were incubated on ice for 3 minutes with 50 ng of purified plasmid. After transferring to pre-chilled ice-cold electroporation cuvettes (1 mm gap), cells were electroporated (*E. coli* 1800 V or *K. pneumoniae* 2500 V, 200 Ω and 100 μF) and then immediately transferred to 900 μL of non-selective LB media. Following a 1-hour incubation, selective solid media was spread plated with 500 -1000 μL of transformation mixture. Following overnight incubation (typically 12-16 hours) at 37 °C, a well-isolated single colony of transformed bacteria was used to inoculate fresh LB (liquid) media for overnight incubation.

2.1 DNA/RNA based techniques

The plasmids and synthesised DNA fragments (gBlocks™) used in this study are listed in Table 2.0.2. Oligonucleotides used in this study are listed in 2.0.3.

Plasmid isolation

Plasmids were isolated from saturated overnight cultures (5 mL) using the Wizard® Plus SV Miniprep DNA purification system (Promega) kit as per manufacturer's instructions.

Genome isolation

Genomic DNA was isolated from saturated overnight cultures using the *GenElute™* Bacterial Genomic DNA Kit (Sigma-Aldrich) as per manufacturer's instructions.

Polymerase chain reaction (PCR)

Oligonucleotide primers used for PCR were designed using SnapGene® viewer (GSL biotech) and were synthesised by Integrated DNA Technologies (USA). For PCR amplifications, Phusion® High-Fidelity DNA polymerase (New England Biolabs) or Taq DNA polymerse (Roche) kits were used as per manufacturer's instructions. The template DNA used for PCR were either isolated genomic DNA (gDNA), synthesised DNA fragments "gBlocks" (IDT), or DNA in

a single colony picked from solid media after an overnight incubation and resuspended directly into the PCR reaction. DNA amplifications were stained using SYBR® Safe (Thermo Fisher Scientific) as per manufacturer's instructions and visualised using a LAS-3000 Imaging system (Fuji). Slices corresponding to the amplified PCR fragments were purified using the Wizard® SV Gel and PCR Clean-Up System kit (Promega).

Restriction digestion and alkaline phosphatase treatment

Restriction enzymes (New England Biolabs) used for digestion of plasmids or purified PCR amplicons were used as per manufacturer's instruction. Linearised plasmids were treated with alkaline phosphatase (Promega, M182A) for 1 hour as per manufacturer's instructions. After alkaline phosphatase treatment and/or restriction enzyme digestion, DNA samples were analysed by gel electrophoresis using 0.5-2.5 % w/v molecular biology grade agarose (Scientifix, 9010E) in TAE buffer (CSH protocols, 2013). DNA was stained using SYBR® Safe (Thermo Fisher Scientific) as per manufacturer's instructions and visualised using a LAS-3000 Imaging system (Fuji). Slices corresponding to the digested vector or insert were extracted and DNA was purified using the Wizard® SV Gel and PCR Clean-Up System kit (Promega).

DNA Ligation

The relative amount of purified plasmid and insert DNA was estimated using the NanoDrop 1000 Spectrophotometer (Thermo Scientific) as per manufacturer's instructions. Ligation using the T4 DNA Ligase (New England Biolabs) system was performed as per the manufacturer's instructions, usually using a 1:3 vector:insert ratio. Ligation mixture (10-20 µL) was used to transform chemically competent *E. coli* DH5α (instead of 1 µL purified plasmid).

DNA sequencing

Plasmid and/or PCR products were sequenced by Macrogen, Inc. (South Korea). The sequence data obtained was then evaluated manually to determine if the desired deletions, insertions or mutations were present.

Total RNA isolation

K. pneumoniae cultures were grown to exponential phase (OD$_{600}$ ~0.6) or stationary phase (OD$_{600}$ ~0.6) phase and 5 mL of cell culture media was added to 10 mL of RNAprotect Bacteria Reagent (Qiagen). Samples were immediately vortexed for 5 seconds and incubated for an additional 5 minutes at room temperature. Samples were then centrifuged ($15000 \times$ g, 10 minutes, room temperature) and the supernatant discarded. Total RNA was then isolated from the bacterial pellets using a RNeasy Plus Mini kit (Qiagen) according to the manufacturer's instructions. Extracted RNA samples were quantified using a NanoDrop 1000 spectrophotometer and the quality of the isolated RNA was checked by visualisation on agarose gels and by measuring the sample's A$_{260}$:A$_{280}$ ratio (>1.8).

qPCR analysis

Two micrograms of isolated total RNA was used for cDNA synthesis, using 70 ng of random hexamers (Invitrogen) and Superscript II Reverse Transcription kit (Invitrogen), all according to the manufacturer's guidelines. Quantification of the expression of the target genes was performed using SYBR Green Master (Roche) with primer pairs used at 0.125 µM each; Cycling conditions using the LightCycler480 instrument (Roche) were as follows: 95°C for 10 min followed by 35 cycles of 95 °C for 30 sec, 55 °C for 60 sec and 72 °C for 15 sec. Primer specificity was determined by melting curve analysis and normalised cycle threshold (Ct) values were obtained using open source qPCR data analyser LinReg (Ruijter *et al.*, 2009). All qPCR sample reactions were performed in triplicate.

Construction of *K. pneumoniae* mutant strains

Knock-out and knock-in mutants, in which target genes were deleted or replaced by allelic exchange with a kanamycin resistance-encoding gene, were constructed in *K. pneumoniae* B5055 using the "gene gorging" technique (Herring *et al.*, 2003, Lee *et al.*, 2009). All oligonucleotide primers: for generation of DNA regions, flanking homology arms, antibiotic cassettes or plasmid backbones, are listed in Table 2.0.2.

'Donor' plasmids carrying the desired mutation were constructed as follows. The kanaymcin gene was amplified from pKD4 using primers kanF and kanR. The resulting product included flanking fragment length polymorphism (FLP) recombinase target (FRT) sites to permit subsequent kanamycin cassette excision through use of a helper plasmid. DNA fragments were joined using the Gibson assembly® Master Mix (New England Biolabs) as per manufacturer's instructions.

Plasmid pACBSR carries genes encoding the I-SceI endonuclease and lambda Red recombinase where their expression can be induced by addition of L-arabinose. Donor plasmids and pACBSR were transformed into electrocompetent B5055 cells (0.1 cm gap-width cuvette; 200 ohms, 25 μF, 1.8 kV) and selected on LB agar containing kanamycin and chloramphenicol. A single co-transformant was inoculated into 1 mL LB containing 0.2% L-arabinose (Sigma-Aldrich) and chloramphenicol and grown in a shaking incubator at 37°C for 16 hours. Cell dilutions were grown on LB agar containing kanamycin, and resultant colonies were screened by colony PCR using primers flanking the targeted region. The loss of pACBSR was induced by 0.2% L-arabinose without selection. When required, the kanamycin resistance gene was excised via the FRT sites using the FLP helper plasmid pCP20. All mutations were confirmed by PCR analysis, using primers flanking the targeted regions. Schematics for donor plasmids used for gene gorging can be found in the following sections: Figures 3.3.2, 3.45, 4.2.2, 5.3.2 and 5.5.1.

Generation of a gain-of-function mutagenesis library and insertion site screening

Separate overnight starter cultures (5 mL) of *E. coli* S17-λ*pir* harbouring suicide vector pUT-mini-Tn*5Km2* and *K. pneumoniae* B5055^{Rif} Δ*bamK*::Amp^R were diluted into 25 mL of fresh LB media to a starting OD$_{600}$ of 0.05 with appropriate antibiotics. Cultures were grown to logarithmic phase (OD$_{600}$ ~0.6) and were mixed in a 1:1 ratio in a final volume of 10 mL for 1 hour at 37 °C static growth for liquid mating. The conjugation mix was then centrifuged ($15000 \times$ g, 10 minutes, room temperature), resuspended in 1 mL LB and spread plated onto non-selective LB agar and incubated at 37 °C overnight for solid mating. The resultant confluent bacterial growth was scraped and resuspended in LB (1.5 mL) and spread plated onto LB agar plates containing kanamycin and rifampicin and incubated at 37 °C overnight. Single *K. pneumoniae* B5055^{Rif} Δ*bamK*::Amp^R colonies which showed kanamycin (transposon marker) and rifampicin resistance were replica plated to LB plates that contained kanamycin, rifampicin and ampicillin plates incubated at 37 °C overnight. The *K. pneumoniae* B5055^{Rif} Δ*bamK*::Amp^R colonies which showed kanamycin (transposon marker), rifampicin and ampicillin resistance (*bamK* proxy expression) were subsequently store at -80 °C. The sequences flanking the Tn*5Km2* transposons were amplified by Y-linker ligation PCR (Kwon & Ricke, 2000).

DNA binding reactions and Electrophoretic mobility shift assay (EMSA)

DNA-binding activity was determined by mixing cleared cell lysate, containing putative DNA-binding proteins, with double-stranded DNA amplified by PCR to represent regions of the putative *bamK* promoter. Cleared cell lysate was isolated from *K. pneumoniae* B5055 using

detergent-based protein extraction reagent B-PER II (Thermo-Fisher) as per manufacturer's instructions. A constant amount of purified PCR amplicons (250 nmol) acted as the "bait" titrated against an increasing amount of "prey" DNA-binding protein present in the cleared cell lysate. Binding reactions were performed in 20 µL final volume with a binding buffer (10 mM Hepes pH 8.0, 50 mM NaCl, 20 mM MgCl, 20 mM DTT). The binding reaction was performed at room temperature for 30 minutes. The synthetic polynucleotide, poly(dI-dC) (Roche), was used as a competitor for nonspecific DNA binding and was added to the purified DNA amplicons for 30 minutes before the addition of cleared lysate. Prior to loading for TBE-PAGE, 6× EMSA loading buffer (0.25% bromophenol blue, 10 % glycerol (v/v)) was added to the binding reaction and ran on a TBE gel for 120V until the bromophenol blue marker migrated three quarters down the gel. DNA-protein bands were detected by placing the TBE-PAGE gel into a GelRed™ DNA staining solution for 10 minutes then visualised using a Bio-Rad Chemidoc XRS+ as per manfacturer's instructions.

2.2 Protein based techniques

Polyacrylamide Gel Electrophoresis

Proteins were analysed by polyacrylamide gel electrophoresis (PAGE) using Mini- PROTEAN® Tetra vertical electrophoresis cell (Bio-Rad) equipment as per manufacturer's instructions. Polymerization of in-house gel solutions was initiated by the addition of 50 µL of APS (20% w/v) and 10 µL of TEMED, per 10 mL of gel solution, immediately before casting the gel.

For sodium dodecyl sulphate (SDS)-PAGE analysis previously described methods were used but with some modifications (J Simpson, 2006). Gels were cast using a 0.75 mm spacers and combs. Stocks for the two gel layers were at 4× resolving gel buffer (1.5 M Tris pH 8.8, 0.4 % w/v SDS and 2 mM EDTA) and 4× stacking gel buffer (1.0 M Tris pH 6.8, 0.08% w/v SDS and 2 mM EDTA). For resuspension or dilution of protein samples, the final concentration for SDS sample buffer was: 100 mM Tris-HCl pH 6.8, 10 % v/v glycerol, 1 % w/v SDS, 100 mM DTT and 0.01 % w/v bromophenol blue. A 40% acrylamide solution (Bio-Rad, #1610146) was adjusted accordingly for the resolving (gradient 4-16 % or uniform 8 %, 10 %, 12 % or 15 %) for specific applications. The same 40 % stock was used to prepare the stacking gel (3 %) solutions. Gradient gels were prepared using the Hoefer™ SG 50 gradient maker (GE Healthcare Life Sciences) as per manufacturer's instructions. For molecular weight markers, Dual colour Precision Plus Protein™ Pre-stained standards (Bio- Rad) were diluted 1:10 with SDS sample buffer (without

boiling) and 5 µL was loaded into polyacrylamide gels for SDS-PAGE analysis. SDS running buffer comprised of 25 mM Tris, 192 mM glycine and 0.1 % SDS was used to separate proteins by electrophoresis. Protein samples loaded per well were normalised where appropriate and typically volumes of 1-10 µL or 5-50 µg were used unless otherwise stated.

For semi-native (SN)-PAGE analysis, a similar method to SDS- PAGE was used but that buffers were of a modified composition. Instead of SDS running buffer, MES running buffer was used, containing 50 mM 2-(*N*- morpholino)ethanesulphonic acid (MES) (50 mM MES pH 7.3, 2 mM EDTA and 0.2 % w/v SDS). Instead of SDS sample buffer, SN sample buffer (40 mM Tris-HCl pH 6.8, 10 % v/v glycerol, 0.2 % w/v SDS, 0.01 % w/v bromophenol blue) was used. Instead of 4× resolving and stacking gel buffers, 4× SN gel buffer was used (2.0 mM EDTA and 1.5 M Tris pH 8.8) with no stacking layer. SN-PAGE analysis used 4-16 % gradient gels for pulse-chase assay experiments. Gradient gels were prepared using the Hoefer™ SG 50 gradient maker (GE Healthcare Life Sciences) as per manufacturer's instructions. For molecular weight markers, Kaleidoscope™ Pre-stained standards (Bio- Rad) were diluted 1:10 with SN sample buffer (without boiling) and 5 µL was loaded into polyacrylamide gels for SDS-PAGE analysis. Protein samples loaded per well were normalised where appropriate and typically volumes of 1-10 µL or 5-50 µg were used unless otherwise stated.

For TBE-PAGE analysis non-denaturing and reducing conditions were used to facilitate protein-DNA interactions. Gels were casted using a 0.75 mm spaced plates and well combs. Similar to SN-PAGE analysis, TBE-PAGE analysis also did not feature a stacking gel layer. TBE gel solution was comprised of a final concentration of 0.5× TBE buffer (CSH Protocols, 2006) adjusted with 40 % acrylamide solution (Bio-Rad, #1610146) and deionised water (MilliQ®). TBE-PAGE analysis was performed using 4-8 % gradient gels for DNA-protein binding experiments. Gradient gels were prepared using the Hoefer™ SG 50 gradient maker (GE Healthcare Life Sciences) as per manufacturer's instructions.

Blue native (BN)-PAGE was used to resolve OM protein complexes of *E. coli* and *K. pneumoniae* under native conditions. Gels were cast using 1.5 mm width spacers and well combs. A 3× BN-gel buffer (150mM Bis Tris pH 7.0, 200 mM n-amino caproic acid) was used for the resolving and stacking gel solution. A 40% acrylamide solution (Bio-Rad, #1610149) was adjusted accordingly for the resolving and stacking solutions. Based on previously published work (Webb *et al.*, 2012b), a BN-PAGE sample buffer comprising 5% Brilliant blue G, 500 mM amino caproic acid in 100 mM Bis-Tris pH 7.0 and 5-16 % gradient gels were used to visualise the BAM

complex. Gradient gels were prepared using the Hoefer™ SG 50 gradient maker (GE Healthcare Life Sciences) as per manufacturer's instructions. For molecular weight markers, 0.6 mg mL^{-1} ferritin, 0.6 mg mL^{-1} BSA 0.5 mL^{-1} was diluted with BN sample buffer to 1 mL final volume (without boiling) and 50 µL was loaded into polyacrylamide gels for BN-PAGE analysis. An anode buffer comprised of 50 mM Bis-Tris pH7.0 and cathode buffer comprised of 15 mM Bis-Tris pH 7.0, 50 mM Tricine, 0.02% Brilliant Blue G250, 0.03% DDM were used to separate native proteins by electrophoresis.

Protein samples of 50 µg of total membranes were prepared by adding ACA750 buffer (50 mM Bis-Tris pH 7.0, 750 mM n-amino caproic acid, 0.5 mM EDTA) to a final volume of 18 µL then solubilised with 2 µL 10% DDM for 20 minutes on ice with gentle mixing. Samples were cleared by centrifugation (20,000 $\times$ g, 10 minutes, 4 °C) to pellet any non-solubilised membranes. The solubilised portion was added to 5 µL of 5$\times$ BN-sample buffer 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.

Coomassie staining and destaining of polyacrylamide gels
For some applications, following SDS-PAGE analysis gels were stained with Coomassie staining solution, comprising 50 % v/v methanol, 10 % v/v acetic acid and 0.05 % w/v Coomassie Brilliant Blue R-250. Gels were incubated in Coomassie staining solution for at least 1 hour with gentle shaking (40-50 rpm, 32 mm orbit), and the stain was then gradually removed with a destaining solution comprising 5 % v/v methanol and 7 % v/v acetic acid. Gels were usually destained for up to 24 hours at 40-50 rpm (32 mm orbit), until the background staining was minimal and the proteins of interest remained visible.

Western transfer and immunoblotting
Following PAGE analysis, proteins were transferred from gels to 0.45 µm hydrophobic immobilon-P PVDF transfer membranes (Merck Millipore), or transferred to 0.45 µm Protran nitrocellulose hybridisation transfer membranes (PerkinElmer). Western transfer was performed using mini Trans-Blot® electrophoretic transfer cell (Bio-Rad) as per manufacturer's instructions, except that transfer buffer A also contained 10 % v/v methanol. In the case of membrane transfer BN-PAGE analysis, gels were first soaked in 10$\times$ SDS running buffer for 10 minutes (additional denaturation of proteins aiding antibody detection) then transferred as detailed above.

For immunoblotting, antibodies and conjugates used as detection reagents in this study are listed in Table 2.0.4. Primary antibodies, secondary antibodies and conjugates were diluted in blocking

solution (5 % (w/v) instant skim milk powder (Coles)) in Western washing buffer. Western washing buffer was prepared as per CSH Protocols (2013) except the final concentration of Tween-20 was 0.2 % v/v. All incubations were performed at room temperature at 40-50 rpm (25 mm orbit) unless otherwise indicated. Primary and secondary antibodies were stored at -80 °C and -20 °C, respectively, and when required aliquots were thawed on ice and subsequently stored at 4 °C. Antibodies (either concentrated or diluted in blocking solution) were usually supplemented with 0.025 % w/v sodium azide for prolonged storage at 4 °C, except for antibodies conjugated to horse-radish peroxidase (HRP). For direct HRP conjugates, Precision Protein™ StrepTactin- HRP conjugate (Bio-Rad) and HisProbe™-HRP Conjugate (Thermo Scientific), these probes substituted as the primary antibody incubation but did not require a secondary antibody incubation.

Following Western transfer, membranes were incubated for 1 hour in blocking solution at room temperature. After removing the blocking solution, the membrane was incubated for one hour in the diluted primary antibody solution. The primary antibody solution was removed and stored at 4 °C for reuse, and the membrane was washed thrice for 5 minutes in Western washing buffer. The membrane was incubated for 45 minutes in secondary antibody solution and washed as before. Strep-Tag II or a polyhistidine tags were fused to the Biorad™ pre-stained markers, thus lanes containing the standards were removed from the membrane prior to immunoblotting, and then reattached (for size referencing) before exposure to film. This was to ensure the extremely strong signal from the markers did not interfere with the much weaker signal from the protein(s) of interest.

Protein detection was performed using the enhanced chemiluminescence method with the ECL Prime Western blotting detection reagent (GE Healthcare Life Sciences) as per manufacturer's instructions. Chemiluminescent membranes were then exposed to Super RX-N film (Fujifilm) in an Amersham Hypercasesette™ (GE Healthcare Life Sciences) for up to 1 hour (usually 5 minutes). Immunoblot images were scanned using CanoScan 8600F with software version 5.0.1.2 (Canon).

Mass spectrotrometry (MS) analysis
DNA-protein gel bands were analysed by liquid chromatography-mass spectrometry (LC-MS) to identify DNA binding protein candidates. Gel pieces were washed, and the samples reduced and alkylated, and digested overnight with Trypsin (Promega) in a 20mM ammonium bicarbonate buffer with a total sample volume of total of 50 μL. Tryptic digests were analysed by LC-MS/MS

using the Ultimate 3000 nano HPLC (Thermo Scientific) coupled with a MicroTOFq mass spectrometer (Bruker Daltonics). For separation of peptides, digested samples were loaded onto a *75* μm i.d. PepMap™ 100 (C18) nanocolumn and eluted at a flow rate of 300 nL/minute using a linear gradient from 95% buffer A (0.1% Formic acid) to 70% B (80% Acetonitrile 0.1% formic acid). The eluted peptides were nebulised and ionised using the nanoESI source with a capillary voltage of 4000V. The chromatographic system was interfaced to an amaZon ion trap (Bruker-Daltonics) operating in AutoMSn (with $n=2$) in Enhanced Resolution (maximum speed=8100 m/z per s) for MS mode and Ultrascan Mode for MS/MS. Prior to analysis, the qTOF mass spectrometer was calibrated using tune mix.

Data from LCMSMS run was exported in Mascot generic file format (*.mgf) and searched against an in-house curated database of protein sequences obtained from Uniprot using the MASCOT search engine (version 2.4). The following search parameters were used: missed cleavages, 1; peptide mass tolerance, ± 20ppm; peptide fragment tolerance, ± 0.04 Da; peptide charge, 2+, 3+ and 4+; fixed modifications, carbamidomethyl; Variable modification, oxidation (Met).

Pulse-Chase Assay

Radioactive labelling of proteins and pulse-chase assays were performed as previously described (Stubenrauch *et al.*, 2016), with some modifications. Overnight cultures of *E. coli* BL21 Star™ (DE3) strains harbouring expression plasmids (inserts for LptD, LptE, LptDE or PhoE) were diluted 1:50 with fresh LB and incubated to an OD_{600} of ~0.6 at 37 °C with shaking. Cells were subjected to centrifugation (4600 × g, 10 min, 4 °C), washed and resuspended in M9-S media. Following a 30-minute incubation at 37 °C with shaking, cells were diluted with an equal amount of glycerol (40 % (v/v), aliquoted and snap frozen in liquid nitrogen. Aliquots from the same batch were considered to be technical replicates, and aliquots from different, independent batches were considered to be biological replicates.

When required, aliquots were thawed on ice, subjected to centrifugation (3000 × g, 5 min, 4 °C) and resuspended in M9-S media. Following addition of 0.5 mg/mL rifampicin, cells were incubated for 1 hour at 37 °C, 400 rpm (3 mm orbit), 0.2 mM isopropyl-β-D-thiogalactoside (IPTG) was added and cells were incubated for a further 5 minutes. Samples were 'pulsed' by addition of 22 μCi mL−1 [^{35}S]-methionine and [^{35}S]-cysteine (NEG072, Perkin Elmer) for 60 s without shaking, subjected to centrifugation (3000 × g, 5 min, 4 °C), and then 'chased' by resuspension in M9+S media for up to 32 min (the chase temperature was 37 °C (static) for both PhoE and LptDE assembly assays).

To track PhoE or LptDE assembly, aliquots were removed to SN-sample buffer at 10 seconds, 2, 4, 8, 16 and 32 min. Samples were incubated for 10 minutes at 37 °C or at 100 °C as a control to denature any assembled oligomers. Samples were analysed by SN-PAGE using 4–16% gradient gels. Proteins were transferred to 0.45 µm nitrocellulose membranes using standard techniques, membranes were air-dried, and radiation was captured using a storage phosphor screen (GE Health Sciences) and detected using Typhoon Trio (at 320 nm).

From at least three biological replicate gels (n=3), PhoE or LptDE band densities were measured using ImageQuant™ TL software version 7.0 (GE Health Sciences) with the local average background correction tool. To be able to compare biological replicates, 'normalised density' values were calculated by dividing the density measurements of the six timepoints by the greatest density measurement among those timepoints. To calculate the observed rate constant, the 'normalised density' values were subsequently determined using the one-phase decay or exponential nonlinear regression tool from GraphPad Prism software version 7.02 (GraphPad software, Inc.).

Isolation of total membranes from *E. coli* or *K. pneumoniae*

A 5 mL overnight starter bacterial culture of E. coli or K. pneumoniae was diluted into 400 mL of fresh LB media to a starting OD600 of 0.05. Cultures were grown to logarithmic phase (OD600 ~0.6) and cells were harvested by centrifugation (E. coli $5000 \times g$ or K. pneumoniae $15000 \times g$, 10 min, 4 °C). For strains harbouring an inducible plasmid, cultures were grown to logarithmic phase (OD600 ~0.6) then induced with arabinose to a final concentration of 0.05 % (w/v) for 3 hours. Cell pellets were then resuspended in TS buffer (10mM Tris-HCl, 0.75 M sucrose pH 7.5). To aid in cell lysis, addition of lysozyme (50 µg/ml), PMSF (2mM) and 2 volumes of 1.65 mM EDTA (pH 7.5), were added sequentially before homogenising cells (2-3 passes) with an EmulsiFlex (Avestin Inc.) at 15,000 psi. Debris was pelleted ($15000 \times g$, 10 minutes, 4 °C) and soluble membranes were collected by ultracentrifugation ($100000 \times g$, 10mM Tris-HCl, 1 hour, 4 °C) and washed in TES (3.3 mM Tris-HCl, 1.1 mM EDTA, 0.25 M sucrose pH 7.5) buffer. Membranes were pooled (~8 ml) with TES and collected by centrifugation ($100000 \times g$, 1 hour, 4 °C) and then resuspended in ~400 µl 25% (w/v) sucrose in 5 mM EDTA (pH 7.5) and stored at -80 °C.

Sucrose density fractionation

To separate outer and inner membrane fractions based on density differences, total membrane samples (~400 µl) resuspended in 25% (w/v) sucrose were overlain on a six-step sucrose gradient (1.9 mL each of 35:40:45:50:55:60% (w/v) sucrose in 5 mM EDTA (pH 7.5). After isopycnic

density centrifugation using ultracentrifugation in a SW40 Ti rotor (34,000 rpm, 17 hours, 4 °C), the sucrose gradient was fractionated into 1 mL fractions, which were stored at − 80 °C until required for analysis by SDS-PAGE and immunoblotting.

Identification of exoproteins in culture supernatants

A 5 mL overnight starter culture of *K. pneumoniae* B5055 harbouring pBAD-bcLV was sub-cultured into 25 mL of fresh LB media to a starting OD_{600} of 0.05. Cultures were grown to logarithmic phase (OD_{600} ~0.6) then induced with arabinose to a final concentration of 0.05 % (w/v) for 3 hours. Two different methods were used to enrich for proteins in the culture supernatants which were presumed to represent secreted exoproteins. The first method was direct precipitation of proteins in the culture supernatant using trichloroacetic acid (TCA). Culture supernatants were initially passed through a 0.2 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 ice cold 100% methanol. Pellets were air-dried to remove excess methanol and resuspended in 50 µl SDS sample buffer. The second method was enrichment of exoproteins in the culture supernatant using a centrifugal filter. Bacteria were centrifuged at $4000 \times g$ and supernatants were aspirated and clarified by using 0.2 µm filters. Exoproteins were concentrated by centrifugal filtration with 3 kDa cut-off Amicon ultra filtration tubes (EMD Millipore) to a final volume of 1 mL. Enriched exoprotein samples were analysed by 3-14% gradient SDS-PAGE gels for subsequent Coomassie blue staining and immunoblotting.

Urea Extraction of peripherally associated membrane proteins

Total membrane samples (50 µg protein) from *K. pneumoniae* B5055 harbouring pBad24-bcLV were treated with 1 mL of 1× PBS (CSH protocols, 2006) or 5M urea and allowed to mix on a rotary wheel for 1 hour at room temperature. Insoluble and extracted fractions were separated by ultracentrifugation (100,000 x g, 50 minutes, 4 °C). The soluble fraction containing the urea-extracted proteins was precipitated with TCA (10% final concentration) on ice for 30 min. Precipitated proteins were collected by centrifugation (20000 x g, 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 SDS sample buffer. As controls, equivalent amount of total membranes (50 µg) of the tested strains were loaded in parallel to demonstrate the initial material amount prior to PBS or urea treatment. Proteins were separated by SDS-PAGE (12%) and analysed by immunoblotting.

Protein-protein interaction: affinity purification

To assay for binding between BamL and LupV, affinity purification was undertaken using the introduced C-terminal tags and affinity resins (i.e. Strep-Tactin® Superflow® 50% suspension (IBA lifesciences) and Ni-NTA resin (Thermo scientific) for polyhistidine tags) (Refer to Figure 4.5.1). Total membrane samples (50 µg protein) isolated from *K. pneumoniae* B5055 harbouring pBad24-bcLV were initially solubilised using 1% DDM. To remove insoluble material, the solubilised membranes were centrifuged (20,000 × g, 15 min, 4 °C) and the resulting soluble fraction was added to 50 µL of Strep-Tactin® Superflow® or Ni-NTA resin, and the sample gently rotated at 4 °C for 1 hour. The resin samples were washed with wash buffer (20 mM Tris-HCl, 0.15 DDM, 50 mM NaCl, pH 7.5, 1 mM phenylmethyl sulphonyl fluoride (PMSF) and 10% (w/v) glycerol) three times before the bound protein fraction was eluted directly with SDS–PAGE sample buffer. In the case of affinity purification assays utilising Strep-Tactin® Superflow® 50% suspension (IBA), proteins were instead eluted using Buffer E (IBA) a desthiobiton based elution buffer. The eluted proteins were analysed by SDS–PAGE and immunoblotting.

BamL Protein purification

Unless stated otherwise, the entire purification protocol was performed on ice. *E. coli* C41 cultures harbouring plasmid pET-20b(+)-BamL were routinely cultured to OD_{600} 1.0 before addition of 0.1mM IPTG for overnight protein expression induction. Cells were harvested by centrifugation at (5000 × g, 10 minutes, 4 °C), resuspended in lysis buffer (50 mM Tris-HCl pH 7.5, 0.15 mM DDM, 150mM NaCl, 300 µg/mL lysosome, 1 µg/mL DNase solution) and then disrupted using a high-pressure homogeniser (Emulsiflex-B15, Avestin). Cell debris was pelleted by centrifugation at (15,000 × g, 20 minutes, 4 °C) and crude membrane fractions were collected from the remaining cell extract by ultracentrifuge (100000 × g, 1 hour, 4 °C).

For solubilisation of membrane-bound proteins, Elugent (Calbiochem) was added to the crude membrane preparations to a final concentration of 2% in a final volume of ~60 mL. The preparation was left to solubilise for 30 minutes at room temperature with end-over-end rotation. Insoluble debris was removed from the preparations by centrifugation (15,000 x g, 20 minutes, 4 °C) and soluble preparation was affinity purified using a HisTrap column (GE Healthcare) on an ÄKTAexplorer system (GE Healthcare). The column was equilibrated with a buffer consisting of 50 mM Tris–HCl pH 8.0, 300 mM NaCl, 10 mM imidazole. The target protein was eluted with a buffer consisting of 50 mM Tris–HCl pH 8.0, 100 mM NaCl and 1 M imidazole using a linear gradient from 0.1 to 50%. The presence and purity of the BamL in the resulting elution fractions was confirmed by SDS-PAGE (12%).

In order to further purify the recombinant BamL, fractions from the HisTrap column were pooled and concentrated utilising a centrifugal filter with a 30 kDa cutoff. The salt concentration of the final sample was set to 150 mM through buffer exchange to suit the buffers utilised in Size exclusion chromatography. The buffer exchanged and concentrated recombinant BamL samples were loaded onto a Superdex 200 column (GE Healthcare) connected to an ÄKTAexplorer system (GE Healthcare) and gel filtration buffer (50 mM Tris-HCl, 0.15 mM DDM, 100mM NaCl pH 7.5) was pumped through at a flow rate of 1 mL/min at room temperature. The resulting elution fractions were tested by 12% SDS-PAGE to confirm the presence and purity of recombinant BamL, and then combined for antibody production, protein crystallisation or saved in -80°C for future analyses.

If detergent exchange was required, eluted recombinant BamL from the size exclusion step were pooled and bound to 50 µL of Ni-NTA resin (Thermo scientific). Samples were gently rotated at 4 °C for 1 hour. The bound resin samples were then washed with wash buffer (50 mM Tris-HCl, 0.15 mM DDM, 100mM NaCl, 10 mM imidazole pH 7.5) three times before the bound protein fraction was eluted directly using elution buffer (50 mM Tris-HCl, 0.10 mM C_8E_4, 100mM NaCl, 400 mM imidazole pH 7.5). Imidazole was buffer exchanged out of the protein sample using an Amicon Ultracel 30K cut-off concentrator (Merck Millipore) using the above elution buffer omitting imidazole.

BamL immunisation for antibody production

For antibody production, 150 µg of purified recombinant BamL protein (DDM micelle) was emulsified using Freund's complete adjuvant (Sigma) for the primary immunisation of a rabbit. Subsequent booster immunisations were administered in Freund's incomplete adjuvant (Sigma) and utilised 50 µg of purified recombinant BamL protein (DDM micelle). An equal volume of adjuvant and purified recombinant antigen (i.e. BamL) were mixed into an emulsion using a stopcock (DISCOFIX®) and 1 mL syringes. New Zealand White rabbits free of leptospiral antibodies were injected subcutaneously on day 1 and as boosts to the immune response, immunisation was performed again on day 21, day 42 and day 63. An antiserum was retrieved prior to the day 1 injection (pre-bleed), and thereafter on day 28 (test bleed 01) and day 49 (test bleed 02) in 1 mL aliquots. The third and final bleed (test bleed 03) was performed on day 65. The BamL-antisera were stored at −20 °C until required. Animals were housed in accordance with the ethical principles and experimental procedures with animals were approved by the Animal Care and Use Committee of the Faculty of Medicine, Nursing and Health Sciences, Monash University (Ethics approval MARP/2016/147, Chief Investigator Dr. Steve Comber)

X-ray crystallography trials

For crystallisation, purified recombinant BamL (DDM or C_8E_4 micelles) was concentrated to 5 mg/mL using an Amicon Ultracel 30K cut-off concentrator (Merck Millipore) without precipitation. High-throughput crystallisation experiments were performed at 20 °C and setup as sitting-drops by the Mosquito® (TTP LabTech Ltd). Commercial crystallisation screens selected for use were PACT *premier*™ and JCSG-*plus*™ (Molecular Dimensions). Crystallisation drops were setup in a Swissci (MRC) 96-well 2-drop UVP plates at 200 nL volume (100 nL BamL + 100 nL reservoir condition) with a 50 μL reservoir volume. Crystallisation drops were monitored and imaged automatically over a 90-day period.

2.3 Cell based techniques

Automated growth curves

In order to examine bacterial growth, bacterial strains were grown overnight in LB broth (M9 minimal media, or DMEM) with orbital shaking at 200 rpm, diluted to a starting optical density at 600 nm (OD600) of 0.05 into fresh LB broth (M9 minimal media, or DMEM), placed into 96-well plates (Falcon-BD), and shaken at 200 rpm for 24 hours. Values for the OD600 of each sample were recorded at regular intervals using the Tecan Spark™ 10M, and all strains were assayed in three independent biological replicates in technical triplicates.

Construction of gfp reporter fusions and GFP fluorescence measurements

DNA fragments 500bp directly upstream of the translational start site of *bamA* or *bamK* were amplified from *K. pneumoniae* B5055 genomic DNA and were cloned into reporter plasmid pPROBE′-gfp[tagless] digested with BamHI and EcoRI. The vector is a broad-host range reporter plasmid with a "promoterless" open-reading frame encoding the green fluorescent protein (GFP; Miller et al., 2000). To modify the *bamK* promoter segment ("non-palindrome" mutant), nucleotide gene blocks (gBlocks, from IDT) were synthesised and cloned into pPROBE′-gfp[tagless] as above. Resultant gfp reporter fusion plasmids were transformed into *K. pneumoniae* B5055 via electroporation.

Overnight cultures of the reporter strains were diluted to a starting OD600 of 0.05 into fresh LB broth and incubated for 3 h at 37 °C on an orbital incubator shaker (180 rpm). From the mid-logarithmic reporter strain cultures, 1 mL aliquots were then harvested (15000 × g, 10 minutes, room temperature) and washed three with PBS. Washed cells were then resuspended in PBS back

to 1 mL and GFP fluorescence emission was measured using filter set at 488 nm excitation and 512 nm using a Spark™ 10M multimode plate reader (Tecan). All GFP measurements were performed in triplicate and were normalised to its cell density (OD_{600}).

Mouse infection and analysis

Mice (6-week-old, male BALB/c) were infected by intranasal delivery with $8\text{-}9 \times 10^3$ CFU of B5055 wildtype or the isogenic Δ*bamK* mutant strain. After 4, 24, 72 and 96 hours post- infection, mice were euthanised and bacterial counts from lungs and liver determined. All animal experiments were conducted at The University of Melbourne, approved by The University of Melbourne Animal Ethics Committee, and were conducted in accordance with the Prevention of Cruelty to Animals Act (1986) and the Australian National Health and Medical Research Council Code of Practice for the Care and Use of Animals for Scientific Purposes (1997).

Macrophage uptake assays and live cell imaging

Macrophages were isolated from the femurs and tibias of 6-8 weeks old C57BL/6 mice. Bone marrow were flushed from the bones using a 26G x 1/2 in needle and 10 mL of macrophages medium (RPMI-1640 media supplemented with 10% FBS, 15% L-cell supernatant, 25 mM HEPES and 100 U/mL Penicillin/Streptomycin (Sigma)). Cells were then placed in a T25 flask overnight at 37°C, 5 % CO_2, allowing adherent cells to be separated from non-adherent monocytes. Monocytes were then transferred to petri dishes and differentiated in macrophages medium for 7-14 days, at 37 °C, 5 % CO_2. For culture periods longer than 7 days, the medium was replaced after one week. Macrophages were gently scraped from petri dishes using a cell scraper and were seeded in tissue culture-treated 24-well plates at a density of 2.5×10^5 cells/well.

For infections with *K. pneumoniae*, bacterial cultures were pelleted ($15000 \times g$, 15 minutes, room temperature) and washed with ×1 PBS (CSH protocols. Bacteria were then resuspended in macrophages medium without penicillin/streptomycin and added to the cells at a multiplicity of infection (MOI) of 25 or 50. The infected cells were then centrifuged ($300 \times g$, 5 minutes, room temperature) and incubated at 37 °C for 2 hours. Macrophages were then washed thrice with PBS. After washing, macrophages were incubated with macrophages medium containing 200 µg/mL gentamicin at 37 °C for 40 min to prevent extracellular bacterial growth. Finally, macrophages were replaced with macrophages medium, and infections proceeded at 37°C for 2 or 6 h. For all experiments, control cells were mock-infected with PBS.

To harvest bacteria from cells, culture medium was removed, and macrophages were lysed in 200 μL of PBS containing 1% (w/v) Saponin at room temperature for 10 minutes. Serial dilutions (10^{-1} to 10^{-5}) of the cell lysates were then plated drop onto LB agar plates and colonies forming units (CFU) were enumerated after incubation at 37°C for 12-18 hours.

Live cell imaging was used to follow the cell death events of *Klebsiella*-infected macrophages in real-time. Infected macrophages were replaced with culture medium containing 600 nM Draq7 (Abcam) after gentamicin incubation. Experiments were performed on a Leica AF6000 LX epi-fluorescence microscope equipped with an incubator chamber set at 37 °C, 5 % CO_2 and an inverted, fully-motorised stage driven by Leica Advanced Suite Application software. Time-lapse images were acquired with bright-field and Y5 filters every 30 minutes for up to 24 hours using a 10×/0.8-A objective. The images were analysed in ImageJ software.

Sequence analysis

DNA regions were scanned for promoter elements using pattern locator algorithm PATLOC (Mŕazek & Xie, 2006). Calculations for predicting DNA curvature of DNA sequences utilisied the bend.it® online bioinformatics service (Vlahoviček *et al.*, 2003). Calculations for predicting the folding and hybridisation of nucleic acid sequences utilised the Mfold web server (Zuker, 2003).

To determine the primary sequence of mature protein following signal peptide cleavage, analysis of the preprotein by SignalP version 4.1 was performed (Petersen *et al.*, 2011). To predict the number and orientation of IM transmembrane α-helices, TMHMM server version 2.0 was used (Krogh *et al.*, 2001).

ExPASy ProtParam (Gasteiger *et al.*, 2003) was used to estimate molecular weight and calculate the distribution of amino acids or atoms within a protein. Multiple sequence alignments were performed using the MUSCLE algorithm (Edgar, 2004).

Protein structures were predicted using Phyre2 (Kelley *et al.*, 2015, Kelley & Sternberg, 2009) or utilised actual structures deposited into the RCSB protein data bank (PDB) (http://www.rcsb.org). The models were further evaluated using I-TASSER (Iterative Threading Assembly Refinement) (Roy *et al.*, 2010, Zhang, 2008) to provide validation to the predicted Phyre2 generated structures. Rendering of protein images to PNG files, including superimpositions, of the models and crystal structures were generated using PyMOL™ Open-Source version 1.8.2.1 (Schrödinger).

Figures

Unless otherwise stated, all figures were prepared using Adobe Illustrator (version 2018.22.1) and where required, images were cropped or re-sized using Adobe Photoshop (version 2018.19.1.5).

Table 2.0.1.: Strains used in this study.

Strains	Genotype or description	Reference/Source
E. coli BL21 Star™ (DE3)	F⁻ *ompT hsdS*$_B$ (r$_B$⁻m$_B$⁻) *gal dcm rne* λ(DE3)	Invitrogen
E. coli DH5α	F- *end*A1 *hsdR*17 (r$_k$2⁻,m$_k$⁺) *sup*E44 thi-1 λ⁻ *recA*1 *gyrA*96 *rel*A1 deoR Δ(*lacZYA-argF*)-U169 Φ80*dlac*ZΔM15; Nal^R	Invitrogen
E. coli S17 *λpir*	*recA, thi, pro, hsdR-M*⁺RP4::2-Tc::Mu::Km Tn7 λpir, Tp^R Sm^R	Richard A. Strugnell, University of Melbourne
E. coli BL21 Star™ (DE3) Δ*bamA*::Kp*bamA*	Native *bamA* gene functionally replaced with B5055 *bamA*	This study
E. coli BL21 Star™ (DE3) Δ*bamA*::Kp*bamK*	Native *bamA* gene functionally replaced with B5055 *bamK*	This study
K. pneumoniae AJ218	K54:O? serotype, hospital isolate, Amp^R	Richard A. Strugnell, University of Melbourne
K. pneumoniae AJ218 Δ*hns*	AJ218 mutant, *hns* gene deleted	Richard A. Strugnell, University of Melbourne
K. pneumoniae MGH78578	K52:O? serotype, laboratory reference strain	Richard A. Strugnell, University of Melbourne
K. pneumoniae B5055	K2:O1, mouse lethal clinical isolate, laboratory reference strain	Richard A. Strugnell, University of Melbourne
K. pneumoniae B5055^Rif	B5055 spontaneous rifampicin mutant, isolated by gradient plate method (Carsenti-Etesse *et al.*, 1999)	This study
K. pneumoniae B5055^Rif Δ*bamK*::Amp^R	B5055 spontaneous rifampicin mutant, *bamK*gene replaced with Ampicillin resistance gene from pGEM-T easy	This study
K. pneumoniae B5055 Δ*bamK*	B5055 mutant, *bamK* gene deleted	Abigail Clements, Imperial College London
K. pneumoniae B5055 Δ*bamA*::Kp*bamK*	Native *bamA* gene functionally replaced with B5055 *bamK*	This study
K. pneumoniae B5055 Δ*dksA*::Kan^R	B5055 mutant, *dskA* gene deleted	This study
K. pneumoniae B5055 Δ*hupA*::Kan^R	B5055 mutant, *hupA* gene deleted	This study
K. pneumoniae B5055 Δ*lupV-bamL*::Kan^R	B5055 mutant, *lupV-bamL* genes deleted	This study

Table 2.0.2.: Plasmids and synthesised DNA fragments (gBlock™) used in this study.

Plasmids or gBlock	Genotype or description	Reference/Source
Plasmids		
pET-15b	Empty vector, IPTG inducible, Amp^r	Novagen
pKS07 (pET-15b)	*phoE* from from *E. coli* K-12 strain MG1655 cloned into MCS	Stubenrauch *et al.* (2016)
pETDuet-1	Empty vector, IPTG inducible, Amp^r	Novagen
pCJS42 (pETDuet-1)	*lptD* from from *E. coli* K-12 strain MG1655 cloned into MCS1	Stubenrauch et al. unpublished
pCJS43 (pETDuet-1)	*lptE* from from *E. coli* K-12 strain MG1655 cloned into MCS2	Stubenrauch et al. unpublished
pCJS44 (pETDuet-1)	*lptD* and *lptE* from from *E. coli* K-12 strain MG1655 cloned into MCS1 and MCS2, respectively.	Stubenrauch et al. unpublished
pACBSR	Ara promoter controlling I-*Sce*I & λ Red	Herring *et al.* (2003)
pKD4	FRT-*kan*-FRT *oriR6K* Amp^r	Datsenko and Wanner (2000)
pCP20	cI857 λPR *flp* pSC101 *oriTS* Amp^r Cam^r)	Cherepanov and Wackernagel (1995)
pBAD24	*ori* pMB1, Amp^r	Guzman *et al.* (1995)
pBAD24-bcLV	Duet expression of B5055 Omp85 genes *lupV* and *bamL* under arabinose control	This study
pET-20b (+) modified	N-terminal modifications: pelB signal sequence, a His$_{10}$-tag and a TEV cleavage site. IPTG inducible, Amp^R	Susan Buchanan, NIH, USA
pET-20b (+) modified-BamL	pET-20b (+) modified that expresses B5055 BamL without its native signal peptide	This study
pKpBK-KpBA donor_Kp	Donor plasmid for replacing B5055 native *bamA* with *bamK*	This study
pKpBA-EcBA donor_Kp	Donor plasmid for replacing B5055 native *bamA* with B5055 *bamA*	This study
pKpBK-EcBA donor_Kp	Donor plasmid for replacing BL21 Star native *bamA* with B5055 *bamK*	This study
pAmp-KpBK donor_Kp	Donor plasmid for replacing B5055 native *bamK* with ampicillin resistance from pGEM-T easy	This study
pKan-dksA donor_Kp	Donor plasmid for replacing B5055 *dksA* with kanamycin cassette from pKD4	This study
pKan-hupA donor_Kp	Donor plasmid for replacing B5055 *hupA* with kanamycin cassette from pKD4	This study
pKan-bamLALL donor_Kp	Donor plasmid for replacing B5055 *bamL* and *lupV* with kanamycin cassette from pKD4	This study
pGEM-T Easy	Cloning vector Amp^R	Promega
pJET (1.2)	Cloning vector Amp^R Blunt end cloning for gBlocks™	Thermo Scientific
pUA139	Reporter plasmid promoter, Cam^R	Zaslaver *et al.* (2006)
pUA139[bamA]	pUA139 with *bamA* promoter from B5055 in front of *gfp* reporter CDS	This study
pUA139[bamK]	pUA139 with *bamK* promoter from B5055 in front of *gfp* reporter CDS	This study
pPROBE′-gfp[tagless]	Broad-host range reporter plasmid promoter, Kan^R	Miller *et al.* (2000), Addgene
pPROBE'-gfp[bamA]	pPROBE′-gfp[tagless] with *bamA* promoter from B5055 in front of *gfp* CDS	This study
pPROBE'-gfp[bamK]	pPROBE′-gfp[tagless] with *bamK* promoter from B5055 in front of *gfp* CDS	This study
pPROBE'-gfp[tamA]	pPROBE′-gfp[tagless] with *tamA* promoter from B5055 in front of *gfp* CDS	This study
pPROBE'-gfp[bamL]	pPROBE′-gfp[tagless] with *bamL* promoter from B5055 in front of *gfp* CDS	This study
pPROBE'-gfp [bamK non-palindrome]	pPROBE′-gfp[tagless] with *bamK* promoter from B5055 containing a disrupted palindrome sequence in front of *gfp* CDS (see Figure 3.2.2)	This study
pBR:*BamK + prom #15*	B5055 *bamK* with 500 bp putative promoter region in pBR322	Abigail Clements, Imperial College London
gBlock™		
bamK non-palindrome	Contains 500 bp upstream of B5055 *bamK* CDS with disrupted palindrome (see Figure 3.2.2),	Integrated DNA technologies (IDT)

	cloned into pJET then sub-cloned into pPROBE′-gfp[tagless]	
LupV_BamL_polycis	Contains CDS for l*upV* and *bamL* from B5055 with native signal sequences for localisation and biochemical studies, cloned into pJET then sub-cloned into pBAD24 (see Figure 4.5.1)	Integrated DNA technologies (IDT)

Table 2.0.3.: Oligonucleotides used in this study.

Function	Primer	Sequence (5'→3')
gene knock-in:		
***bamA* replaced by B5055 *bamK*, in B5055**	VVT029	tagggataacagggtaataggtaaggcggaagggttcg
	VVT022	ctgatgatatgagtcttttttaacatcgttattatgcgttcttcc
	VVT023	atgttaaaaaagactcatatcatcagc
	VVT024	gaagcagctccagcctacacactaccaggttttgccgatattaaac
	VVT030	ctaaggaggatattcatatgttgttggcggcaacggaatg
	VVT031	tagggataacagggtaatcgctctgtacggcagtctggatac
***bamA* replaced by B5055 *bamA*, in BL21 Star™ (DE3)**	VVT150	cgttattatgcgttcttcctaacta
	KpBamA-F	tagttaggaagaacgcataataacgatggcgatgaaaaagttgctcatagc
	KpBamA-R	gaagcagctccagcctacacattaccaggttttaccaatgttaaactggaac
***bamA* replaced by B5055 *bamK*, in BL21 Star™ (DE3)**	VVT021	tagggataacagggtaataatggtaaagcgattggttttg
	VVT027	ctaaggaggatattcatatggtgttctccacaaaggaatgtagtgg
	VVT028	tagggataacagggtaatcggatttcacagcagtctggatacg
***bamK* replaced by ampicillin resistance gene, in B5055**	VVT096	tagggataacagggtaataggcgaccctactgtattctgc
	VVT097	gggcgacacggaaatgttgaatactcatgaatgtgtcgtctacgagataatgagcaatag
	VVT098	atgagtattcaacatttccgtgtcgccc
	VVT099	gaagcagctccagcctacacattaccaatgcttaatcagtgaggcacc
	VVT100	ctaaggaggatattcatatgggctccaacggggcctgatag
	VVT101	tagggataacagggtaatttcattgttataaatcgcaagcacccctg
gene knock-outs:		
***dksA* knock-out**	VVT120	tagggataacagggtaatatgaccgacaaaactgctatattg
	VVT121	Gaagcagctccagcctacacagttgcttctccttaacacgcactatc
	VVT122	ctaaggaggatattcatatgtcccctgctgcctcactctctggcg
	VVT123	tagggataacagggtaatgcgcagacggtcgctaaactgcagcaccg
***hupA* knock-out**	VVT200	tagggataacagggtaatccatagacgcttgcgtggcgttga
	VVT201	gaagcagctccagcctacacaaagttatccttacaatgtgtttatcgc
	VVT202	ctaaggaggatattcatatggaccgcgtggcagtgaacagtttta
	VVT203	tagggataacagggtaatttccagcgcccctgccgcaggacga
***lupV-bamL* locus knock-out**	VVT204	tagggataacagggtaatatgatgattgatgagccggcaatcc
	VVT205	gaagcagctccagcctacacagaaaagtatgatgttagttctgagt
	VVT206	ctaaggaggatattcatatgcgcctcctccccttaattttctctc
	VVT207	tagggataacagggtaatcatagtcattttccagtacccactg
Sequencing primers for genomic integration:		
***bamK* replaced by ampicillin resistance gene, in B5055**	VVT102	ttagaaccatgcgcttcagaaaaagttcc
	VVT103	gaagagaggaccatagcaggcagctattta
***hupA* knock-out**	VVT094	gaggctggcgagagcaatattggtataa
	VVT095	cattgagttcgataagcgccagttgat
***dksA* knock-out**	VVT092	ggaactttctacgacgaggatgactctga
	VVT093	tcaagctggcgcaggatggtat
***lupV-bamL* locus knock-out**	VVT086	gaaccttgaagatggcgttgcc
	VVT087	gagtccttgtgctctatgtgtactgc
***bamA* replaced by B5055 *bamK*, in B5055**	VVT063	ccagttagagattcggagtccgttctg
	VVT064	ggacgcaacgccggtg
***bamA* replaced by B5055 *bamA* or *bamK*, in BL21 Star™ (DE3)**	VVT148	gccgcaaattgaacctgtactggaaa
	VVT149	tgtgacctgtttgtgcagattgcat
Promoter regions for gfp reporter pProbe, unless otherwise stated:		
bamA	VVT007	gcggcgaattcaggtaaggcggaagggttcg
	VVT008	ccgcgggatcccgttattatgcgttcttcctaacaaactctcttataacc

bamK	VVT003	ccggcggtaccaggcgaccctactgtattctgc
	VVT004	gcgcgggatccgaatgtgtcgtctacgagataatgagcaatagc
tamA	VVT015	gcgggaattcagtcacgccggatcgtcg
	VVT016	ccgcgggatccggtcgccctctttaaacggcac
bamL	VVT011	gccgcgaattcgggatgcaggatctcgactttcgtctc
	VVT012	cggcggatccaagcgatactcgccggcaataaag
bamA (pUA139)	VVT005	gccgggatccaggtaaggcggaagggttcg
	VVT006	ccggcctcgagcgttattatgcgttcttcctaacaaactctcttataacc
bamK (pUA139)	VVT001	gccgggatccaggcgaccctactgtattctgc
	VVT002	ggcggctcgaggaatgtgtcgtctacgagataatgagcaatagc
EMSA PCR amplicons from putative *bamK* promoter:		
Fragment 1(220 bp)	VVT069	ttagaaccatgcgcttcagaaaaagttc
	VVT072	catgaatgtgtcgtctacgagataatgagc
Fragment 2(140 bp)	VVT069	ttagaaccatgcgcttcagaaaaagttc
	VVT071	aaaatatgcgacagaaaataacctgagcg
Fragment 3 (83 bp)	VVT069	ttagaaccatgcgcttcagaaaaagttc
	VVT070	caaaatggagaaaaaatggatgttaaagcatctt
Cloning for BamL protein expression		
pET-20b (+) modified - bamL	VVT059	gcgcgtcatgattgaggcgctcagccgcg
	VVT050	gcggccgagctcttaaaacgcttcgccgacctgg
Kanamycin cassette gene from pKD4:		
	VVT025	gtgtaggctggagctgcttc
	VVT026	catatgaatatcctccttag
pGEM T-easy backbone for gene doctoring:		
	VVT019	attaccctgttatccctagtcatagctgtttcctgtgtg
	VVT053	attaccctgttatccctaactggccgtcgtttacaacg
qRT-PCR:		
bamA	KpB5_01_qPCR	cccgacgccgtttattagt
	KpB5_02_qPCR	ttttgaacggttgggcgtag
bamK	KpB5_03_qPCR	gtcgtgaaatgcgccaaatg
	KpB5_04_qPCR	ccaggcccacattgaatgag
tamA	KpB5_05_qPCR	tttatcctggggtctcggtg
	KpB5_06_qPCR	accttgtcaaagttgtcgc
bamL	KpB5_07_qPCR	gacgatcggccgaagataga
	KpB5_08_qPCR	tatcgctacctaatccgggc
dnaG	KpB5_09_qPCR	gatctgagtactccggacgg
	KpB5_10_qPCR	gccattttccgcctgttttg
Glk	KpB5_11_qPCR	ttaccgctgtttcaatggcc
	KpB5_12_qPCR	aatcaacatgaccgccttcg
gyrA	KpB5_13_qPCR	tatcaggtgaacaaagcgcg
	KpB5_14_qPCR	gagctaccatgttgatgccg
rpoD	KpB5_15_qPCR	caatgaccatctgccggaag
	KpB5_16_qPCR	cttcggcagcatcttcatcc

Table 2.0.4.: Antibodies and conjugates used in this study.

Antibody	Type	Dilution used	Reference/Source
Mouse αBamK (MK1)	Primary	1:100	This study
Mouse αBamK (MK2)	Primary	1:100	This study
Mouse αBamK (MK3)	Primary	1:100	This study
Mouse αBamK (MK4)	Primary	1:100	This study
Mouse αBamK (MK5)	Primary	1:100	This study
Rabbbit αBamL (TB1-TB3)	Primary	1:10000	This study
Rabbit αF1β (Mitochondrial)		1:5000	In-house, Lithgow Laboratory
Mouse αBamA	Primary	1:20000	Susan Buchanan, NIH, USA
Mouse αBamB	Primary	1:40000	Susan Buchanan, NIH, USA
Mouse αBamC	Primary	1:20000	Susan Buchanan, NIH, USA
Mouse αBamD	Primary	1:20000	Susan Buchanan, NIH, USA
Mouse αBamE	Primary	1:5000	Susan Buchanan, NIH, USA
Rabbit αBamB	Primary	1:20000	In-house, Lithgow Laboratory
Rabbit αBamC	Primary	1:20000	In-house, Lithgow Laboratory
Rabbit αBamD	Primary	1:20000	In-house, Lithgow Laboratory
Rabbit αBamE	Primary	1:20000	In-house, Lithgow Laboratory
PrecisionProtein™ StrepTactin	Primary (HRP Conjugate)	1:5000	Bio-rad
HisProbe™	Primary (HRP Conjugate)	1:5000	Thermo Scientific
Goat αMouse-HRP	Secondary	1:20000	Sigma (A4416)
Goat αRabbit-HRP	Secondary	1:20000	Sigma (A6154)

3

Results: Genetic expression of Omp85 family genes in _K. pneumoniae_

3.0 Introduction

Like many Gram-negative bacteria, the outer membrane of _K. pneumoniae_ is crucial as it provides the first line of defence against a range of host innate defensive mechanisms, antimicrobials and reactive oxygen species but also plays a critical role in host–pathogen interactions (Nikaido, 2003, Nikaido, 1989, Kuehn & Kesty, 2005). Associated with these protective and virulence-associated functions, are the β-barrel OMPs embedded within this asymmetric lipid bilayer, which are almost exclusively assembled and inserted by the β-barrel assembly machinery (BAM) complex (Knowles _et al._, 2009, Silhavy & Malinverni, 2011). The core protein of the BAM complex is essential OMP BamA, a member of the Omp85 protein family (Hagan _et al._, 2011). Members of the Omp85 protein family function as protein translocases or as outer membrane protein assembly factors and are defined by unique shared sequence motifs (Mazar & Cotter, 2006, Hagan _et al._, 2011). An interesting observation found in many bacterial species is that whilst BamA is essential for the assembly of β-barrel OMPs, additional orthologous copies are also maintained within the same genome of some species most with unknown functional roles. One such example can be found in _K. pneumoniae_ B5055, which harbours 4 distinct _omp85_ family members: _bamA_, _tamA_, _bamK,_ and _bamL._

Gene expression of bacterial OMPs are differentially regulated by environmental signals (e.g. the abundance or absence of specific nutrients/cofactors), enabling these bacteria to adapt to a range of different hosts and environments (Guilhen _et al._, 2016, Li _et al._, 2015a, Van Laar _et al._, 2015). A gap in knowledge exists for the differential gene expression and regulatory networks which control Omp85 expression in bacterial species that have varied copy-numbers of these genes. Are certain Omp85s expressed only during certain environmental conditions? What transcriptional/genetic factors are regulating their expression? In this chapter, I present results investigating the gene regulation and expression of these genes, with a particular focus on _bamK_ expression, to better understand how they contribute to outer membrane biogenesis.

3.1 Investigating gene expression and promoter activity of Omp85 family genes in *K. pneumoniae* B5055

To determine the general level of gene expression of Omp85 family genes in *K. pneumoniae* B5055, I performed quantitative PCR (qPCR). The qPCR technique is a method for assaying gene expression under a different growth conditions and is favoured because of its simplicity, reproducibility and accuracy (Thellin *et al.*, 2009). For this study, relative transcripts of target genes were normalised to a reference gene illustrating its fold change under different conditions. The genes used as references are often referred to as "housekeeping genes" and are assumed to be constitutively expressed under the selected growth conditions or treatments. However, from a study by Rocha *et al.* (2015), expression levels of these so called "housekeeping genes" varied in different bacterial species and certain growth conditions. Therefore, for gene expression analyses of the *omp85s* in *K. pneumoniae* B5055, I first evaluated the expression levels of a selection of candidate reference genes for their reproducibility and level variations in the intended growth conditions and culture media. Shown in Figure 3.1.1, expression profiles of the candidate reference genes were investigated at logarithmic ($OD_{600} = {\sim}0.6$) and stationary growth phases ($OD_{600} = {\sim}1.6$) cultured in LB media. The genes *gyrA* and *rpoD* were most consistent in transcript abundance within the same growth phase but were expressed at differential levels when comparing one growth phase to another. Of the four candidate references, *rpoD* was selected for gene expression normalisation in downstream studies based on its invariability expression pattern in the tested conditions Additional validation qPCR assays can be found in Appendix 1 for qPCR assays performed during this study.
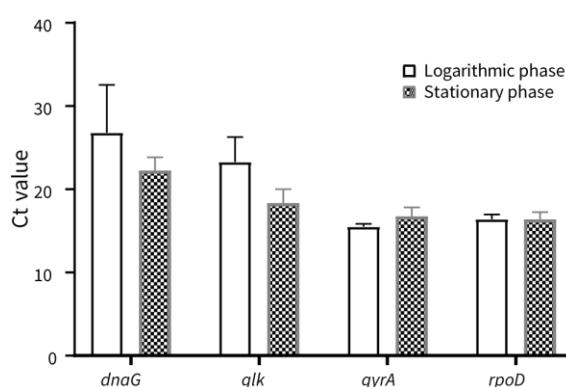
Figure 3.1.1: Comparison of candidate reference gene expression in *K. pneumoniae* B5055 cultured in LB media. Transcript levels for Omp85 genes in *K. pneumoniae* B5055 were determined through qPCR at logarithmic and stationary growth phase grown in rich LB media. The expression levels of candidate reference (*dnaG, glk, gyrA* and *rpoD*) genes were determined using cycle threshold (Ct) values through real-time qPCR. A "Ct value" represents the amplification cycle number at which the fluorescence signal of the reporter dye reaches above an arbitrarily placed baseline threshold (Schmittgen & Livak, 2008). Thus, a lower Ct value corresponds to a higher abundance of a transcript, whereas a higher Ct value corresponds to a lower abundance of transcript. (n=3, error bars represent SD).

After selecting *rpoD* as the preferential reference for gene expression studies in *K. pneumoniae* B5055, I determined the relative expression levels of the genes encoding Omp85 family members under standard planktonic growth conditions. Figure 3.1.2 shows that under planktonic growth, *bamA* was the highest in relative expression of the four Omp85 genes in both mid-logarithmic and stationary growth phases when grown in nutrient rich LB media. Furthermore, when shifting from logarithmic to stationary growth phase, the relative expression of *bamA* decreased but *tamA* had a notable increase. Relative expression of *bamK* and *bamL* transcripts were very low and growth phase did not seem to modulate these levels.
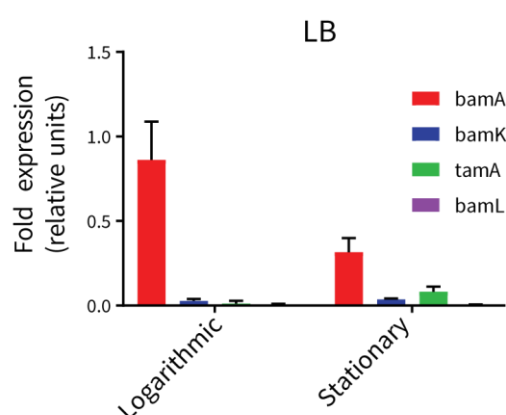
Figure 3.1.2: Transcript levels of Omp85 genes at different growth phases in *K. pneumoniae* B5055. Transcript levels for Omp85 genes in *K. pneumoniae* B5055 were determined through qPCR at logarithmic and stationary growth phase grown in rich LB media. Data is expressed relative to control transcripts from the constitutively expressed housekeeping gene: *rpoD*. (n=3, error bars represent SD).

In a parallel assay, qPCR was also performed on *K. pneumoniae* B5055 cultured in DMEM, which has been used to mimic "in-host" growth conditions and macrophage development (Trouplin *et al.*, 2013, Kawakami *et al.*, 2016). It was again observed relative expression levels of the Omp85 genes mirrored that of the LB planktonic growth conditions (Figure 3.1.3). For all qPCR assays, transcript levels of each gene were normalised to a reference gene selected by its mRNA stability in the tested conditions. For relative gene expression calculations, *rpoD* was selected as the preferential reference based on similar evaluation experiments as described in Figure 3.1.1 (Appendix 1, B).
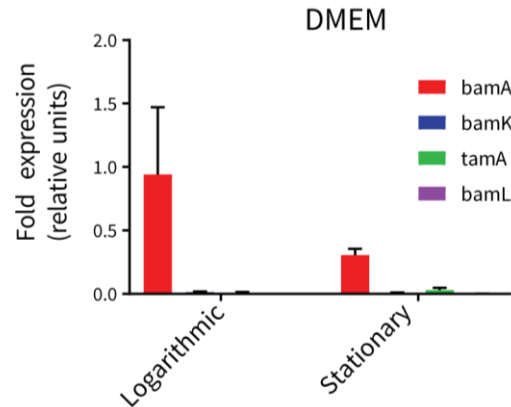
Figure 3.1.3: Omp85 transcript levels at different growth phases in *K. pneumoniae* B5055. Transcript levels for Omp85 genes in *K. pneumoniae* B5055 were determined through qPCR at logarithmic and stationary growth phase grown in "host-like" culture conditions by growth in DMEM. Data is expressed relative to control transcripts from the constituently expressed housekeeping gene -*rpoD*. (n=3, error bars represent SD)

Genes of the BAM complex (*bamA* through *bamE*) in *E. coli* (Dartigalongue *et al.*, 2001, Onufryk *et al.*, 2005), and *Salmonella enterica* subsp. *enterica* serovar Typhimurium (Li *et al.*, 2015b) are known to be part of the σ^E regulon, which maintains outer membrane homeostasis. The σ^E-regulon is an adaptive response, which can be triggered by growth at non-optimal temperatures (Mecsas *et al.*, 1993, Raina *et al.*, 1995). To that end, qPCR assays were performed to examine relative expression levels of the genes of interest in response to temperature "shocks". Log-phase (OD_{600}= ~0.6) *K. pneumoniae* B5055 cultures grown at 37 °C were pelleted and then resuspended in pre-heated media for an immediate temperature "shock" and incubated for 1 hour at temperatures of 30, 37 or 42 °C. The rationale was to minimise the time for cultures to adapt to the change in temperature. Shown in Figure 3.1.4, the transcript levels of essential gene *bamA* was the highest in the three temperature conditions tested, with expression of *bamK, tamA*, and *bamL* found at dramatically lower levels (Figure 3.1.4). For relative gene expression calculations, *rpoD* was elected to be a preferential reference based on evaluation experiments described in Figure 3.1.1 (Appendix 1, C).
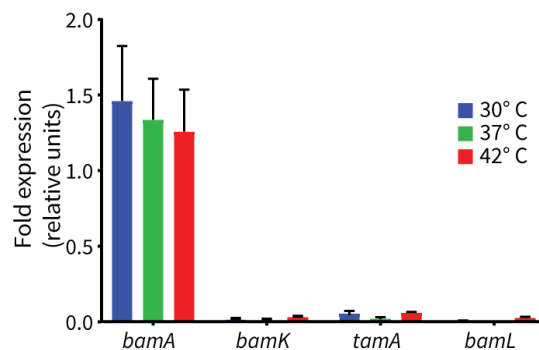
Figure 3.1.4: Relative gene expression of Omp85 genes of *K. pneumoniae* B5055 subjected to different temperatures shocks. Transcript levels for *bamA, bamK, tamA* and *bamL* of *K. pneumoniae* B5055 cultures subjected to different temperatures "shocks". The temperature "shock" was the transfer of mid-log cultures grown at 37 °C subjected to pre-heated media at three different growth temperatures 30 °C, 37 °C or 42 °C. Data is expressed relative to control transcripts from constitutively expressed housekeeping gene *rpoD*. (n=3, error bars represent SD).

As it was clear the levels of *bamK*, *tamA*, and *bamL* remained at or below the limit of detection in each of the test conditions, I interpreted this to mean either the transcript itself is unstable (short half-life) and/or the transcript is not being produced (incorrect inducing conditions). To test whether the transcripts themselves had short half-lives, reporter gene fusions were constructed. Reporter gene fusions are powerful experimental tools that allow effective monitoring of gene expression, by fusing a relatively stable gene, such as the green fluorescent protein (GFP) gene to a promoter of interest (Southward & Surette, 2002, Miller *et al.*, 2000). Each of the four upstream putative promoters were amplified using primers for regions upstream of the corresponding CDS (except for *bamL*, where the primers were designed to amplify the upstream region of *lupV*, which shares the promoter with *bamL*) into broad-host range reporter plasmid pProbe (Figure 3.1.5) (Miller *et al.*, 2000).
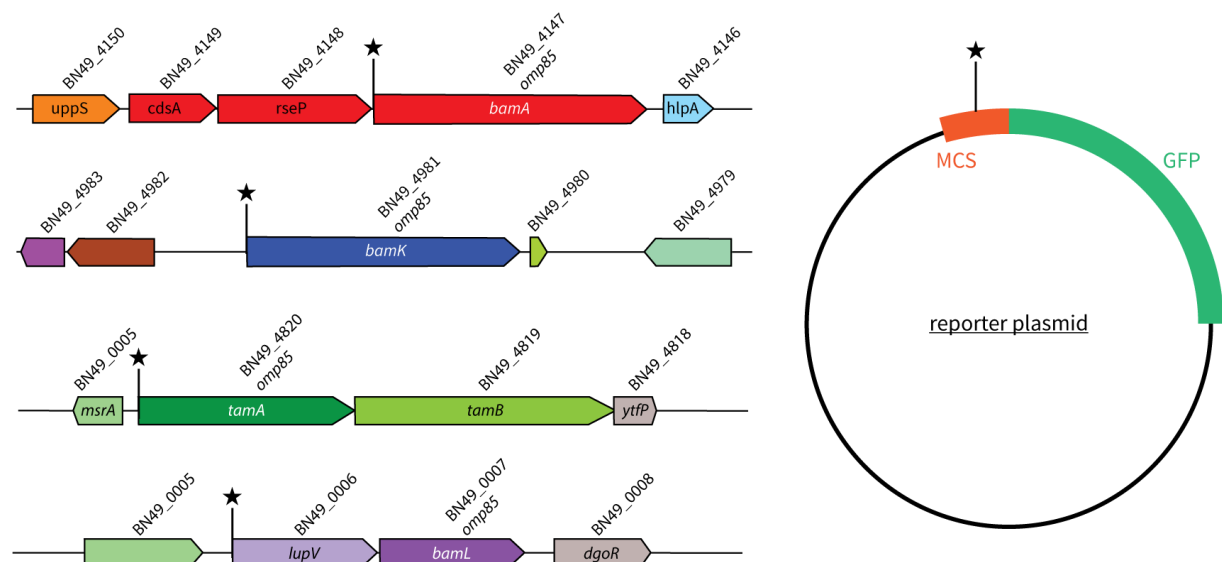
Figure 3.1.5: Schematic outlining the genomic region cloned into GFP reporter plasmids for transcriptional studies. Operon organisation of the four genes (white text) encoding for the Omp85s present in *K. pneumoniae* B5055. The star (★) indicates the 3' hydroxyl end of a 500 bp region (i.e. immediately upstream of Omp85 translation start sites) cloned into the reporter plasmids used in this study (pProbe or pUA139).

Under the identical growth conditions performed for qPCR transcript quantification, promoter activity of *K. pneumoniae* B5055 genes *bamA* and *tamA* were found to be the highest whilst genes *bamK* and *bamL* were below the level of detection. This was observed for both growth phases when cultured in nutrient rich LB media (Figure 3.1.6). Controls included non-transformed (wildtype) *K. pneumoniae* B5055 and *K. pneumoniae* B5055 transformed with the control plasmid.
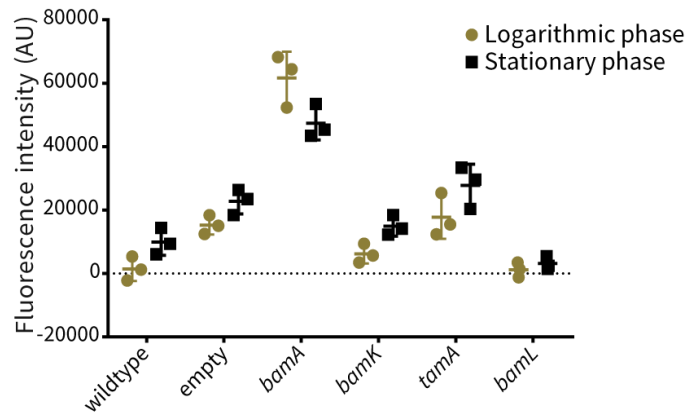
Figure 3.1.6: Promoter activity of Omp85 genes at two different growth phases in *K. pneumoniae* B5055. Analysis of the expression of *bamA*, *bamK*, *tamA*, and *bamL* in *K. pneumoniae* B5055 carrying pProbe transcriptional fusion (*bamA*::*gfp, bamk*::*gfp*, *tamA*:;*gfp*, and *bamL*::*gfp*) reporter plasmids. Bacteria grown in LB were measured for promoter activity at different growth phases plotted as paired colour-coded data sets. Wildtype – B5055 no harboured reporter plasmid. empty – B5055 containing empty vector pProbe (n = 3, where error bars represent SD and horizontal bar represents the mean).

Previous studies in *E. coli* showed that osmoregulation is an intricate process that involves concerted control of membrane associated transcription factors (e.g. OmpR) and OMPs such as the porins, OmpC and OmpF (Yoshida *et al.*, 2006). The assembly of most OMPs in this species requires the BAM complex, and as previously mentioned, *bamA* and the other subunits of the BAM complex (i.e. *bamB-bamE*) are subject to control by the σ^E regulon. *K. pneumoniae* is closely related to *E. coli* and is thought to share many gene regulatory networks (Kim *et al.*, 2012, Seo *et al.*, 2012), I hypothesised that *bamA* and other genes in *K. pneumoniae* B5055 may be osmoregulated. To test this, I cultured *K. pneumoniae* B5055 harbouring Omp85 GFP-promoter fusion plasmids in low-osmolarity and high-osmolarity media. As shown in Figure 3.1.7, the promoter activity of *bamA* was highest followed by *tamA*. For the *Klebsiella* specific Omp85 genes, *bamK* and *bamL,* promoter activity was very low and mirrored the previous set of qPCR and GFP-fusion assay results suggesting the promoters are tightly regulated. Altogether, these data suggest the expression and regulation of *bamK* and *bamL* is cryptic and not expressed under the tested standard laboratory conditions and therefore factors could be inhibiting or repressing their expression.
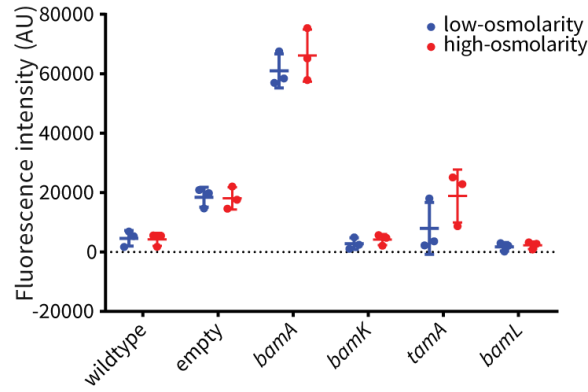
Figure 3.1.7: Promoter activity of Omp85 genes at two different osmotic concentrations in *K. pneumoniae* B5055. Analysis of the expression of *bamA*, *bamK*, *tamA*, and *bamL* in *K. pneumoniae* B5055 carrying pProbe transcriptional fusion (*bamA*::*gfp*, *bamk*::*gfp*, *tamA*::;*gfp*, and *bamL*::*gfp*) reporter plasmids. Mid-logarithmic *K. pneumoniae* B5055 grown in LB supplemented with or without 20% sorbitol (corresponds to high and low osmolarity, respectively). Wildtype – B5055 no reporter plasmid. empty – B5055 containing empty vector pProbe. ($n = 3$, horizontal bar represents the mean). ($n = 3$, where error bars represent SD and horizontal bar represents the mean).

3.2 Promoter analysis of the novel *K. pneumoniae* Omp85, *bamK*

The qPCR and promoter-gene fusions revealed little to no expression of *bamK* using various media, osmolarity, temperature and growth phases. Indeed, the promoter-gene fusion experiments consistently revealed that levels of GFP were equivalent to the empty vector controls. To determine whether the *bamK* promoter was actively being repressed, the *bamK* promoter region was analysed for common regulatory elements. Shown in Figure 3.2.1 are putative promoter elements identified within the 500 bp region upstream from the *bamK* translational start site: (1) The −10 and −35 elements characteristic of σ^{70} promoter binding site were identified, an overlapping (2) stem loop structure and a predicted (3) ribosomal binding site (RBS). While the RBS was identified within 6 bp of the *bamK* start codon, the more distal "−35 and −10"-like promoter elements were found 173 and 197 bp upstream, respectively.

```
-500  AGGCGACCCTACTGTATTCTGCTATTTTGAGGAAAATATGGAGATGTTTG
-450  CCAACGGCGTCTGGCGGTGAATTCGGGGTAGAATGTTGTGATATCAGTTG
-400  CATAATTTATTCAATTTCTCTTCATTAAATCTGCTGGGTTGCAGGTTATT
-350  TTTCTCCCTGTTCCTGGTGATCCCAGAGAGCATCATTCCTAACGAGATCC
-300  TTCCCCCTTTCCATTAATAAGAAAGGGGGCTTTTTATTAATAAAACGTTC
-250  AGGCAAATCTAAATATAGTACTGGTGTTGGCCATTAGAACCATGCGCTTC
-200  AGAAAAAGTTCCATCATTAAATATATTTTTAAAGATGCTTTAACATCCAT
-150  TTTTTCTCCATTTTGTCGCCATAATTGCATGTAATTCTGCGCCTCGCTCA
-100  GGTTATTTTCTGTCGCATATTTTAGAGATGCATCCTGAACAAATAAGCTC
-50   AACGCCAGCAGGACACTGCTATTGCTCATTATCTCGTAGACGACACATTCATG
```

Figure 3.2.1: Putative transcription elements of 5' region of *bamK*. Sequence (500 bp) upstream of the *bamK* open-reading frame. The putative "−35 and −10" like elements of the predicted σ^{70} promoter is labelled in red, a palindromic sequence overlapping the −10 element is underlined potentially forming a stem loop structure. The predicted ribosome-binding site (RBS) is shown in blue boldfaced and the ATG start codon is highlighted in green. Performed by Dr. Ji Yang (The University of Melbourne, Australia)

The presence of stem-loop structures in prokaryotic genomes are known to be involved in gene regulation at the transcriptional and translational levels (Diwa *et al.*, 2000, Kozak, 2005). In the case of the putative *bamK* promoter region, a palindromic sequence was noted because of its potential to form a stem loop structure, which would overlap the −35 element. Such a structure might repress gene expression. Therefore, to examine if the predicted stem loop structure influences *bamK* expression, I constructed another series of GFP reporter plasmids. This construct (–palindrome) contained a synthetic 500 bp promoter sequence, where the palindromic sequence was disrupted by modifying select nucleotides that would prevent the ability of the stem loop structure to form hydrogen bonds. As shown Figure 3.2.2 A, the proposed stem loop structure nucleotide sequence within the *bamK* upstream sequence, CATCATTAAATATATTTTTAAAGATG, was modified to CATCATTAAATATATT**AAATTACTAC**. This mutant sequence (–palindrome) was introduced into the same reporter-fusion plasmid system. As seen in Figure 3.2.2 B, the promoter activity of *bamK* with the native or modified (– palindrome) stem loop structure sequence did not influence *bamK* gene expression. Together, these data suggest the predicted stem loop structure within the upstream putative promoter region does not play a major role in gene expression of *bamK* under the conditions tested or other sequence determinants may be involved.
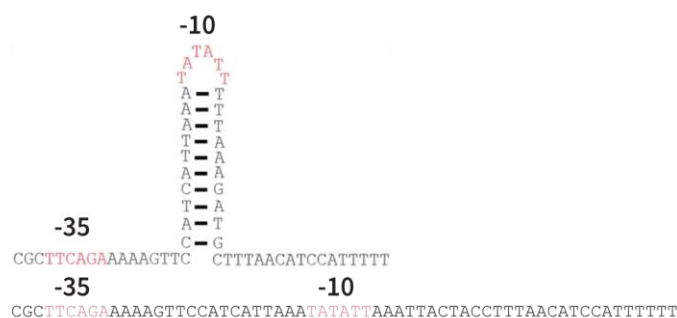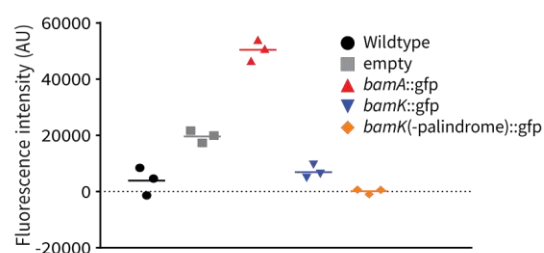
Figure 3.2.2: Disruption of predicted palindrome and its influence on promoter activity of *bamK* in *K. pneumoniae* B5055. (A) Depiction of the predicted palindrome found in the putative *bamK* element which forms a hairpin/stem-loop structure potentially affecting transcriptional regulation. Below, the introduced synthetic sequence aiming to disrupt the hydrogen bonds forming the hairpin structure. The predicted ΔG for the hairpin loop was −3.17 kcal/mol whilst the synthetic sequence was at 0.21 kcal/mol using the Mfold RNA folding prediction server (Zuker, 2003). (B) Analysis of the promoter activity in *K. pneumoniae* B5055 carrying pProbe transcriptional fusions with: (i) *bamA*, a (ii) native or (ii) synthetically altered *bamK* promoter (*bamk*::*gfp*, *bamK-palindrome*::*gfp*) reporter plasmids . All bacteria were grown in LB media to mid-logarithmic phase then measured for promoter activity using the Spark™ 10M multimode reader. Wildtype – B5055 no harboured reporter plasmid. Empty – B5055 containing empty vector pProbe. (n = 3, horizontal bar represents the mean).

Intrinsically curved genomic DNA regions can be caused by properties such as poly(A) tracts (Matsushita *et al.*, 1996, Cheema *et al.*, 1999), and such regions have a propensity to be targeted by silencing proteins, including global transcriptional regulators such as H-NS (Owen-Hughes *et al.*, 1992, Yamada *et al.*, 1991). Therefore, to determine if the putative promoter region of *bamK* included a DNA region with high curvature, the publicly available bend.it® server was employed by inputting the 1000 bp upstream region as the query (Vlahoviček *et al.*, 2003). From the output, a highly AT-rich region upstream of the *bamK* locus was predicted to contribute to a highly curved DNA secondary structure (Figure 3.2.3). Based on the predicted high DNA curvature region, it was hypothesised that global gene regulator H-NS, known to bind highly curved DNA regions, might be repressing *bamK* expression (Figure 3.2.4).
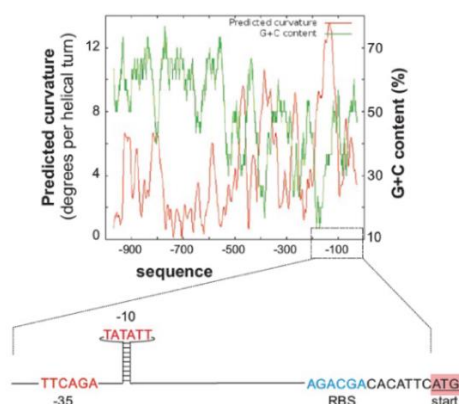
Figure 3.2.3: In silico analysis of intrinsic curvature of the 5' regulatory region of *bamK*. *In silico* analysis of intrinsic curvature of the 5' regulatory region of *bamK* was predicted with bend.it® (Vlahoviček *et al.*, 2003) and regions with >5 degrees per helical turn of DNA (in red) correspond to highly curved regions of DNA. The 200 bp region of the genome immediately upstream of the start start codon of *bamK* is summarised as an inset.
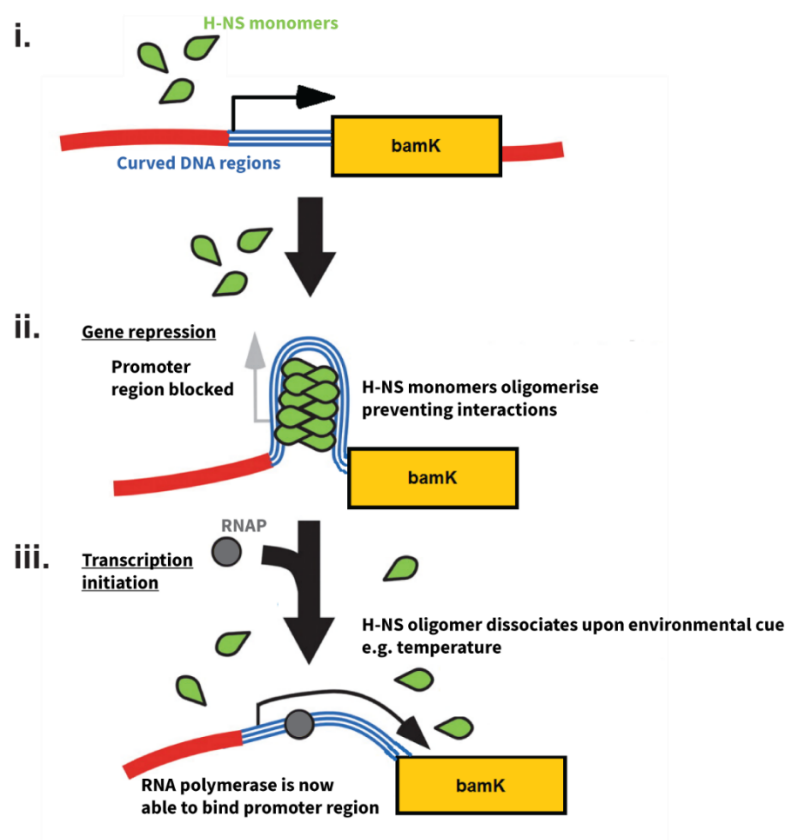
Figure 3.2.4: H-NS schematic showing the hypothetical gene repression of *bamK*. (i) As predicted in Figure 3.2.3, a highly curved region is situated before the translational start site of the *bamK* locus. These DNA regions have been reported previously to bind to global gene regulators like H-NS (Owen-Hughes *et al.*, 1992, Yamada *et al.*, 1991). (ii) H-NS oligomerisation can exclude RNA polymerase (and other co-factors) by obstructing promoter binding elements. (iii) H-NS oligomer dissociates upon environmental cues, such as temperature or metal ion concentrations (Ono *et al.*, 2005, Amit *et al.*, 2003), allowing promoter binding by RNA polymerase (RNAP) for transcription initiation.

This non-coding region is 100 % conserved, suggesting in both *K. pneumoniae* strains B5055 and AJ218 that regulation and binding of any transcription or translation factors would be identical. To investigate if H-NS regulates *bamK* expression, qPCR was employed on *K. pneumoniae* AJ218 Δ*hns* and its isogenic parent, (*K. pneumoniae* AJ218) (strains were obtained from Prof. Richard Strugnell, The University of Melbourne). In Figure 3.2.5, gene expression data revealed that relative transcription levels for *bamK* were unchanged even in the absence of H-NS. These results could suggest in a *hns* mutant strain, the single deletion of the global regulation is not enough to derepress (or activate) transcription. For relative gene expression calculations, *rpoD* was selected as the preferential reference based on similar evaluation experiments as described in Figure 3.1.1 (Appendix 1, D). The putative promoter region of *bamK* (-500bp) is highly conserved in many *K. pneumoniae* species (Appendix 2), and therefore the observed low transcript levels may not be unique to the strains tested (i.e. B5055 and AJ218).
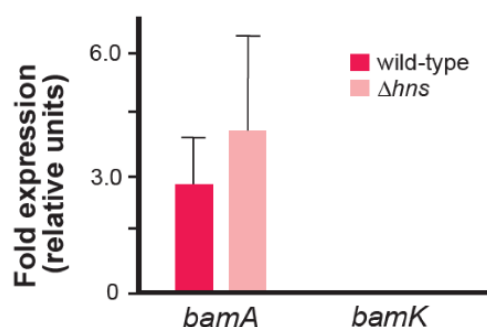
Figure 3.2.5: Gene transcript levels of *bamA* and *bamK* in *K. pneumoniae* AJ218. Relative transcript levels for *bamA* and *bamK* in *K. pneumoniae* AJ218 were determined by qPCR from cells grown to mid-logarithmic phase cultured in rich LB media. Data is expressed relative to control transcripts from the constituently expressed housekeeping gene - *rpoD*. (n=3, error bars represent SD).

To selectively detect expression of BamK and measure protein expression levels, the BamK protein sequence was analysed for a peptide sequences which differed between BamK (KpBamK) and BamA (KpBamA) of *K. pneumoniae* B5055 to be used for monoclonal antibody production. The peptide IALNEGERYRVDRT$_{271}$ located between POTRAs 3 & 4 was selected based on its unique sequence and predicted flexibility and therefore accessibility for immunoblot detection for natively folded proteins. (Figure 3.2.6) (Grant, 2003).
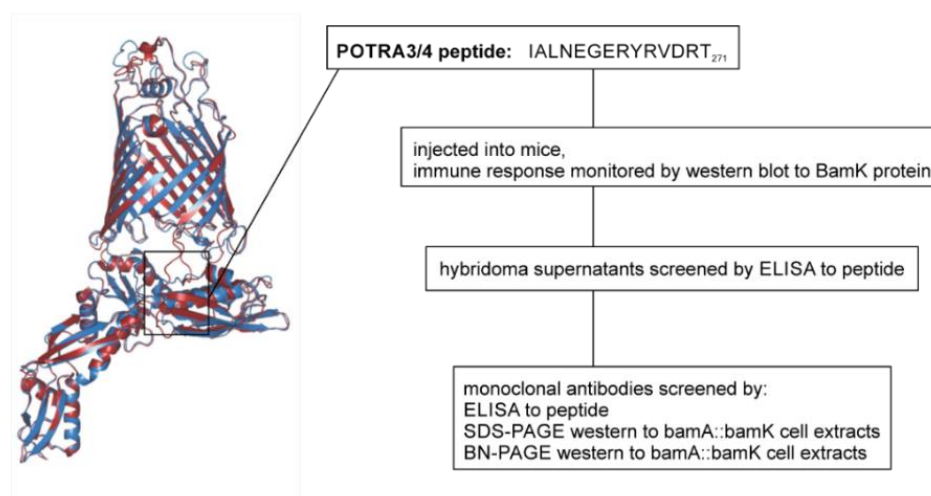
Figure 3.2.6: Flow chart for generation and identification of a suitable monoclonal antibody of BamK. Strategy for the screening of monoclonal antibodies (MK1 through MK5) made to a diagnostic peptide (IALNEGERYRVDRT$_{271}$) of BamK not present in BamA based on multiple sequence alignment analysis. At left, a structural overlay of the modelled structures for *Kp*BamA (red) and *Kp*BamK (blue).

Monoclonal antibodies against this peptide were produced from five hybridomas (MK1 through MK5) with preliminary ELISAs confirming positive detection to the diagnostic BamK peptide (performed by Monash MATF). The antibodies MK 1 through 5 were tested by immunoblotting on *K. pneumoniae* B5055 Δ*bamA*::*bamK* (Explored further in Chapter 5) whole cell lysates analysed by SDS-PAGE (Figure 3.2.7 A) or prepared total membranes analysed by BN-PAGE (Figure 3.2.7 B).

It was found that clones MK 3 and 5 were able to detect BamK under denaturing conditions, whereas only MK3 could detect BamK under native conditions. Consequently, both MK3 and MK5 were subsequently immortalised for large scale production for future assays.
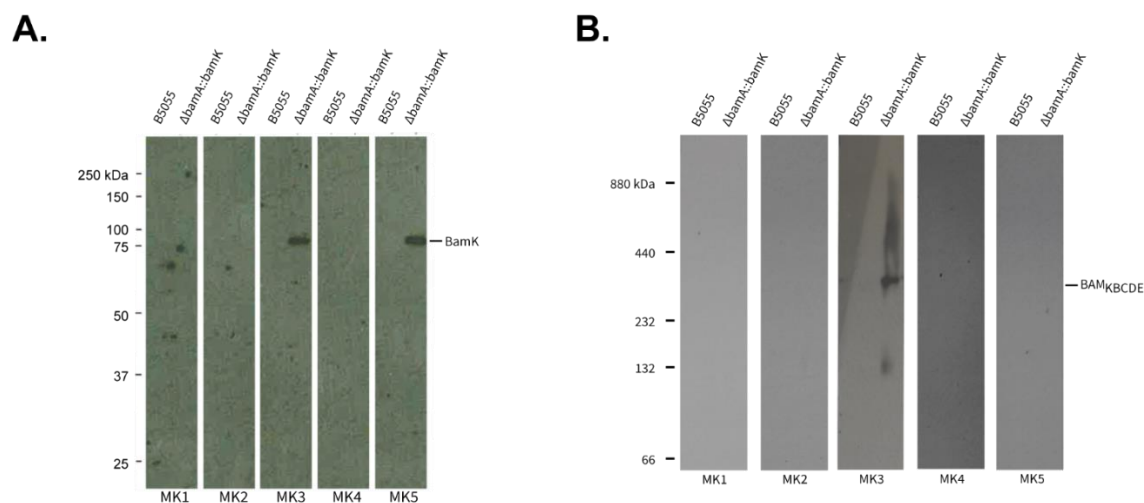
Figure 3.2.7: Generation and identification of a suitable monoclonal antibody for BamK identification. (A) Immunoblotting for BamK with 5 candidate mouse monoclonal antibodies on whole cell lysates of wildtype B5055 and ΔbamA::bamK analysed by (12%) SDS page. (B) Immunoblotting for BamK with 5 candidate mouse monoclonal antibodies on prepared total membranes isolated from wildtype B5055 and ΔbamA::bamK analysed by (5-16%) BN-PAGE.

The putative *bamK* promoter region is highly conserved, and I hypothesised that the corresponding low transcript levels could be limited by the strength or likely gene repression of this promoter region. To test this, I performed immunoblot analysis looking at select *K. pneumoniae* strains that shared an identical upstream *bamK* region. Shown in Figure 3.2.8, BamK was not detected in the selected strains and only expressed in a control strain where the putative promoter region was replaced with the known constitutive promoter of *bamA* (Refer to Chapter 5.0, section 5.3). Together, these data suggest that the putative *bamK* promoter could be a rate-limiting factor in its transcription and translation.
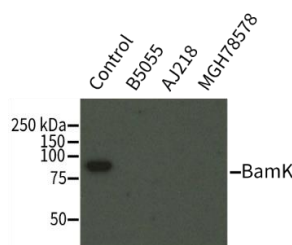
Figure 3.2.8: Western blot analysis of *K. pneumoniae* strains harbouring BamK. Whole cell lysates from select *K. pneumoniae* strains analysed by SDS-PAGE (12%) then detected with a monoclonal BamK antibody Whole cell lysates were prepared from mid-logarithmic cultures grown in nutrient rich LB media. Control – *bamK* under the promoter of *bamA* B5055 – *K. pneumoniae* B5055, *K. pneumoniae* AJ218 and *K. pneumoniae* MGH78578. The putative promoter region of *bamK* (500 bp upstream region) of all three strains have 100 % sequence identity. The protein coding region has >99.5 % identity between all three strains.

3.3 Generation of a mutagenesis library to identify potential repressor(s) of *bamK*

Transposon mutagenesis is a powerful genomic technique that can identify essential genes or fitness conferring genes in defined environmental conditions (Mazurkiewicz *et al.*, 2006, Baba *et al.*, 2006). The basis of this technique is the mobility of genetic elements, termed transposons, that are capable of self-excision and reinsertion within a target genome at random or in a site-specific manner (Reznikoff, 2003). In the case of *bamK*, I hypothesised that its expression is inhibited by transcriptional repressors under standard laboratory conditions; consequently, negative selection for its functional role would not be discernible. Therefore, I instead aimed to generate a library screening for gain-of-function mutations where *bamK* expression is detected by proxy of antibiotic resistance. To identify genomic regulators of *bamK* by way of a gain-of-function screen, a transposon mutagenesis library was generated in *K. pneumoniae* B5055 where the coding region of *bamK* was precisely replaced with a β-lactamase gene (hereafter referred to as Amp^R) (Figure 3.3.1). The gene replacement and subsequent excision of a kanamycin cassette (here after Kan^R) was confirmed by PCR and sequencing, using primers flanking the targeted regions (Figure 3.3.2). The resultant strain, *K. pneumoniae* B5055 Δ*bamK*::Amp^R, expressed the β-lactamase gene under the control of the original *bamK* promoter. In Chapter 5.0, a *K. pneumoniae* B5055 Δ*bamK* strain is further investigated for phenotypic and molecular characterisation.
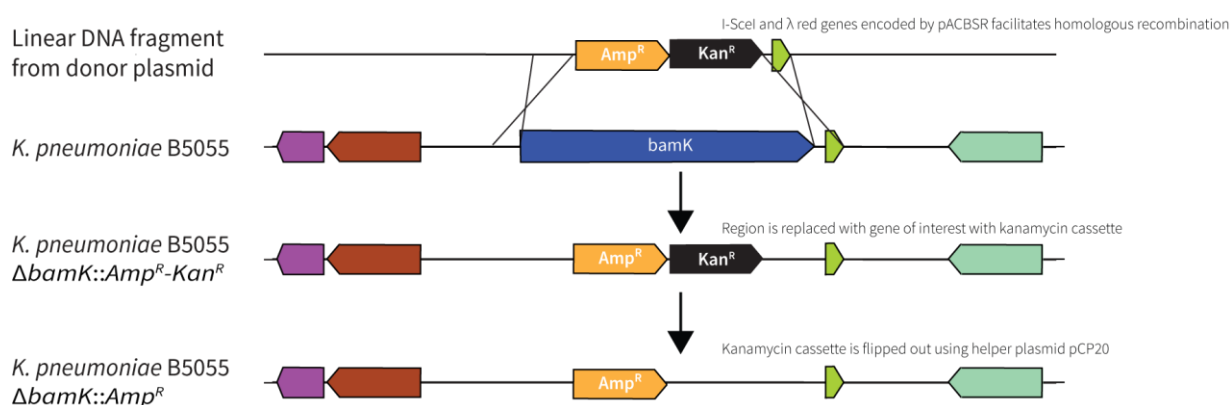
Figure 3.3.1: Strategy for genomic gene replacement of *bamK* with Amp^R in *K. pneumoniae* B5055. Schematic showing the replacement of the *bamK* coding sequence with an ampicillin resistance coding sequence to produce gene replacement strain *K. pneumoniae* B5055 Δ*bamK*::Amp^R. The 'gene-gorging' method utilises a two-plasmid system for genetic manipulation in broad range of bacterial species (Herring *et al.*, 2003). Briefly, a donor plasmid (plasmid 1) containing the kanamycin cassette with flanking homology regions is excised into a linear fragment by I-SceI by recombineering plasmid (plasmid 2) pACBSR. The recombineering plasmid pACBSR carries the λ-Red and I-SceI endonuclease genes, under the control of an *araBAD* promoter. The obvious advantage of this system is that multiple copies of the homologous DNA are present in the bacterial cell, which increases the number of potential recombination events. The gene replacement of the *bamK* CDS with the ampicillin resistance gene (Amp^R) is under control of the putative *bamK* promoter.
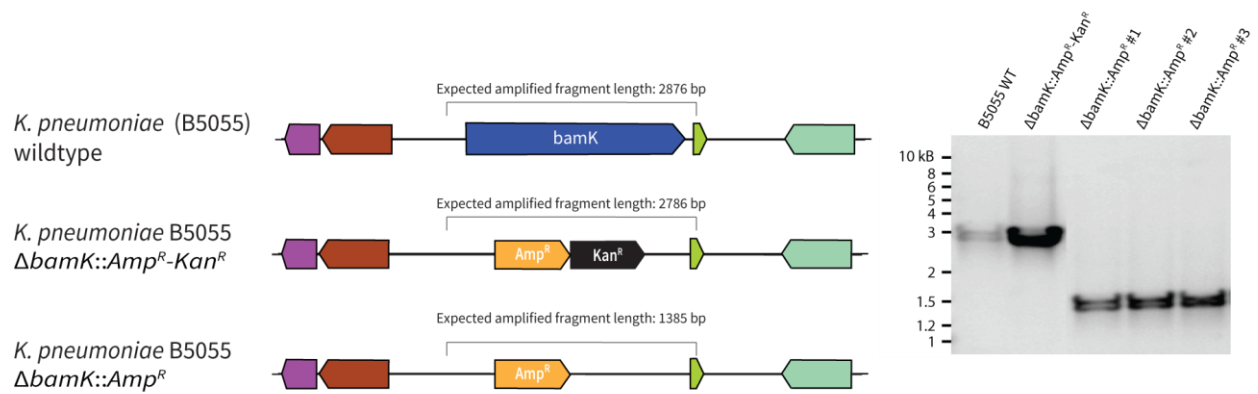
Figure 3.3.2: PCR confirmation of *bamK* replacement with Amp^R in *K. pneumoniae* B5055. Left hand side, cartoon showing the expected fragment lengths of *K. pneumoniae* B5055 at different stages during the gene knock-in process (refer to Figure 3.3.1). Right hand side, PCR confirmation illustrating the intermediate gene knock-in strains by PCR amplification using primers, VVT0102 and VVT103, which lie outside the region of genomic integration. PCR products were analysed by 1 % TAE agarose gel and sequenced to confirm the gene replacement.

Details pertaining to the generation of transposon library can be found in Chapter 2.0 'Materials and Methods'. The creation of a preliminary transposon insertion library was successfully achieved in recipient strain *K. pneumoniae* B5055 Δ*bamK*::Amp^R via a bacterial conjugation-based system utilising donor *E. coli* S17-λ*pir* harbouring the suicide vector, pUT-mini-Tn*5Km2* (Figure 3.3.3). The spontaneous mutant *K. pneumoniae* B5055^Rif Δ*bamK*::Amp^R, was isolated via gradient plate method to allow antibiotic selection from susceptible donor *E. coli* S17-λ *pir* after solid and liquid mating techniques. The rationale for this strategy was that a triple antibiotic selection screen would be able to identify *K. pneumoniae* B5055^Rif (rifampicin-resistant), integrated with a Tn5 transposon (kanamycin marker), with reporter β-lactamase expression (ampicillin resistance) as a proxy for *bamK* expression (Figure 3.3.4). A total of 150 ampicillin-resistant transposon mutants (*bK*-candidates) were archived for further analysis.
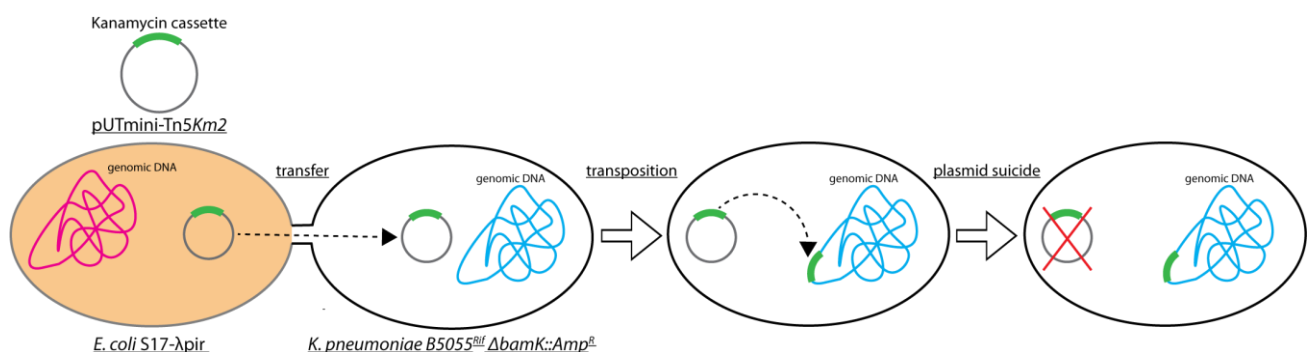
Figure 3.3.3: Tn*5Km2* transposon integration into *K. pneumoniae* B5055^Rif Δ*bamK*::Amp^R. Schematic illustrating the generation of transposon mutants using a conjugation strategy. The mini-transposon plasmid pUTmini-Tn*5Km2* is a suicide vector that is unable to replicate upon conjugal plasmid transfer into its recipient cell. A spontaneous rifampicin mutant of *K. pneumoniae* B5055^Rif Δ*bamK*::Amp^R was isolated to allow selection for *K. pneumoniae* mutants from *E. coli* S17-λ*pir* harbouring the suicide vector. The resulting transconjugant *K. pneumoniae* B5055^Rif Δ*bamK*::Amp^R mutants would theoretically be kanamycin and rifampicin mutant, and only ampicillin resistant if *bamK* expression was de-repressed/activated.
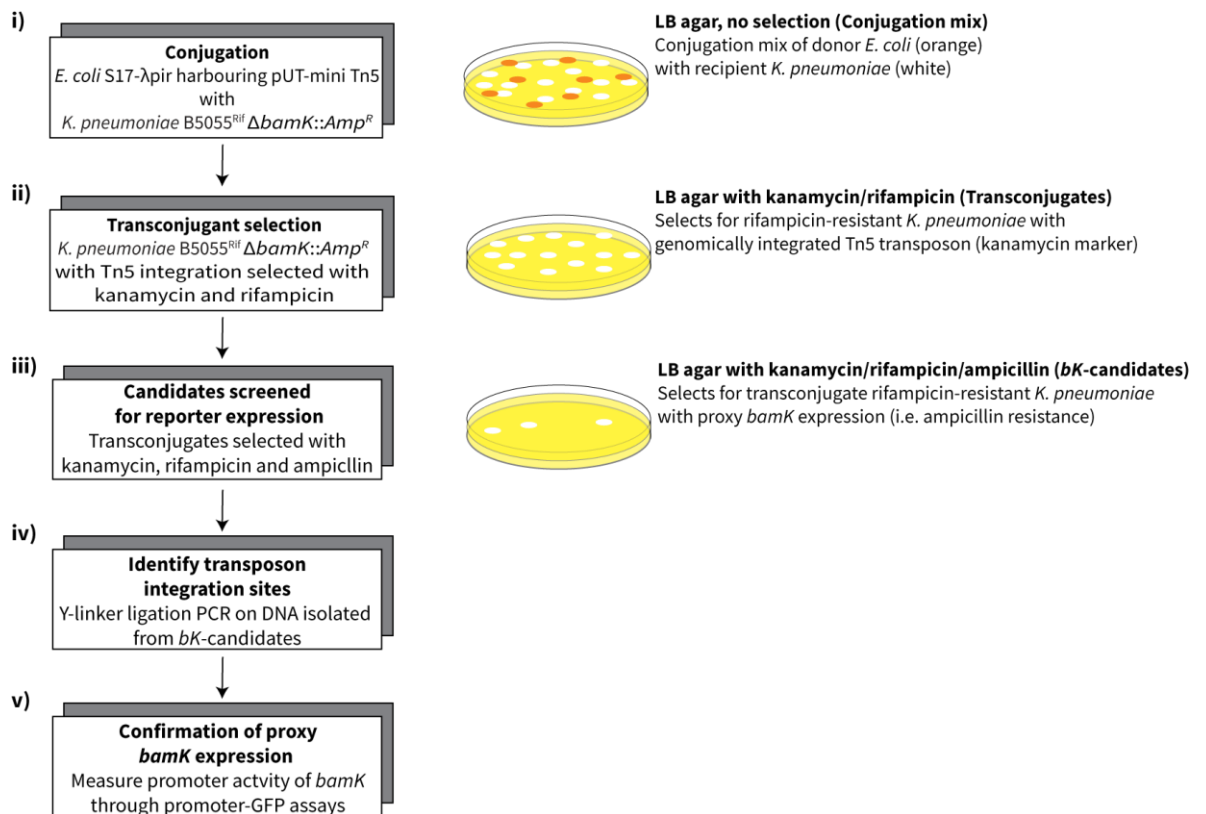
Figure 3.3.4: Schematic flowchart of generating a mutagenesis library to identify proxy *bamK* expression in *K. pneumoniae* B5055^{Rif} Δ*bamK*::Amp^R. A flow chart illustrating the strategy to generate and identify gain-of-function mutants as a reporter proxy for *bamK* expression. i) Donor strain, *E. coli* S17-λpir harbouring suicide transposon plasmid pUT-mini Tn5, was conjugated with *K. pneumoniae* B5055^{Rif} Δ*bamK*::Amp^R. *K. pneumoniae* B5055^{Rif} Δ*bamK*::Amp^R is a spontaneous rifampicin resistant mutant isolated by gradient plate method as described previously (Carsenti-Etesse *et al.*, 1999). The donor *E. coli* S17-λpir is susceptible to rifampicin. ii) Mutant *K. pneumoniae* B5055^{Rif} Δ*bamK*::Amp^R, with successful integration of transposon Tn5 (kanamycin marker), were selected by plating on LB-agar supplemented with kanamycin and rifampicin. iii) Transconjugant colonies from (ii) were replica plated to LB-agar supplemented with kanamycin, rifampicin and ampicillin. The inclusion of ampicillin would select for transconjugate *K. pneumoniae* B5055^{Rif} Δ*bamK*::Amp^R mutants with a gain-of-function resistance mutation (i.e. to ampicillin) as a proxy for *bamK* expression. These mutants were termed *bK*-candidates. iv) DNA was isolated from *bK*-candidates and were subjected to Y-linker ligation PCR to identify the genomic integration sites of the Tn5 transposons (Kwon & Ricke, 2000). v) Confirmation of *bamK* expression was screened promoter-fusion GFP assays.

To establish how effective this system was in generating gain-of-function mutants, I determined the frequency of ampicillin-resistant *bK*-candidates (iii) within the initial *K. pneumoniae* B5055 transconjugate pool (ii) (refer to Figure 3.3.4). To elucidate this, appropriately diluted conjugation mixtures (i) were replica-plated to LB-kanamycin/rifampicin or LB-ampicillin/kanamycin/rifampicin agar plates (Figure 3.3.5). Replicate plating revealed ~5% (29/562) of transconjugants were ampicillin, kanamycin and rifampicin resistant, which were the suspected gain-of-function *bK*-candidates that would need to be validated and screened. Assuming the *Tn*5 transposon was randomly integrating into the *K. pneumoniae* B5055^{Rif} Δ*bamK*::Amp^R genome without bias, I posited that a 5 % frequency for gain-of-function mutants were the result of selecting false-positives with spontaneous mutations that conferred ampicillin-resistance.
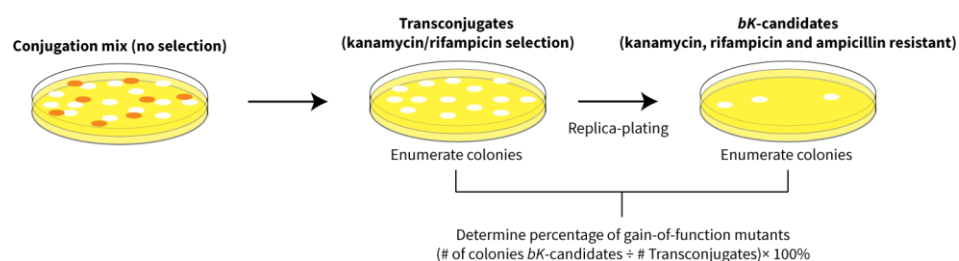
Figure 3.3.5: Flowchart and strategy to determine gain-of-function mutants as proxy for *bamK* expression. Schematic showing the approach taken to determine the percentage of gain-of-function mutants from the transconjugant mix (see Figure 3.3.3). Using this transposon mutagenesis system method, a library of 25,787 transconjugant clones would theoretically saturate the genome of *K. pneumoniae* B5055 with 99% confidence using a commonly used Poisson-statistics based formula (calculations found in Appendix 4) (Zilsel *et al.*, 1992).

To determine if this was the case, i.e. that the designed gain-of-function screen generated false-positive mutants for screening, five randomly selected *bK*-candidates were further investigated by a series of PCR reactions. As a preliminary screen, total DNA was isolated then analysed by PCR for presence of a Tn5 transposon and the upstream promoter region of *bamK* as proof for transposon integration into the genome *K. pneumoniae* B5055 Δ*bamK::*Amp^R (Figure 3.3.6 A-B). As all five *bK*-candidates contained both elements, the flanking sequences of the transposons were amplified using Y-linker ligation PCR (Kwon & Ricke, 2000). As shown in Figure 3.3.6 C, flanking regions of the transposons were successfully amplified from the five *bK*-candidates and were then gel-purified for sequencing to allow mapping of transposon integration sites to the *K. pneumoniae* B5055 chromosomal and plasmid sequences (GenBanK™ accessions: FO834904, FO834905 and FO834906).
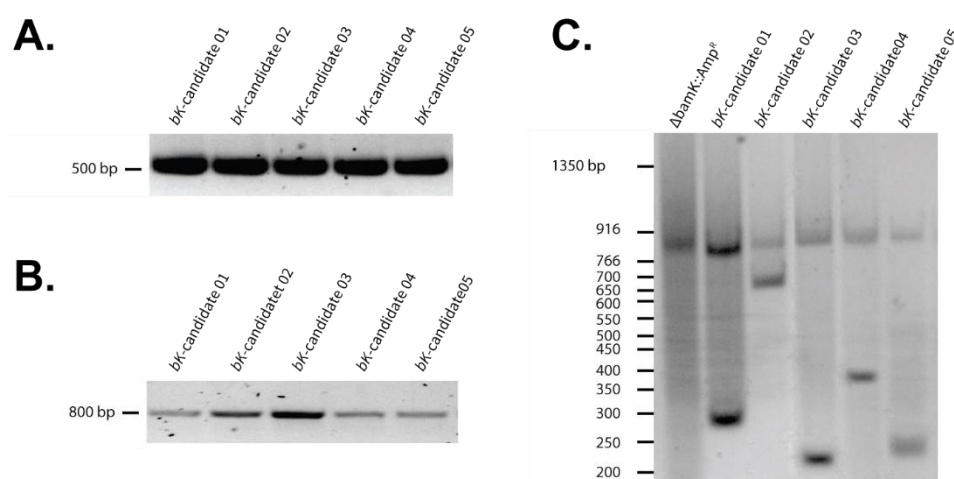
Figure 3.3.6: PCR analysis of randomly selected *bK*-candidates to confirm *K. pneumoniae* transconjugate and Y-linker ligation. (A) Amplification *bamK* upstream region (only found in *K. pneumoniae* B5055) or (B) genome integrated *Tn*5 transposon analysed using 1 % TAE agarose gel electrophoresis. (C) PCR amplification of the DNA sequences flanking mini-Tn5 insertions in the chromosomal DNA of *K. pneumoniae* B5055 Δ*bamK::*Amp^R. Included as a control was parental *K. pneumoniae B5055* Δ*bamK*::Amp^R analysed using 2 % TAE agarose gel electrophoresis.

As presented in Table 3.3.1, the transposon integration sites of the five randomly selected *bK*-candidates all mapped to gene coding regions in the bacterial chromosome and neither of the plasmids of *K. pneumoniae* B5055. Of the five insertions, none were linked to known transcriptional regulators and presumably inserted and disrupted non-essential genes in the tested growth conditions (i.e. solid and liquid culture in nutrient rich LB media). Amplification, sequencing and mapping of DNA reads to the *K. pneumoniae* B5055 genome was successful, but the identified gene targets did not reflect products involved in gene-specific repression or transcription.

Table 3.3.1: Transposon insertion sites in *K. pneumoniae* B5055^{Rif} Δ*bamK*::Amp^R. Y-linker ligation PCR amplifying the flanking sequences of the Tn5 transposon insertion sites were sequenced then mapped to the chromosome of *K. pneumoniae* B5055 (GenBanK™ accessions: FO834904). Of the five analysed transposon insertions: (i) three mutants were part of amino acid biosynthesis pathways, (ii) one mutant part of the proUVW operon involved in cellular osmoregulation and (iii) one was a putative exonuclease harbouring an YqaJ domain involved in dsDNA digestion.

Gene name	Brief function description	*K. pneumoniae* B5055 locus tag, location	Library identifier
moeA	Involved in the pathway for molybdopterin biosynthesis	BN49_1911, chromosome	*bK*-candidate 01 (r-test)
ilvD	Involved in the pathways for L-isoleucine. and L-valine biosynthesis	BN49_4589, chromosome	*bK*-candidate 02 (r-test)
Uncharacterised protein	Contains YqaJ - putative exonuclease, viral origin	BN49_1464, chromosome	*bK*-candidate 04 (r-test)
thrB	Involved in the pathway for L-threonine biosynthesis	BN49_4337, chromosome	*bK*-candidate 03 (r-test)
proV	Part of the ProU ABC transporter complex involved in glycine betaine and proline betaine uptake	BN49_1115, chromosome	*bK*-candidate 05 (r-test)

To validate the *bK*-candidates, promoter fusion plasmids were transformed into each of them and their *bamK* promoter activity was monitored. As shown in Figure 3.3.7, *bamK* promoter activity in all five suspects did not have a marked difference compared to the controls. As none of these gene disruptions targeted gene regulatory transcription factors and the promoter fusion results showed no differences, these data suggest the strategy to use a gain-of-function mutagenesis screen was generating false-positives mutants and an alternative strategy was devised to address the activation of *bamK*.
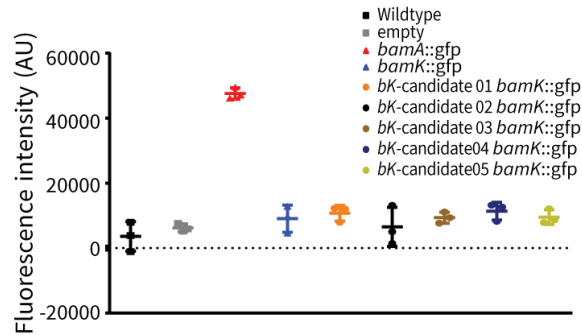
Figure 3.3.7: Promoter activity of bamK in candidate *K. pneumoniae* B5055^{Rif} Δ*bamK*::Amp^R transconjugates. Analysis of the expression of *bamK* in *K. pneumoniae* B5055 strains carrying pUA139 transcriptional fusion (*bamA*::*gfp* or *bamk*::*gfp*) reporter plasmids. All bacteria were grown in LB media to mid-logarithmic phase then measured for promoter activity using Spark™ 10M multimode reader. Wildtype – B5055 no harboured reporter plasmid empty – B5055 containing empty vector pUA139 ($n = 3$, horizontal bar represents the mean, error bars represent SD).

3.4 Identification of DNA-binding proteins that may interact with the putative promoter of *bamK*

To identify whether other DNA-binding proteins may be regulating the expression of *bamK*, a multistep strategy using a one-dimensional electrophoretic mobility shift assay (EMSA) coupled with mass spectrometry and immunoblotting was employed. Firstly, to determine if the putative promoter of *bamK* had any interacting DNA-binding proteins, an EMSA was carried out between amplified PCR fragments (bait) of the promoter elements mixed with *K. pneumoniae* B5055 cleared cell lysate (Figure 3.4.1).
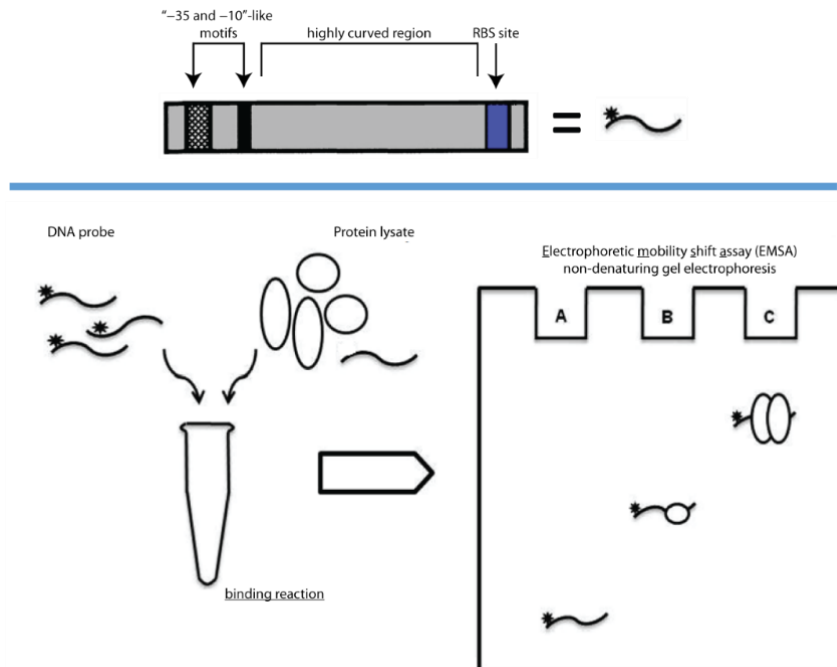
Figure 3.4.1: Identification of DNA-binding protein(s) on the putative *bamK* promoter using EMSA. Depicted here is a cartoon schematic outlining the strategy to identify DNA-binding proteins to the putative *bamK* promoter. A 220 bp amplified fragment (above) comprised of the identified promoter elements of *bamK* is mixed with *K. pneumoniae* B5055 protein lysate. This binding reaction is then analysed under native PAGE conditions and then stained with a DNA gel stain. The amplified nucleotide fragment (a), if bound (b-c), results in a shifted DNA::protein complex (below).

As shown in Figure 3.4.2 A, a mobility shift occurred when the cleared cell lysate was mixed with the amplified 220 bp fragment containing putative *bamK* promoter elements. A control experiment utilising a synthetic polynucleotide, poly(dI-dC), was performed in parallel where the lysate was initially mixed with poly(dI-dC) to act as a competitor for non-specific DNA-binding proteins before the addition of the intended amplified promoter fragments demonstrating that the observed mobility shift was due to a DNA-protein binding interaction (Figure 3.4.2 B).

A.

| DNA probe (250 nmole) | + | + | + | + | - |
| Klebsiella lysate (µg total) | 0 | 8 | 16 | 32 | 32 |

DNA::protein complex

free DNA probe

DNA from lysate

B.

Poly (Di-Dc) (1µg)	+	+	+	+	+
DNA probe (250 nmole)	+	+	+	+	-
Klebsiella lysate (µg total)	0	8	16	32	32

DNA::protein complex

free DNA probe

DNA from lysate
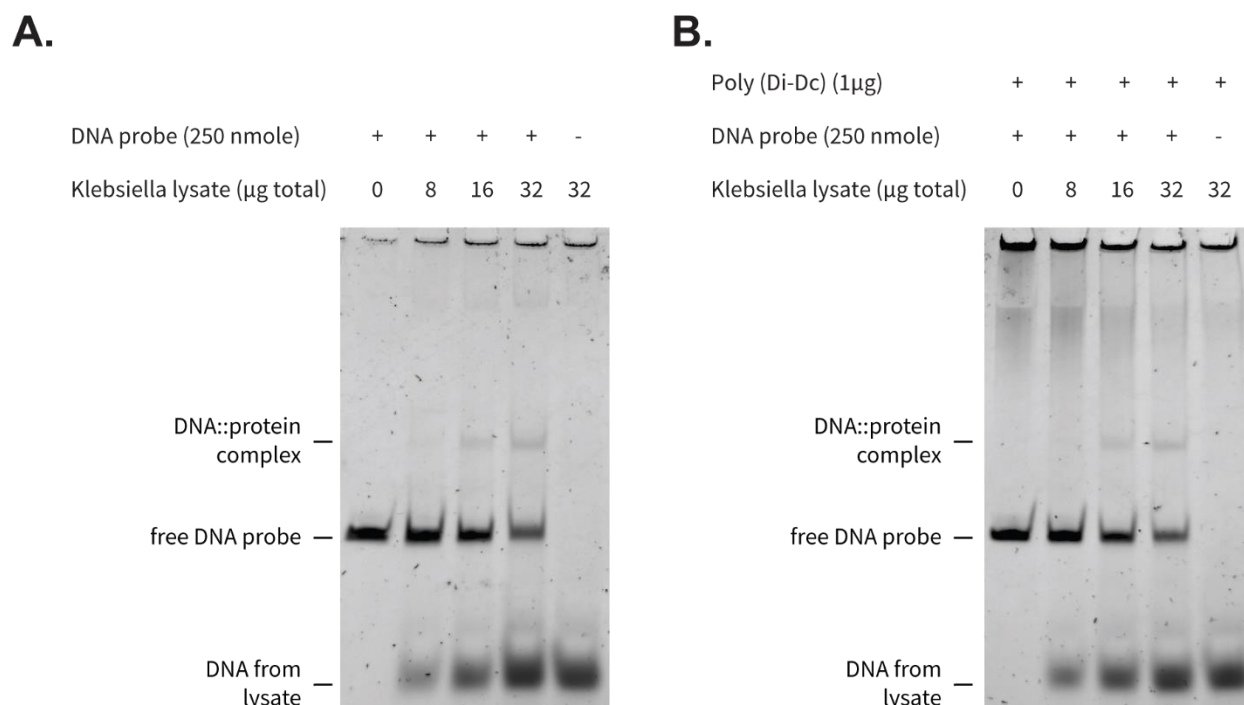
Figure 3.4.2: EMSAs identifying proteins from *K. pneumoniae* B5055 lysate binding with an amplified *bamK* promoter fragment (220 bp). (A) Bait promoter fragments PCR amplified from *K. pneumoniae* B5055 were subjected to EMSA (TBE 4-8%) using cleared cell lysate (mid-log cultures lysed with detergent-based protein extraction reagent *B-PER*™ (Thermo Scientifc ™) as prey. (B) Even in the presence of synthetic polynucleotide poly(dI-dC) as a competitor, formation of a DNA::protein complex formation was still observed.

To further characterise which elements of the putative *bamK* promoter were interacting with the *K. pneumoniae* B5055 protein lysate, additional EMSAs were performed with 3' promoter region truncations of the original 220 bp fragment (Figure 3.4.3 A). From these complementary pull-down assays, it was observed that mobility shifts were seen with PCR fragments, which retained the predicted highly curved AT-rich region (Figure 3.4.3 B). These results suggest that a protein is selectively binding the amplified fragment representing the putative *bamK* promoter region and could be the basis of its suspected gene repression.

A.

"−35 and −10"-like motifs

highly curved region

RBS site

Fragment 1

Fragment 2

Fragment 3

B.

	Fragment 1				Fragment 2				Fragment 3				
Poly (Di-Dc) (1μg)	+	+	+	+	+	+	+	+	+	+	+	+	+
DNA probe (250 nmole)	+	+	+	+	+	+	+	+	+	+	+	+	-
Klebsiella lysate (μg total)	0	8	16	32	0	8	16	32	0	8	16	32	32

DNA::protein complex

free DNA probe (fragment 1)

free DNA probe (fragment 2)

free DNA probe (fragment 3)

DNA from lysate
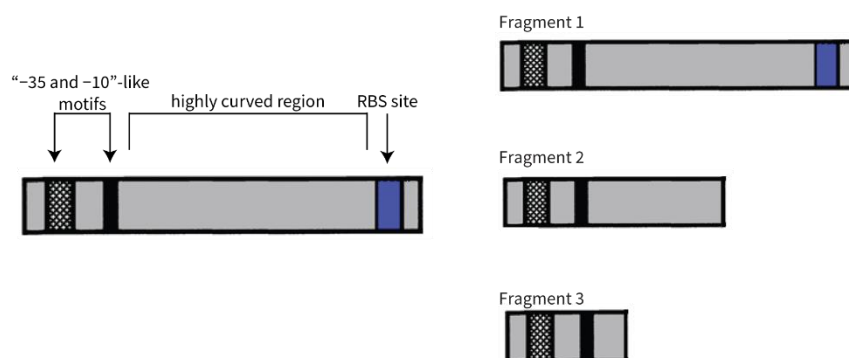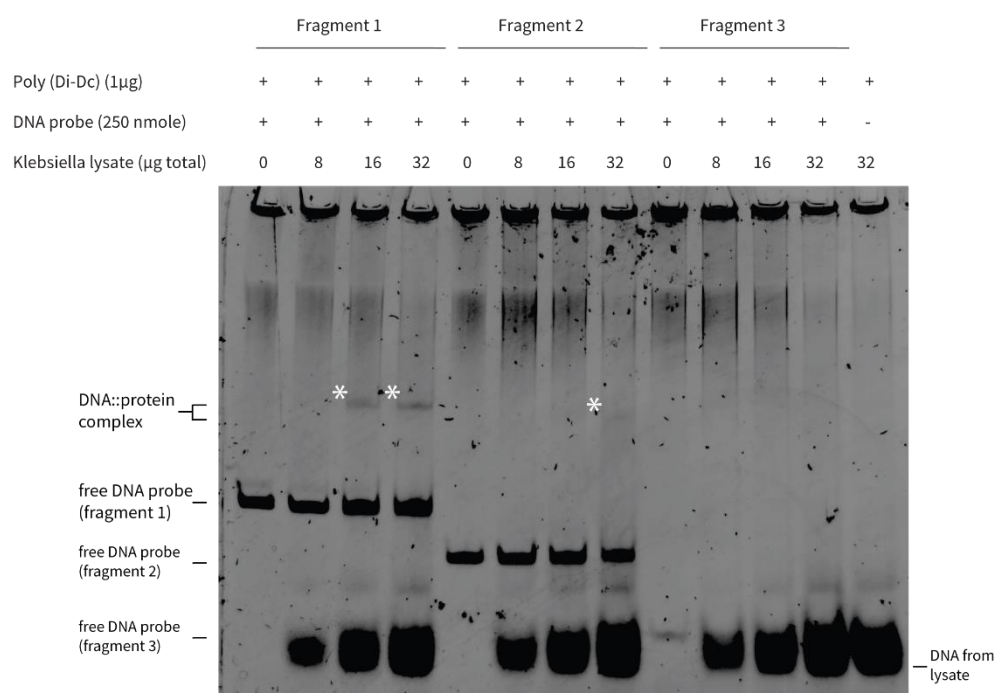
Figure 3.4.3: Profiling the specific promoter elements required for DNA::protein binding. (A) Cartoon representations of the truncated PCR promoter fragments used to elucidate the key elements interacting in the DNA-protein complex formation. (B) Truncated bait promoter fragments PCR amplified from *K. pneumoniae* B5055 were subjected to EMSA (TBE 4-8%) using cleared cell lysate (mid-log cultures lysed with detergent-based protein extraction reagent *B-PER*™ (Thermo Scientific ™) as prey. Protein-DNA complex are marked with white asterisks. DNA::protein band formation was seen for fragments 1 and 2 but was not seen in 3. It should be noted, that at the utilised DNA probe concentrations observation of a 'Fragment 3' with lysate formatting DNA::protein may be difficult, as the visualised signal of the GelRed™ DNA stain intercalates nucleotide substrates in a concentration dependent based manner (Crisafuli *et al.*, 2014).

The candidate DNA::protein complex gel band (Figure 3.4.3) was excised and analysed by mass spectrometry (MS) to identify DNA-binding proteins. Included as a control, a gel band in the corresponding migration distance but without DNA probe (i.e. poly(dI-dC) and 32 μg lysate) was also excised for analysis as a comparative baseline. The DNA::protein complex was enzymatically digested with trypsin and the resulting peptides were determined by liquid chromatography-tandem mass spectrometry mass spectrometry (LC-MS/MS). Using proteomics search program MASCOT (Koenig *et al.*, 2008), peptide hits were analysed then matched against protein entries from reference

K. pneumoniae Uniprot proteomes (UniProt, 2015). From the list, two approaches were used to refine the output: (i) a "Black and White" approach – where proteins exclusively found in the DNA::protein sample but not the control, and (ii) comparative upregulation – noting proteins that are at least 2 fold increased in the DNA::protein sample compared to the control. This combined meta-analysis resulted in a list of 218 protein candidates Appendix 3. Presented in Table 3.4.1, proteins likely to be involved in DNA-binding or transcriptional regulation were considered as candidates for gene knockouts to assess for *bamK* de-repression. Two such candidates, *dksA* and *hup*A, were selected for subsequent gene deletion studies based on their potential involvement in *bamK* gene regulation. The others were disregarded, as they were either essential genes in *E. coli* or likely to be a promiscuous DNA-binding factor not specifically involved in promoter binding and gene repression.

Table 3.4.1: Suspect DNA-binding proteins potentially regulating *bamK* expression identified by MS analysis. Mass spectromry analysis generated a preliminary list of protein hits potentially interacting with the amplified PCR fragment corresponding to the putative *bamK* promoter. Suspects selected for deletion studies were chosen based on their known or predicted plausible functional roles in transcriptional regulation and their suspsected essentiality. Assumed gene essentiality was based on the *E. coli* Keio deletion library study (Baba *et al.*, 2006).

Gene name	Known function?	MW (kDA)	Lethal in KEIO
DNA-binding protein HU-alpha (*hupA*)	DNA binding/pleitropic transcriptional regulator, Chromosome condensation and role in stationary phase adaptive mutations (Prieto *et al.*, 2012, Williams & Foster, 2007).	9.4768	NO
RNA polymerase-binding transcription factor (*dksA*)	Transcription factor that acts by binding directly to the RNA polymerase (RNAP). Required for negative regulation of rRNA expression and positive regulation of several amino acid biosynthesis promoters. Also required for regulation of fis expression. (Paul *et al.*, 2004, Parshin *et al.*, 2015).	12.456	NO
Type I restriction-modification system, restriction subunit R (*hdsR*)	Prokaryotic defence against foreign DNA (Loenen *et al.*, 2014)	111.98	NO
DNA polymerase III subunit beta (*dnaN*)	DNA polymerase III is a multi-subunit complex, subunit is the DNA clamp component which maintains stability of the overall DNA polymereaase complex (Stukenberg *et al.*, 1991).	40.481	YES
Single-stranded DNA-binding protein (*ssb*)	Plays an important role in DNA replication, recombination and repair. Binds to ssDNA and to an array of partner proteins to recruit them to their sites of action during DNA metabolism. (Meyer & Laine, 1990).	17.867	YES

To identify if deletion of either *hupA* or *dksA* modulated the expression of *bamK¸* I generated gene deletion mutants of each in *K. pneumoniae* B5055 using the gene gorging method (Herring *et al.*, 2003). This resulted in strains *K. pneumoniae* B5055 Δ*dksA*::Kan[R] and *K. pneumoniae* B5055 Δ*hupA*::Kan[R] (Figure 3.4.4). The deletion strains were confirmed by PCR analysis and sequencing, using primers flanking the targeted regions (Figure 3.4.5).
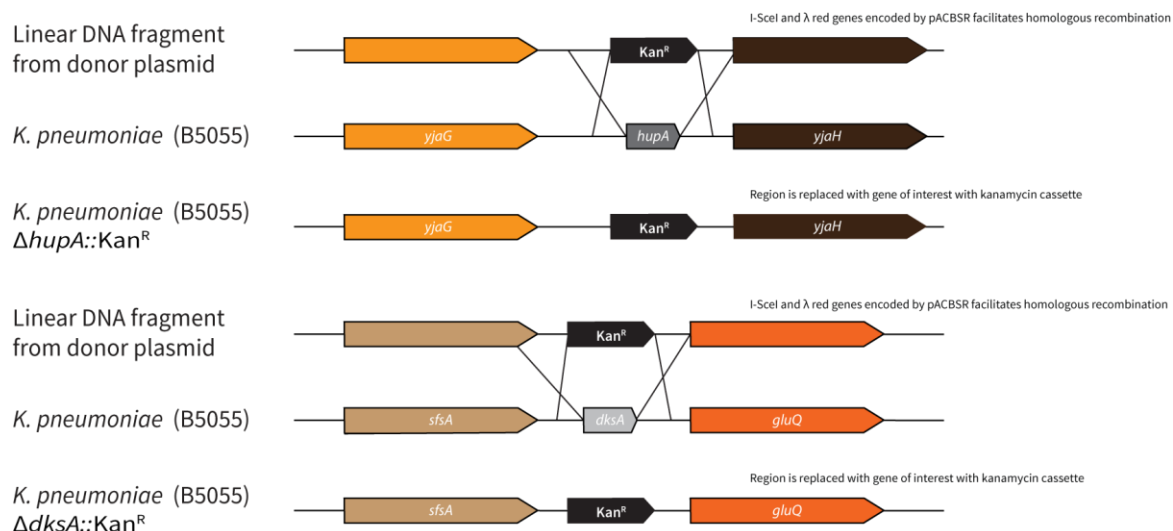
Figure 3.4.4: Strategy for deletion of *hupA* or *dksA* in *K. pneumoniae* B5055. Schematic showing the replacement of the *hupA* or *dksA* in *K. pneumoniae* B5055 with an excisable kanamycin cassette. The 'gene-gorging' method utilises a two-plasmid system for genetic manipulation in broad range of bacterial species (Herring *et al.*, 2003). Briefly, a donor plasmid (plasmid 1) containing the kanamycin cassette with flanking homology regions is excised into a linear fragment by I-SceI by recombineering plasmid (plasmid 2) pACBSR. The recombineering plasmid pACBSR carries the λ-Red and I-SceI endonuclease genes, under the control of an *araBAD* promoter. The obvious advantage of this system is that multiple copies of the homologous DNA are present in the bacterial cell, which increases the number of potential recombination events.
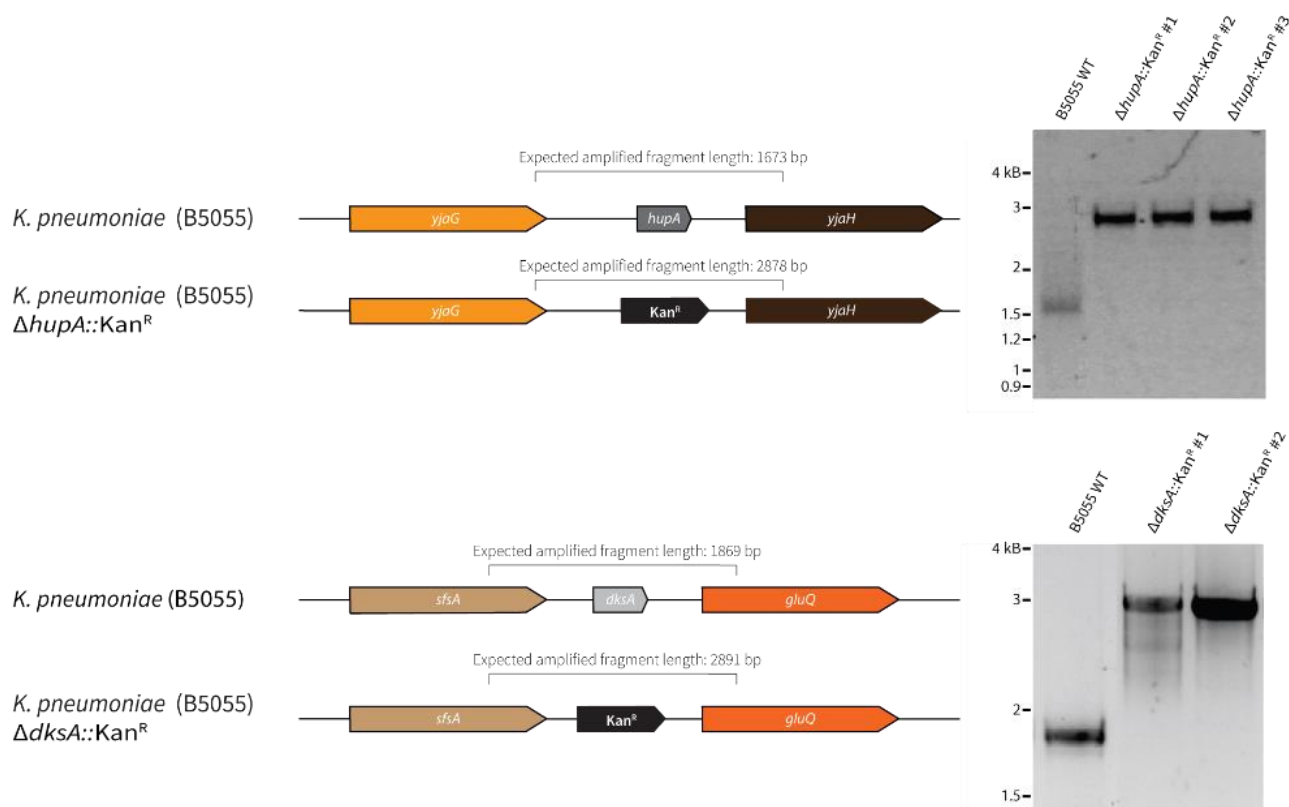
Figure 3.4.5: PCR confirmation of *K. pneumoniae* B5055 Δ*hupA*::Kan^R and Δ*dksA*::Kan^R. Left hand side, cartoon schematics of the expected fragment lengths of *K. pneumoniae* B5055 before and after the kanamycin cassette integration (refer to Figure 3.4.4). Right hand side, PCR amplification of *K. pneumoniae* B5055 before and after the kanamycin cassette integration. For *hupA* and *dksA*, primer pairs VVT094 and VVT095 or VVT092 and VVT093, which anneal outside the regions of genomic integration were used, respectively. PCR products were analysed by 1 % TAE agarose gel electrophoresis and sequenced to confirm the gene deletion.

To confirm whether the *hupA* or *dksA* deletion strains de-repressed *bamK* expression, promoter fusion plasmids were transformed into the relevant strains (*K. pneumoniae B5055* Δ*hupA*::Kan[R] and Δ*dksA*::Kan[R]) and their *bamK* promoter activity was monitored. As shown in Figure 3.4.6 A, *bamK* promoter activity in the deletion strains did not have a marked difference compared to the controls. As an independent assessment of expression, whole cell lysates were examined by immunoblot for BamK expression from logarithmic and stationary phase cultures of the deletion strains. In previous studies, *hupA* and *dksA* has both been shown to have growth phase-dependent regulation (Claret & Rouviere-Yaniv, 1997, Mallik *et al.*, 2006). As shown Figure 3.4.6 B, expression in either deletion strain did not result in detectable levels of BamK expression. Together, these results suggest that single deletions of either *hupA* or *dksA* does not modulate the transcription or translation of *bamK*.
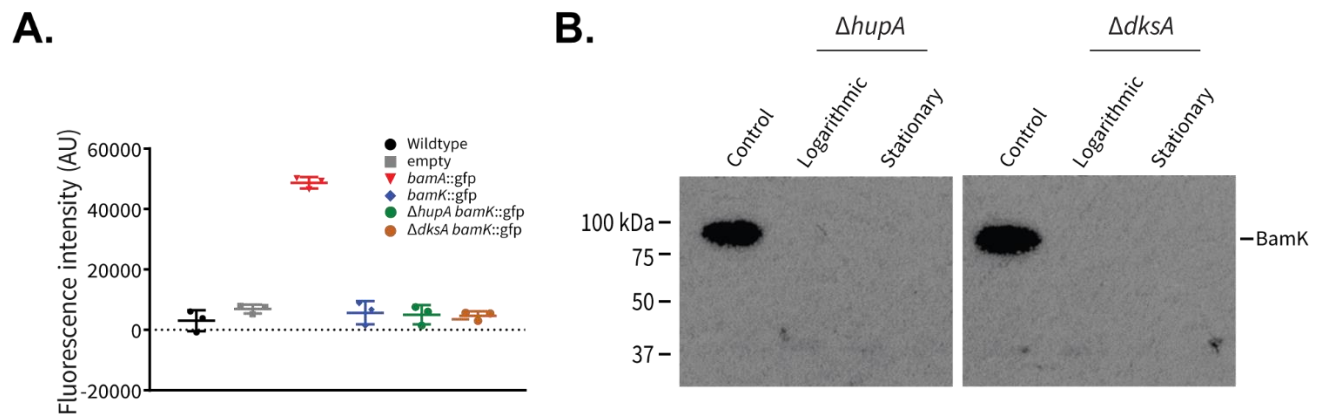
Figure 3.4.6: Promoter activity of bamK and immunoblot analysis for the expression of BamK in *K. pneumoniae* B5055 Δ*hupA* and Δ*dksA*. (A) Analysis of the expression of *bamK* in *K. pneumoniae* B5055 strains carrying pUA139 transcriptional fusion (*bamk*::*gfp*) reporter plasmids. In deletion strains, *K. pneumoniae* B5055 Δ*hupA* and Δ*dksA*, promoter activity of *bamK* was also measured. All bacteria were grown in LB media to mid-logarithmic phase then measured for promoter activity using Spark™ 10M multimode reader. Wildtype – B5055 no harboured reporter plasmid. empty – B5055 containing empty vector pUA139. (n = 3, horizontal bar represents the mean, error bars represent SD). (B) SDS-PAGE (12%) and immunoblot analysis of whole cell lysates from *K. pneumoniae* B5055 Δ*hupA* and Δ*dksA* at either logarithmic (OD_{600} = ~0.6) or stationary (OD_{600} = ~1.6). Included as a control is a *K. pneumoniae* B5055 strain which constitutively expresses BamK (see Chapter 5, section 5.3)

3.5 Discussion

The bacterial σ^E stress response is a regulatory pathway responsible for monitoring misfolded β-barrel sequences in the periplasm and addresses this by differentially regulating factors and OMPs to prevent aberrant protein accumulation causing extracytoplasmic stress (Miot & Betton, 2004, Soltes *et al.*, 2017). Genes of the σ^E regulon include those that encode components of the BAM complex (*bamA* through *bamE*) (Dartigalongue *et al.*, 2001, Onufryk *et al.*, 2005), where *bamA* encodes for the core protein of the BAM complex a member of the Omp85 protein family vital for OMP assembly (Silhavy & Malinverni, 2011). In *K. pneumoniae*, four Omp85 proteins (BamA, BamK, TamA and BamL) are conserved in the core genome (Heinz & Lithgow, 2014, Torres *et al.*, 2018). In previous studies, the genetic regulation and expression of BamA and TamA has been investigated in different bacterial species (Huerta & Collado-Vides, 2003, Mutalik *et al.*, 2009, Rowley *et al.*, 2011); however, the Omp85s in *K. pneumoniae* had remained largely unexamined. In this chapter, I presented gene expression data of *bamA*, *bamK*, *bamL*, and *tamA* present in *K. pneumoniae* B5055, and further explore the genetic elements which could be regulating novel gene *bamK*.

Analyses showed the relative expression of Omp85 members under various growth conditions and demonstrated the expression of novel genes, *bamK* and *bamL*, were detected at very low levels (Figure 3.1.2-3.1.4) or in experiments utilising promoter-fusions approaching the limit of detection (Figure 3.1.5-3.1.6). In a study published during the course of my PhD by Guilhen *et al.* (2016), the authors performed comprehensive transcriptional profiling on *K. pneumoniae* CH1034 to define 'global' gene expression signatures in different growth modes and phases. Transcription data from this study specifically showed the Omp85 gene levels (*bamA,* CH1034_120163; *tamA,* CH1034_6005; *bamK,* CH1034_60057; *bamL*, CH1034_10045) isolated from planktonically grown cells at exponential or stationary phase, reflected similar transcript expression trends with my qPCR and GFP reporter results.

In the absence of post-transcriptional regulation, transcript levels correlate with protein abundance levels (Maier *et al.*, 2009, Guimaraes *et al.*, 2014, Liu *et al.*, 2016). In a study by Brinkworth *et al.* (2015), the authors investigated OMPs and exoproteins that are upregulated in *K. pneumoniae* ST258. BamA was identified by liquid chromatography-tandem mass spectrometry at detectable levels in outer membranes prepared from *K. pneumoniae* ST258 cultured in nutrient rich LB media in both exponential and logarithmic phases of growth. Similarly, TamA was also identified in the outer membranes, but only in exponential phase cells cultured in LB media. BamK and BamL were not detected in this study. These results from this study mirror my qPCR and GFP-reporter results and shows a correlation in Omp85 transcripts and protein levels based on the results demonstrating that

the expression of *bamK* and *bamL* was not detected under various laboratory culture conditions. Future studies could employ more high-throughput screening for other environmental conditions that might induce expression such as broad-spectrum phenotype microarray analysis which includes characterisation by metabolic capabilities or even antimicrobial susceptibility (Shea *et al.*, 2012, Mackie *et al.*, 2014).

In a study by Kim *et al.* (2012), the authors found through chromatin immunoprecipitation combined with microarray hybridization (ChIP-chip) assays, the consensus −35 and −10 DNA binding motifs for σ^{70}-dependent (i.e. RpoD) promoters in *K. pneumoniae* MGH78578 were TTGACA and TATAAT, respectively. These consensus motifs were highly similar to the identified sequence motifs found in *E. coli* and *Salmonella* Typhimurium (Typas *et al.*, 2007, Kroger *et al.*, 2012). Furthermore, the authors did not report binding of RpoD near the *bamK* locus and its putative promoter region, consistent with a low or no transcription scenarios. Interestingly, the predicted "−35 and −10"-like sequence motifs (TTCA**G**A TATA**T**T, respectively) of *bamK* from *K. pneumoniae* B5055 were near identical suggesting regulation pathways of sigma factors is conserved not only in an intra-species manner but also an inter-species manner. It should also be noted, unlike other BAM components controlled by the σ^E regulon, *bamK* had a consensus motif for σ^{70} (RpoD) therefore its expression is most likely regulated by a different set of environmental cues, cell stresses and regulatory cofactors.

Promoter analysis also highlighted two regions in the putative *bamK* promoter that could be negatively modulating *bamK* expression, a stem-loop structure and an AT-rich region. In prokaryotes the base-pairing of single stranded DNA has been reported to form secondary structures such as hairpin and stem-loop structures, and have been reported to be involved in recombination, replication and transcription (Henkin & Yanofsky, 2002, Berg *et al.*, 1991, Jagodnik *et al.*, 2017). A predicted stem loop structure overlapped the −10-like/TATA-box sequence potentially inhibiting or stalling transcription. However, through promoter activity assays, I demonstrated that abolishing the stem loop structure did not increase promoter activity of *bamK* compared to its native sequence (Figure 3.2.2). Histone-like nucleoid-structuring (H-NS) protein is a global transcriptional regulator and its function is usually to repress horizontally acquired genes (based on its higher AT-content) by binding and silencing of transcriptional regions (Lucchini *et al.*, 2006, Navarre *et al.*, 2007). The regulation of H-NS has been linked to temperature, osmolarity and metal ion levels (Ono *et al.*, 2005, Amit *et al.*, 2003). Unlike classical regulatory repressor and activator proteins, H-NS does not recognise specific nucleotide sequences but is inclined towards curved DNA regions (Owen-Hughes *et al.*, 1992, Yamada *et al.*, 1991). Due to inherent AT-rich regions the putative promoter region of *bamK* was predicted to be highly curved (Figures 3.2.3) and was therefore suspected to be regulated by H-

NS. Through qPCR assays, I found that in a Δ*hns* strain *bamK* transcript levels at below detectable levels (Figures 3.2.5). It should be noted that these assays were performed in *K. pneumoniae* AJ218 and not *K. pneumoniae* B5055; it was assumed regulation between the two strains would be identical as the upstream elements (500bp) was identical and the *bamK* CDS was 99.75% identical. Immunoblot analysis confirmed that upon changing the *bamK* promoter to a known constitutive promoter (i.e. *bamA*), a functional protein was able to be detected (further explored in Chapter 5). From these analyses pertaining to the putative *bamK* promoter, I hypothesised that the upstream region of *bamK* contained elements maintaining transcriptional repression despite the open-reading frame encoding a functional Omp85 protein.

Transposon mutagenesis is one of several high-throughput strategies for discovery of virulence genes, essential genes and even functional gene associations in *K. pneumoniae* (Wilksch *et al.*, 2011, Tomás *et al.*, 2015). In this study, I generated a transposon mutagenesis library into a genomically modified *K. pneumoniae* B5055 strain. The B5055 (CIP 52.145) strain is peculiar, as it one of the few β-lactamase-negative strains *K. pneumoniae* (Bialek-Davenet *et al.*, 2014). Thus, as a screening strategy I developed a system where the *bamK* coding region was precisely replaced with a β-lactamase gene (Amp^R) resulting in strain *K. pneumoniae* B5055 Δ*bamK*::Amp^R (Figure 3.3.3-3.3.4). If *bamK* was expressed due to a transposon insertion disrupting an associated regulatory factor, an ampicillin-resistant phenotype would reflect a gain-of-function mutation. Generation of transposon mutants was shown to generate false-positives as indicated by (i) the frequency of gain-of-function mutants, (ii) the identity of disrupted genes (Table 3.3.1) and (iii) the consequent promoter-fusion assays (Figure 3.3.7). I suspect the gain-of-function mutants were the result of spontaneous mutations conferring an ampicillin-resistance phenotype. Like many other successful bacterial pathogens, *K. pneumoniae* has evolved to acquire a suite of antibiotic resistance mechanisms (Lee *et al.*, 2017, Khan *et al.*, 2018). In many *K. pneumoniae* clinical isolates, modulation of its membrane permeability by mutational loss of its major porin(s) diminishes permeability for antibiotics (Hernandez-Alles *et al.*, 1999, Sugawara *et al.*, 2016). In one such study by Tsai *et al.* (2011), the authors found that several clinical isolates of *K. pneumoniae* are deficient of major porins, OmpK35 and OmpK36, and found that the loss of these genes correlated with reduced susceptibility to antibiotics. A different antibiotic resistance mechanism also reported in *K. pneumoniae*, is the hyper expression and/or activation of efflux pump systems which expel a broad range of antibiotics (Padilla *et al.*, 2010, Srinivasan *et al.*, 2014). In a study by George *et al.* (1995), the authors found that mutations in the positive transcriptional activator *ramA*, or its regulatory regions, correlated with the upregulated expression of the AcrAB efflux pump. This was further substantiated in future studies (Ruzin *et al.*, 2008, Bratu *et al.*, 2009, Rosenblum *et*

al., 2011, De Majumdar *et al.*, 2015), which showed overexpression of *ramA* or clinical isolates with transcriptional activator mutations resulted is ubiquitous in several multidrug resistant *K. pneumoniae* isolates. Taken together, my data and the aforementioned studies suggest that the antibiotic selection strategy utilised may have induced compensatory mutations in the engineered *K. pneumoniae* B5055[Rif] Δ*bamK*::Amp[R] strain resulting in the emergence of false-positive transposon mutants. While it would be interesting to investigate how the β- lactamase-negative *K. pneumoniae* B5055 strain was able to acquire a mechanism for ampicillin resistance, future studies will instead need to utilise alternative antibiotic selection or even a different reporter system such GFP, luciferase and β-galactosidase.

Several previous studies have used mass-spectrometry (MS) to identify transcriptional regulators from DNA-protein binding studies (Hellman & Fried, 2007, Jiang *et al.*, 2009). To that end, I attempted to identify DNA-binding proteins binding to the putative promoter of *bamK*, proteins which may be modulating its transcriptional expression, by combining EMSA coupled with MS analysis. Using *K. pneumoniae* B5055 cleared cell lysate as a source of prey protein, I was able to identify a shifted complex DNA::protein complex which represented protein(s) bound to the an amplified to a putative *bamK* promoter DNA fragment (Figure 3.4.2). Furthermore, I was also able to also show the most pertinent DNA-binding domain in the putative *bamK* promoter was most likely the predicted highly curved DNA region as shown by EMSA which utilised truncated promoter regions as bait (Figures 3.4.3). Accessory roles of the other promoter elements (i.e. −35 and −10-like consequence motifs and RBS site) should not be excluded as these other domains may provide additional binding interfaces for stabilisation, or for other cognate transcriptional factors. Proteomic analysis of the DNA::protein complex from the EMSA revealed a preliminary list of proteins that could be involved in binding (Appendix 3). Since only few proteins were expected to be true transcriptional regulators of the *bamK* gene, I further refined the list to 5 protein DNA-binding candidates based on their suspected roles in gene regulation (Table 3.4.1). From this refined list, candidate gene regulators *dnaN* and *ssb* were discounted based on their previously reported essentiality in prokaryotic genome maintenance (Baba *et al.*, 2006, Goodall *et al.*, 2018). Candidate *hsdR*, which encodes a component of the Type I restriction-modification system, was also dismissed as a transcriptional regulator. The Type I restriction-modification system is a prokaryotic defence mechanism which cleaves foreign DNA (Loenen *et al.*, 2014), therefore binding of this subunit could occur based on the fact the bait substrate utilised in the EMSA was an amplified PCR product not synthesised from a native *K. pneumoniae* system. Of the remaining candidates, *dksA* and *hupA*, I generated single-gene deletions in *K. pneumoniae* B5055 resulting in the strains *K. pneumoniae* B5055 Δ*dksA*::Kan[R] and

$\Delta hupA$::Kan^R. The RNA polymerase-binding transcription factor DksA has been reported previously to be involved in aspects of σ^E activation for gene regulation in response to intracellular signals related to growth phase and nutrient availability (Costanzo *et al.*, 2008). The other candidate, HupA, has been reported to be an abundant nucleotide associated protein which binds preferentially to AT-rich regions/highly curved regions for stabilisation reducing accumulation of mutations during stationary phase (Krylov, 2001, Wojtuszewski & Mukerji, 2003, Williams & Foster, 2007). Through promoter-fusion assays, it was found that the promoter activity of *bamK* was consistently low in either *K. pneumoniae* B5055 $\Delta dksA$::Kan^R or $\Delta hupA$::Kan^R deletion strains (Figures 3.4.6 A). As further validation, immunoblot analysis on whole cell lysates prepared from the same deletion strains revealed BamK was not detectable levels in either logarithmic or stationary phase (Figures 3.4.6 B). Together, these data suggest *bamK* transcription and protein expression is not modulated solely by either $\Delta dksA$ or $\Delta hupA$ under the conditions tested.

Hierarchical control between genetic regulatory systems has been reported in Gram-negative bacterial species and allows the fine-tuned control for protein expression in respect to specific stress responses and adaptation to specific environments (Madan Babu *et al.*, 2007, Ishihama, 2010, Flores-Kim & Darwin, 2014). Curli biogenesis is one such example found in *E. coli*: a process that describes the production of cell-surface amyloid structures critical for biofilm formation and surface adhesion (Barnhart & Chapman, 2006). The curli biogenesis system is encoded by the *csg* operon with *csgD* being the master transcriptional regulator (Chirwa & Herrington, 2003, Brombacher *et al.*, 2006). In a study by Ogasawara *et al.* (2010), the authors reported an interplay between at least five transcription factors involved in the activation (OmpR, RstA and IHF) and repression (CpxR and H-NS) of the *csgD* promoter. Naturally, one could speculate that aside from functional redundancy of global gene regulators such as H-NS (Zhang *et al.*, 1996, Müller *et al.*, 2006), expression of *bamK* may involve transcriptional activators in a hierarchical cascade (i.e. repression supersedes positive activation). Hence, a more complex regulatory mechanism for *bamK* gene regulation cannot be excluded at this time. Another explanation of the tight regulation of *bamK* could be due to mRNA transcript instability. As replacement of the 5'region with a constitutively promoter resulted in BamK expression, sequence motifs within this region could be leading to its rapid decay. In prokaryotes, mRNA decay is facilitated by endonucleases which have specificity towards certain sequence motifs or secondary structures such as hairpins and stem loop structure (Moll, 2003, Carrier & Keasling, 1997). Detailed understanding of the gene regulation governing Omp85s are important, especially in genomes with diverse gene copy numbers, as these studies can provide better understanding how each

member contributes to bacterial outer membrane biogenesis allowing adaptation to a range of environments.

4

Results: Identification and functional characterisation of BamL

4.0 Introduction

In a global analysis of the Omp85 protein superfamily, a phylogenetic branch termed "noNterm" was identified as having a sporadic distribution through bacterial species and strains (Heinz & Lithgow, 2014). The noNterm Omp85 subfamily is characterised by a lack of typical N-terminal extensions commonly associated with a C-terminal Omp85 β-barrel domain. The functional role of the noNterm protein subfamily is unclear, as many studies have attributed N-terminal extensions, such as POTRAs as distinguishing and are typically vital for catalytic function (Bos *et al.*, 2007, Simmerman *et al.*, 2014). One of the bacterial strains identified with a member of the "noNterm" group was *K. pneumoniae* B5055. This *K. pneumoniae* Omp85 is hereafter referred to as BamL (Little). In *K. pneumoniae* genomes, I discerned that *bamL* is found in synteny with an upstream partner, hereafter referred to as *lupV*. This genetic organisation is intriguing as it shares similarities to the operon arrangements of two-partner secretion systems (TPS), a sub-class of Type V Secretion Systems (T5SS).

The various subtypes of the T5SS are related, but in seemingly disjointed ways: some subtypes such as T5SSb and T5SSe have Omp85/TpsB β-barrel domains in common, others have mechanistic similarities such as: autotransporters (T5SSa), trimeric autotransporters (T5SSc) and inverse-autotransporters (T5SSe) (Jacob-Dubuisson *et al.*, 2013, Guérin *et al.*, 2017). TPS systems are of the T5SSb subclass, and are comprised of two proteins, TpsA the secreted "effector" and cognate "transporter" TpsB. The transporter, TpsB, adopts a β-barrel structure which features N-terminal POTRA domain(s) though fewer than the archetypal Omp85 protein BamA (Jacob-Dubuisson *et al.*, 2009); but a clear distinction between the Omp85 family (e.g. BamA and TamA) and TpsB family (e.g. FhaC) subfamilies exist at the sequence level because they have evolved in distinct evolutionary trajectories (Heinz & Lithgow, 2014). In this chapter, I present results on the initial bioinformatic analyses, phenotypic assays and biochemical approaches undertaken to identify and initially characterise the function of the novel protein BamL to work towards understanding its function.

4.1 Genomic context and protein predictions of the *bamL* locus

A previous phylogenetic study of the Omp85 family revealed an uncharacterised T5SS/TPS-like operon in *K. pneumoniae* species (Heinz & Lithgow, 2014). Aiming to better understand these uncharacterised genes, I investigated its genomic context. In *K. pneumoniae* B5055, gene locus tags BN49_0006 (referred to as *lupV*) and BN49_0007 (referred to as *bamL*), are both annotated as "putative outer membrane proteins" based on sequence predictions β-barrel topologies of other known OMPs (Figure 4.1.1, A). This genetic organisation is analogous to TPS systems found in Gram-negative bacteria where a translocon Omp85/TpsB member is involved in the export or secretion of a β-helical TpsA (Chapter 1, Figure 1.3.2). Thus, BamL could be a translocon OMP involved in the secretion of LupV. The dual gene locus, henceforward termed the *bamL* locus, overlap by 4 bp in their predicted nucleotide sequences suggesting co-translational regulation. A brief examination looking at the genomic context of the locus shows direct neighbouring genes are not likely involved in OM biogenesis due to their distance and predicted functional roles derived from conserved motifs.
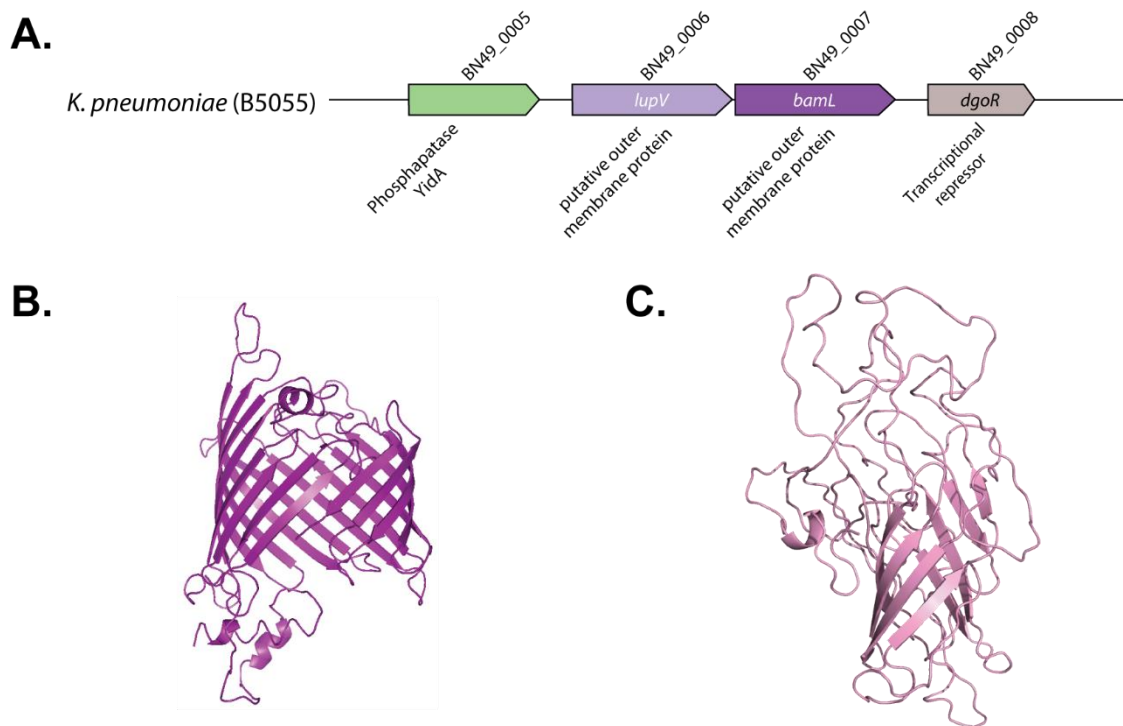
Figure 4.1.1: Genomic context and sequence predictions of the *bamL* locus. (A) Genomic context of the *bamL* locus in *K. pneumoniae* B5055. Putative outer membrane locus tags, BN49_0006 and BN49_0007 were designated as *lupV* and *bamL*, respectively. The *lupV* gene overlaps by 4 bp with *bamL* and might be a co-expressed transcription unit. (B) Using the protein prediction server Phyre2 (Kelley *et al.*, 2015), BamL shows structural similarity to β-barrel domains of BamA and TamA but lacks any N-terminal POTRA domains. (C) For LupV, Phyre2 predicts a disordered N-terminal region with no recognisable homologous protein folds followed by an 8-strand β-barrel similar to *E. coli* OmpA.

The prediction of tertiary structures through homology modelling is a common approach to better understand proteins of unknown function (Baker, 2001). A simplified run down of homology modelling can be the viewed as an iterative process where algorithms consider an input amino acid query sequence. The input sequence is then matched to solved protein structure(s) resulting in a structural model which takes into account sequence and fold alignment to solved structures and other biophysical constraints (Madhusudhan *et al.*, 2005). I generated homology models using protein prediction server Phyre2 (Kelley *et al.*, 2015). The predicted BamL structural models with high confidence (>95 %) as a 16-strand β-barrel with no POTRA domains based on its conserved sequence with solved crystal structures (Figure 4.1.1 B). Structural templates used for the BamL model included: TamA (PDB: 4C00), BamA (PDB: 4K3B, 4N75 and 4N75) and FhaC (PDB:2QDZ). For LupV, the N-terminal region (137 residues) had no detectable domains or available homologous structures but secondary structure prediction servers (PsiPred (McGuffin *et al.*, 2000) and DISOPRED (Ward *et al.*, 2004)) detected a disordered of stretch (24 residues) followed by a putative C-terminal β-barrel domain (>95 % confidence) modelled after: (i) OmpA (PDB: 3NB3, 1QJP), (ii) autotransporter domain of NalP (PDB: 1UYN) and (iii) attachment invasion locus (Ail) protein (PDB: 3QRA) (Figure 4.1.1 C). Validation of the homology modelling for BamL and LupV was by manually by inspecting the output structure and data value ouputs pertaining to confidence in the modelled protein regions and coverage of the query sequence. It should be noted the β-barrel domain structures of both LupV and BamL were modelled with >95 % confidence, but the structural templates utilised had only ~18 % or ~15 % sequence identity, respectively. It should be further noted, "confidence" does not represent the accuracy of the model but is more of a correlated measure that the output homology model is likely to adopt the template structures overall protein fold. To predict whether BamL or LupV were targeted to the periplasm, the protein sequences of BamL and LupV were analysed by SignalP 4.1 and LipoP 1.0 algorithms to determine their subcellular localisations (Petersen *et al.*, 2011, Juncker & Willenbrock, 2003). BamL was confidently suggested to localise to the Gram-negative OM. LupV was predicted to have a signal sequence, but this was not confidently assigned to a specific localisation as it contained characteristics of either an OMP or lipoprotein signal peptide. These data remain suggestive that the *bamL* locus encodes for a novel two-partner system or a completely novel outer membrane imbedded system of unknown function, where one element encodes for an Omp85.

Pseudogenes are segments of DNA which may have its original function deteriorated from accumulated mutations and in some cases being derived from gene duplication events (Mira & Pushker, 2005, Tutar, 2012). In most cases, pseudogenes can be differentiated from functional

genes based on their high prevalence rates and little sequence conservation (Balakirev & Ayala, 2003). Therefore, to examine if *lupV* or *bamL* represented pseudogenes I analysed their nucleotide conservation rates and gene prevalence on a globally diverse genomic dataset of *K. pneumoniae* isolates (Holt *et al.*, 2015b) and additional isolate genomes. A full list of genome accession tags (705 unique identifiers) can be found in Appendix 5. Using an in-house computational pipeline, I analysed the output data of the gene conservation and prevalence rates of *lupV* and *bamL* within a defined set of *K. pneumoniae* genomes. As shown in Figure 4.1.2 A, the average nucleotide conservation rate of *lupV* was found to be 97.55%, and *bamL* was at 98.64% in the analysed 705 genomes. When looking at the prevalence rates of both genes in the same analysed sequences, the co-prevlance of both *lupV* and *bamL* within the same genome was found in 561 genomes (out of 705, ~80%). A less prevalent instance was seen in 87 genomes with *lupV* without *bamL, bamL* without *lupV* in 53 genomes or 4 genomes had neither. However, it should be noted, the automated script excluded *bamL* or *lupV* sequences with any premature stop codons, poor sequence coverage and/or contig breaks (Appendix 5). While pseudogenisation could not be ruled out from the current data, the nucleotide conservation and co-prevalence are likely to be understated due to the scripts thresholds. So, *bamL* or *lupV* can be considered highly conserved at the nucleotide and gene co-prevalence levels.
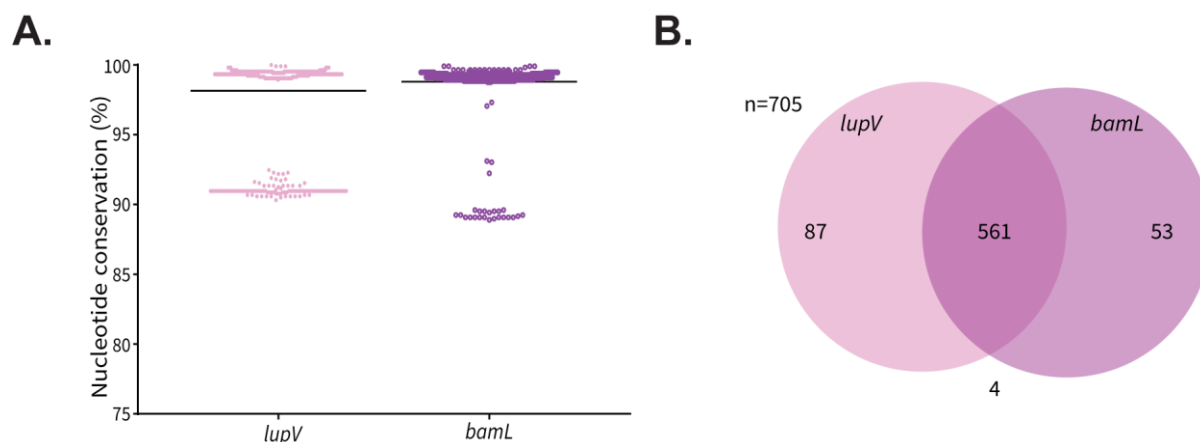
Figure 4.1.2: Nucleotide variation and gene prevalence rates of the *bamL* locus. Using the nucleotide sequences for *lupV* and *bamL* of *K. pneumoniae* B5055 as queries, data generated by an in-house PYTHON script (McIntyre et al., unpublished) analysed a globally diverse genome dataset of *K. pneumoniae* isolates (Holt *et al.*, 2015b) and additional *K. pneumoniae* genomes available from NCBI for their prevalence and nucleotide conservation rates. The list of sequence IDs genome assemblies (n=705) utilised for this analysis with accompanying annotations pertaining to premature stop codons, poor coverage or assembly contig breaks are listed in Appendix 5. (A) Scatter plot detailing the nucleotide variation of *lupV* (n = 648) and *bamL* (n = 614) from *K. pneumoniae* genomes compared to query sequences of *K. pneumoniae* B5055. Data points excluded were any sequences with additional stop codons (compared to the original B5055 sequences), poor coverage or contig breaks. Note the discontinuous Y-axis. (B) Venn diagram depicting the number of *K. pneumoniae* genomes, which harbour either *lupV* and *bamL* sequences (n=705). Some genomes included in the analysis had inherent issues (poor sequence coverage or broken contigs) and therefore may understate both nucleotide conservation and gene prevalence. Data generated by Mr. Liam McIntyre (The University of Melbourne, Australia)

4.2 Investigating the function of the *bamL* locus in *K. pneumoniae* B5055

In Gram-negative bacteria, OMPs are vital in the maintenance of cellular homeostasis, allowing, secretion of toxins, efflux of noxious antibiotics and uptake of nutrients (Jeeves & Knowles, 2015). As predicted *omp85* member *bamL*, and its linkage with *lupV* suggested a co-regulatory relationship exists between the two genes I hypothesised the two genes work together for their function. Therefore, to address whether the *bamL* locus in *K. pneumoniae* B5055 played a role in bacterial fitness, deletion mutagenesis was employed using the gene gorging method resulting in the strain *K. pneumoniae* B5055 Δ*lupV, bamL*::Kan[R] (Figure 4.2.1). This deletion strain retained its selectable kanamycin cassette and was confirmed by PCR analysis and sequencing, using primers flanking the targeted regions (Figure 4.2.2).
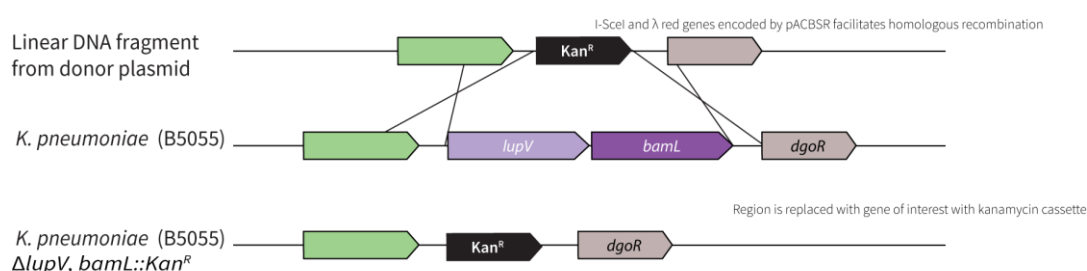
Figure 4.2.1: Strategy for *bamL* locus deletion in *K. pneumoniae* B5055. Schematic showing the replacement of the *bamL* locus coding sequence with a kanamycin cassette. The 'gene-doctoring' method utilises a two-plasmid system for genetic manipulation in broad range of bacterial species (Herring *et al.*, 2003). Briefly, a donor plasmid (plasmid 1) containing the kanamycin cassette with flanking homology regions is excised into a linear fragment by I-SceI by recombineering plasmid (plasmid 2) pACBSR. The recombineering plasmid pACBSR carries the λ-Red and I-SceI endonuclease genes, under the control of an *araBAD* promoter. The obvious advantage of this system is that multiple copies of the homologous DNA are present in the bacterial cell, which increases the number of potential recombination events.
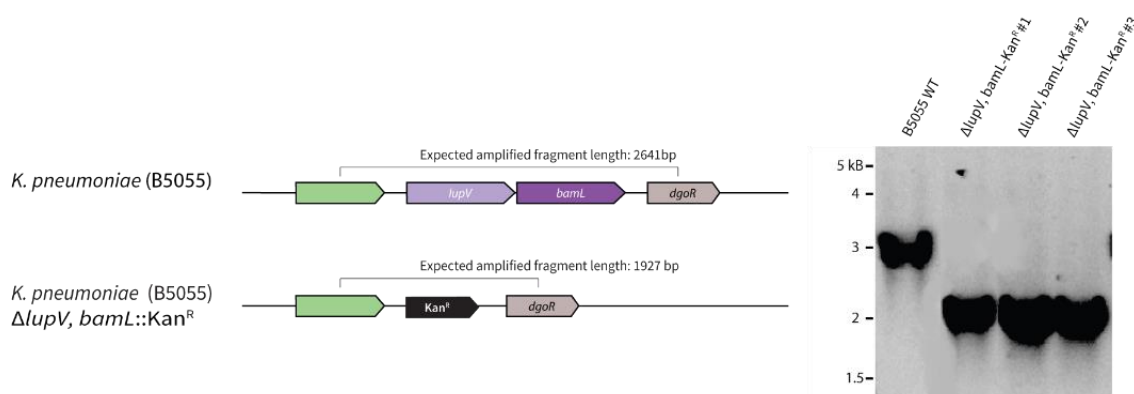
Figure 4.2.2: PCR confirmation of *bamL* locus deletion in *K. pneumoniae* B5055. Left, Cartoon showing the expected fragment lengths of *K. pneumoniae* B5055 before and after the kanamycin cassette integration (refer to Figure 4.2.1). Right, PCR amplification of *K. pneumoniae* B5055 before and after the after the kanamycin cassette integration replacing the *bamL* locus using primers, VVT086 and VVT087, which lie outside the region of genomic integration. PCR products were sequenced to confirm the locus deletion. PCR products were analysed by 1 % TAE agarose gel and sequenced to confirm the gene replacement.

A routine method to measure bacterial fitness is comparative growth curve analysis (Hall *et al.*, 2014, Dykhuizen & Dean, 1990). Figure 4.2.3 shows automated growth curves of *K. pneumoniae* B5055 cultures in select liquid media. In the tested media, no major differences

in growth kinetics were observed between the mutant Δ*lupV, bamL*::Kan[R] and its isogenic wildtype. In rich LB media, a small but reproducible decrease in fitness was observed. These results suggested that the deletion of *bamL* locus does not play a significant role in the planktonic mode growth of *K. pneumoniae* B5055 under standard laboratory conditions. However, the minor growth defect in the mutant Δ*lupV, bamL*::Kan[R] does suggest that BamL contributes to some aspect bacterial fitness for growth in rich media.
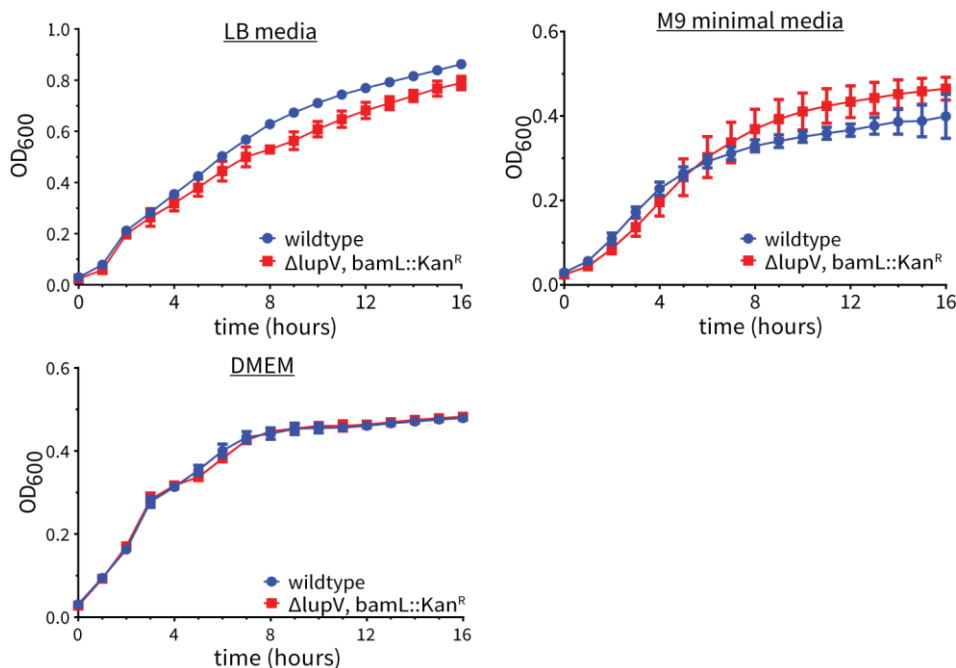
Figure 4.2.3: Effect of the deletion of *bamL* locus on growth kinetics in *K. pneumoniae* B5055. Automated growth curves comparing the parent B5055 strain to *K. pneumoniae* B5055 Δ*lupV*, *bamL*::Kan[R]. Growth kinetic profile trends were generally the same except in rich LB media where a mild growth defect was seen in the deletion strain compared to its corresponding parent strain (n=3, error bars represent SD).

To extend the phenotypic analysis of the Δ*lupV, bamL*::Kan[R] mutant, I tested whether the mutant strains could have an altered uptake rate by mouse derived bone marrow derived macrophages for subsequent phagocytosis. Macrophages were infected with *K. pneumoni*ae B5055 at two different multiplicity of infections (MOIs) (25 or 50) and observed over a 6-hour infection period. The number of phagocytosed *K. pneumoniae* bacteria was determined by osmotic lysis of macrophages and plating of serial dilutions of lysates at the indicated timepoints (Figure 4.2.4). Observed macrophage uptake rate increased over time for the wildtype and Δ*lupV, bamL*::Kan[R] strains. However, the macrophage uptake rate of the mutant Δ*lupV, bamL*::Kan[R] strain was found to be mildly reduced at both time points. Accompanying live cell imaging demonstrated that co-culturing of macrophages with *K. pneumoniae* B5055 strains did not influence the morphology or cell death of the macrophages (Figure 4.2.5). Altogether these results illustrate the genes encoded by the *bamL* locus do not play a major role

in macrophage uptake under the tested conditions, but this does not rule out some other role in virulence.
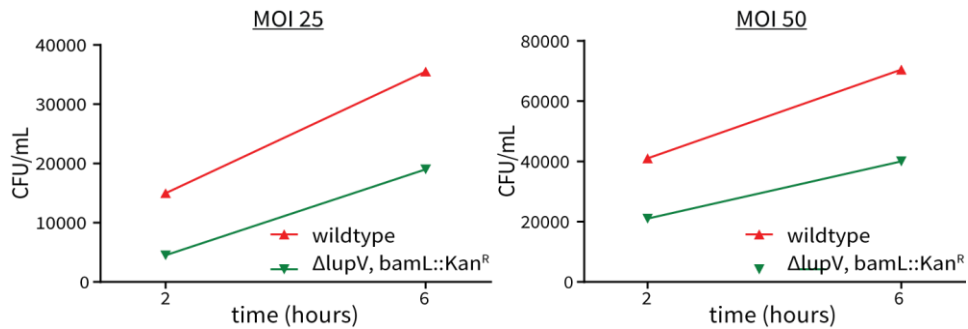
Figure 4.2.4: Effect of the deletion of *bamL* locus on growth kinetics and macrophage uptake rate in *K. pneumoniae* B5055. Macrophages were infected with *K. pneumoniae* B5055 for 30 minutes (MOI: 25 or 50). Cells were then washed and incubated with medium containing gentamicin (300 μg mL^{-1}) for 90 minutes to eliminate extracellular bacteria, and then monitored post-infection with medium containing gentamicin (100 μg mL^{-1}) for up to 6 hours. Intracellular bacteria were quantified by osmotic lysis, serial dilution and viable counting on LB agar plates. Performed by Dr. Seong Chow (Monash University, Australia).
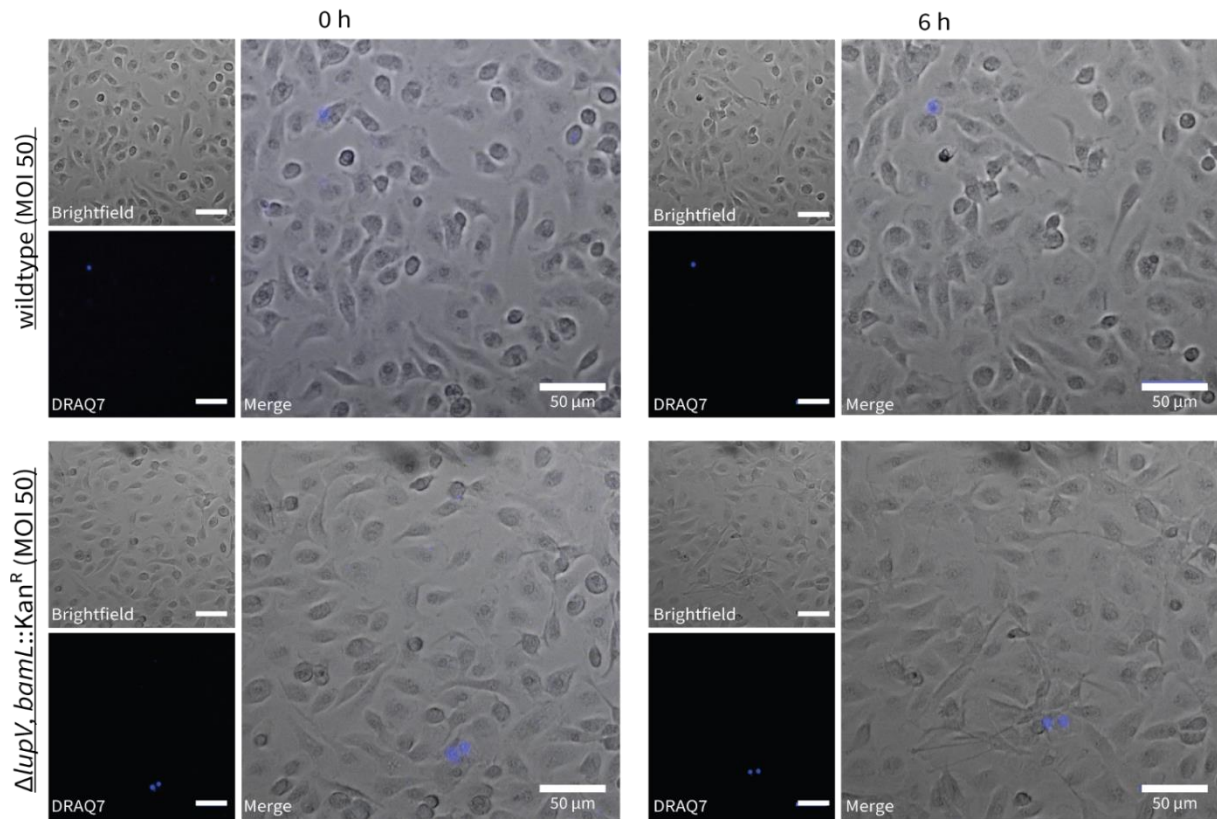
Figure 4.2.5: Live-cell imaging of macrophages infected with *K. pneumoniae* B5055. Time-lapse image snapshots of macrophages infected with *K. pneumoniae* B5055 wildtype or Δ*lupV, bamL*::Kan^R at 0 and 6 hours post-infection. Brightfield images shows unchanged morphology of macrophages during the 6-hour infection assay. DRAQ7™ is a far-red emitting marker that stains (blue) nuclei of dead and permeabilised macrophages. Macrophage viability was generally unaffected over the 6-hour infection assay.

4.3 Production of recombinant BamL

I sought to produce recombinant BamL for crystallisation and other structural analyses as numerous 3D structures of Omp85s have enabled better understanding of their function and mechanisms (Maier *et al.*, 2015, Gruss *et al.*, 2013, Fairman *et al.*, 2011), and so that the purified product could also be used for antibody production for assaying the localisation or processing of BamL under more native conditions. A recombinant BamL (expected molecular weight 43.2 kDa) was expressed in *E. coli*, total cellular membranes were isolated by ultracentrifugation, followed by detergent solubilisation of the membranes, affinity purification and finally size exclusion chromatography. A flow chart depicted in Figure 4.3.1 presents the process undertaken in this study for purification of recombinant BamL for X-ray crystallography and antibody production.
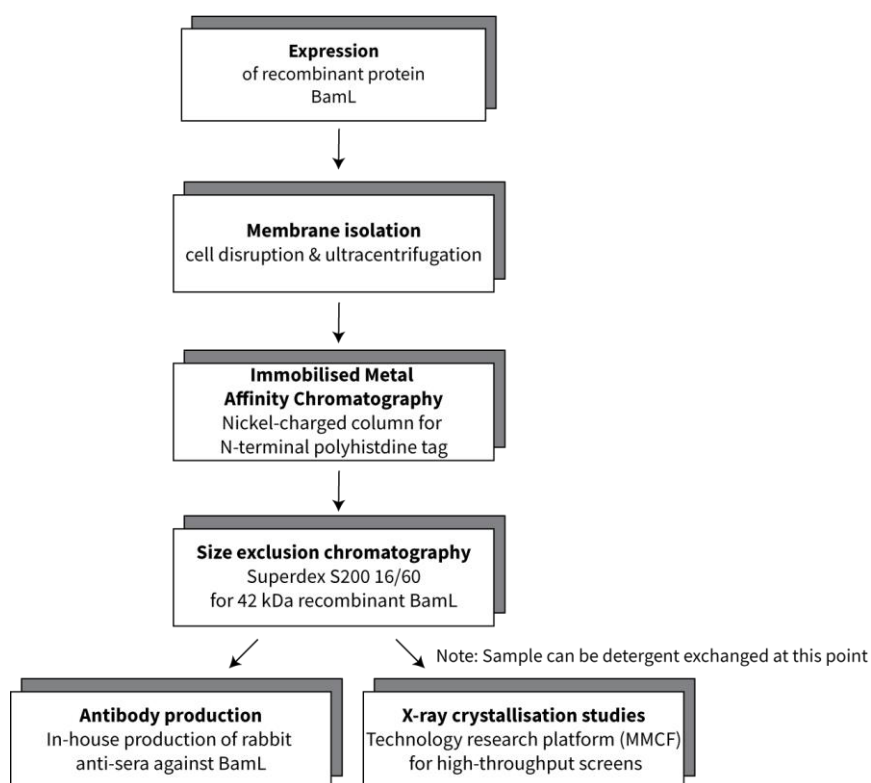
Figure 4.3.1: Purification flowchart of recombinant BamL of *K. pneumoniae* B5055 in an *E. coli* expression system. A flow chart illustrating the purification strategy for BamL. Expression of recombinant proteins were performed *E. coli* C41 (DE3), a strain routinely used for membrane protein overexpression (Dumon-Seignovert *et al.*, 2004). After expression, membranes containing BamL were solubilised, further purified firstly by immobilised metal affinity chromatography then size exclusion chromatography. The resulting purified product was used for antibody production and preliminary crystallisation trials. Predicted molecular mass and chemical/physical parameters were calculated using the online server ExPASy-ProtParam tool (Gasteiger *et al.*, 2003).

Exact details pertaining to: (i) culture and protein expression, (ii) total membrane isolation, (iii) detergent solubilisation and (iv) purification techniques and conditions are as described in Chapter 2.0. Briefly, a modified pET-20b (+) expression vector harbouring *bamL* without its native signal sequence was transformed into *E. coli* C41 (DE3) cells and cultured in TB media. Whole cell lysates were analysed by SDS-PAGE and then immunoblotted with αhexahistidine-HRP conjugate corresponding to BamL. Favourable expression conditions of BamL was found to be better at 0.1 mM IPTG induction for 3 hours upon reaching an initial cell density of OD_{600} = ~1.0. In comparison, the overnight induction condition resulted in a lower yield which echoed the pre-induced sample (Figure 4.3.2).
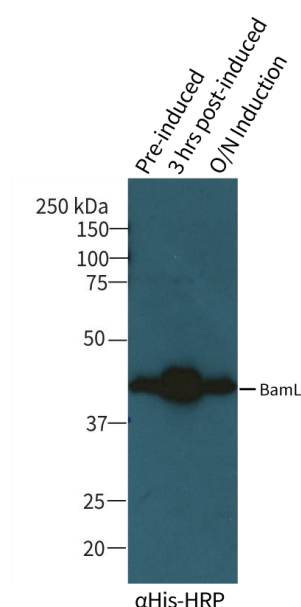
Figure 4.3.2: Small-scale test expression of BamL in *E. coli* C41 (DE3). Depicted here is an immunoblot detecting BamL with αhexahistidine-HRP. Whole cell lysates prepared from *E. coli* C41 (DE3). Pre-induced - sample taken before protein induction with a cell density of OD_{600} = ~1.0; 3hrs post-induced - induction of culture with 0.1 mM IPTG for 3 hours upon reaching a cell density of OD_{600} = ~1.0 and; O/N post-induced - induction of culture with 0.1 mM IPTG overnight upon reaching a cell density of OD_{600} = ~1.0. The overnight expression condition also included a lower growth temperature upon induction (i.e. 25 °C from 37 °C).

Large scale BamL production (i.e 6 litre cultures) employed the use of total membranes where proteins detergent solubilised and loaded onto a His-Trap™ HP column for affinity purification. A broad peak of proteins were eluted in fractions from the nickel affinity chromatography column using a linearly increasing imidazole gradient (0-1M) (Figure 4.3.3 A). Eluted fractions were analysed by SDS-PAGE and Coomassie blue staining (Figure 4.3.3 B) that showed an eluted ~42 kDA protein, likely corresponding to BamL
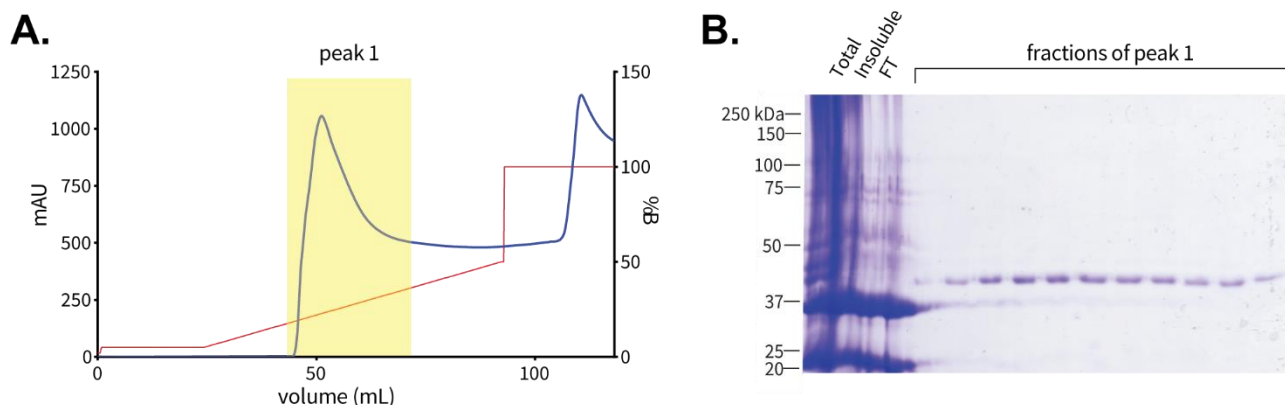
Figure 4.3.3: IMAC purification of BamL using nickel-charged His-Trap™ column. (A) Chromatogram illustrates the elution of a broad-shouldered peak eluted with a linearly increasing imidazole gradient from 0-500 mM then a step to 1M. (B) Analysed fractions of by SDS-PAGE (10%) and Coomassie staining of 'peak 1'. Included for comparison are: (i) total – the initial input solubilised membrane sample loaded onto the nickel-charged column. (ii) Insoluble – material pelleted after initial membrane solubilisation with Elugent. (iii). FT – the flowthrough material which did not bind to the nickel-charged column.

Size exclusion chromatography was employed to purify proteins based based on their mobility corresponding to their native size. The protein material used for SEC consisted of fractions eluted from the former IMAC step. In preparation for SEC, the IMAC fractions containing BamL were pooled and concentrated utilising a centrifugal filter with a 30 kDa cut-off. The salt concentration of the final samples was set to 150 mM through buffer exchange to suit the buffers utilised in SEC. The buffer composition used for SEC for BamL was: 50 mM Tris-HCl pH 7.5, 0.15 mM DDM and 150 mM NaCl After running the protein samples through SEC, chromatograms recorded the retention time of macromolecular species within the pooled IMAC samples (Figure 4.3.4). The shouldered peak elution pattern was further analysed by SDS-PAGE and Coomassie blue straining, where the centre fitted to pure and concentrated BamL with a of ~42 kDa.
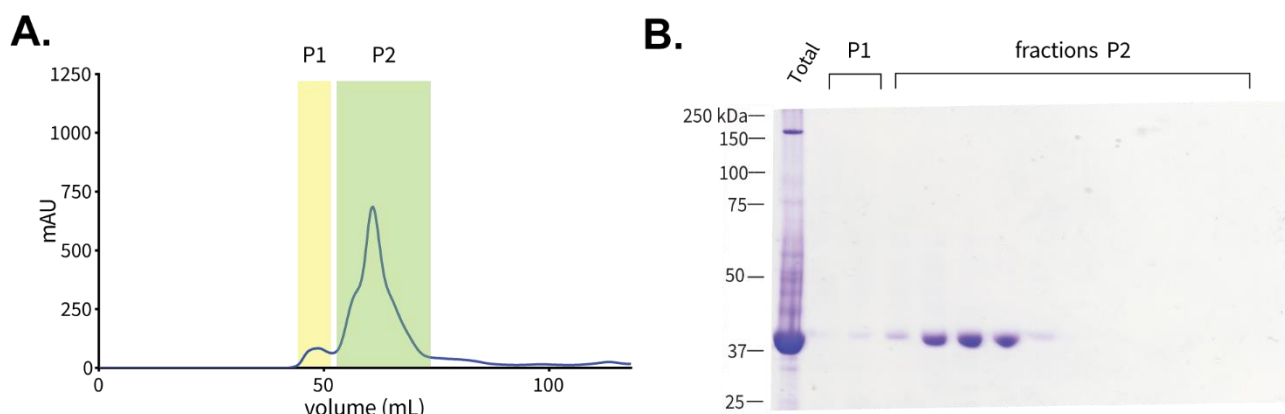
Figure 4.3.4: SEC purification of BamL using Superdex S200 16/60. (A) Chromatogram illustrates the retention time of two broad-shouldered peaks. (B) Analysed fractions of by SDS-PAGE (10%) and Coomassie staining of a single fraction of peak1 (P1) and fractions of peak 2 (P2). Included for comparison is: (i) Total – the sample input of pooled and concentrated fractions from IMAC.

100

Extraction of membrane proteins from biological membranes usually utilises detergents (Jeffery, 2016). The usage of specific detergents is an important consideration in membrane protein biology as some are better at stabilisation which is an important factor in sample homogeneity for crystallisation studies (Anandan & Vrielink, 2016). Therefore, for downstream crystallisation studies I performed a detergent exchange on purified BamL samples within a DDM micelle to C_8E_4. The detergents DDM and C_8E_4 are routinely used for membrane protein crystallisation studies. For example, in a study by (Noinaj *et al.*, 2015), the detergent C_8E_4 has even been used for crystallisation of a truncated BamA from *H. ducreyi,* a protein homologous to BamL. Shown in Figure 4.3.5, BamL was detergent exchanged to a SEC buffer containing C_8E_4 (10 mM) from DDM (0.15 mM) through affinity purification using an Ni-NTA resin. The eluted protein sample was therefore of higher protein concentration, in a smaller volume, as the imidazole concentration necessary for elution (500 mM) was already established from Figure 4.3.2.
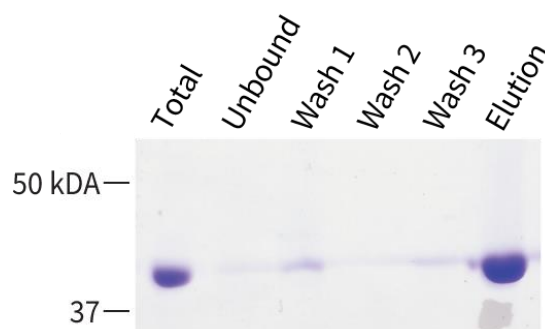
Figure 4.3.5: SDS-page analysis for detergent exchange of BamL from DDM to C_8E_4. Coomassie stained SDS-PAGE (10%) showing the detergent exchange of BamL from a DDM (0.15 mM) detergent SEC buffer to elution in (10 mM) C_8E_4. Some loss of protein can be observed in the unbound and wash fractions, potentially due to the rapid detergent exchange causing aggregation.

4.4 Antibody production for BamL

In Section 4.3, I expressed and purified BamL to generate antibodies to better understand BamL in the context of localisation studies and other biochemical characterisation assays. After immunisation rabbit antiserum was harvested. A pre-immune bleed from the same rabbit was used as a control. Specificity and sensitivity of the anti-serum was tested by western blot analysis using whole lysates of *E. coli* C41 (DE3) expressing or not expressing recombinant BamL. As shown in Figure 4.4.1 A, a ~42 kDA protein corresponding to BamL was observed in all anti-sera but not the pre-immune serum. The specificity of the anti-sera was problematic, particularly the TB2 and TB3 bleeds, where detection of non-specific proteins were also observed in the control whole cell lysates of the *E. coli* that did not express BamL. In a follow up immunoblot assay, duplicate samples were instead detected αhexahistidine -HRP, which revealed a discrete ~42 kDA band corresponding to the smallest of the 3 proteins in TB1 bleed. Altogether, these results indicate a working BamL antibody was generated, but that only the TB1 bleed would be of utility. The identity of the ~45 kDA or ~ 48 kDa forms of BamL are not known, but they may represent post-translational forms, and do not contain the N-terminal polyhistidine tag.
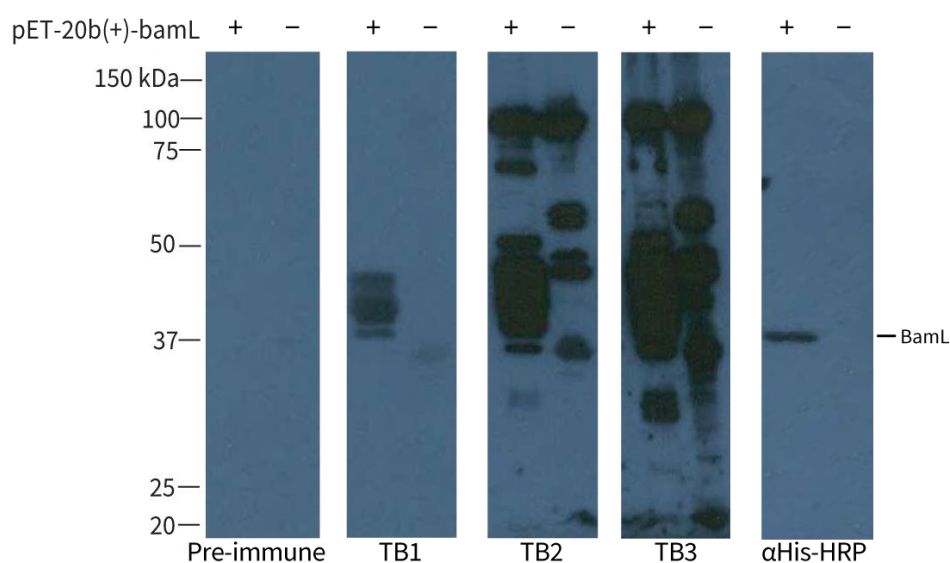
Figure 4.4.1: Evaluating the sensitivity and specificity of rabbit polyclonal anti-sera raised against BamL. Depicted here are immunoblot detections for BamL in *E. coli* C41whole cell lysates. Pre-immune refers to anti-sera harvested before challenging the rabbit with BamL. TB1 refers to 'test bleed 1' harvested at 4 weeks, T2 refers to 'test bleed 2'harvested at 7 weeks and T3 refers to 'test bleed 3' harvested 10 weeks all post-immunisation. The terminal bleed for complete anti-serum exsanguination was conducted at TB3. As the antigen still retained its polyhistidine tag, duplicate samples were detected with αHis-HRP corresponding with a discrete ~42 kDa band presumably BamL. Included as a comparison, whole cell lysates of *E. coli* C41 (DE3) were used to identify any non-specific protein contaminants carried through from the purification protocol of BamL. Samples were analysed by SDS-PAGE (12%).

4.5 Molecular characterisation of the BamL and LupV

As generation of a highly specific polyclonal BamL antibody for downstream assays was not achieved (Section 4.4), I instead designed an epitope-tagging strategy for localisation studies of LupV and BamL. A dual expression vector was designed where *bamL* and *lupV* featured a C-terminal Strep-Tag II or His$_6$-Tag, respectively. The tagged genes were cloned into an arabinose-inducible expression vector (pBAD24) with C-terminal tags, whilst both retaining their predicted native signal sequences (pBAD-bcLV, Figure 4.5.1). These tagged proteins of the *bamL* locus will now be termed LupV-ST and BamL-H6.
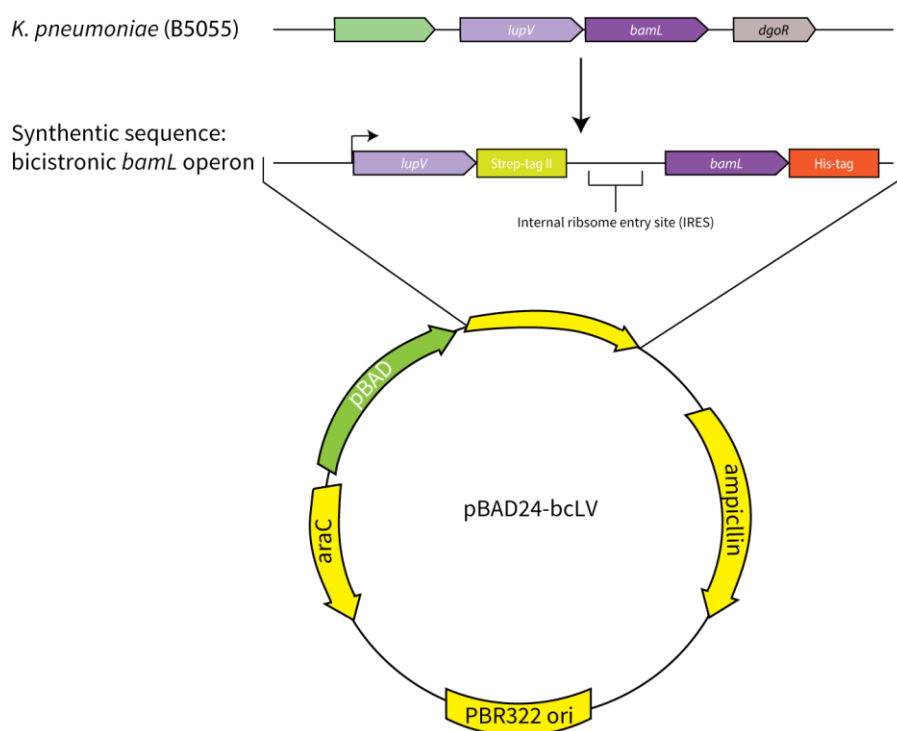
Figure 4.5.1: Cartoon diagram showing the vector map of pBAD24-bcLV which harbors an altered *bamL* locus insert. To produce recombinant *bamL* locus proteins for downstream analysis, a synthesised nucleotide gBlock (IDT) separating the two genes was ligated into pBAD24. The resultant plasmid now theoretically encoded for C-terminally tagged LupV (LupV-ST) and BamL (BamL-H6) under an arabinose inducible promoter. Gene overlap between *lupV* and *bamL* suggested a bicistronic transcription resulting in a single mRNA; thus, an internal ribosome entry site was also introduced to allow co-expression of LupV-ST and BamL-H6.

To detect the recombinant proteins of interest in downstream localisation studies, I first performed a test expression of the pBAD24-bcLV construct transformed in its native host species *K. pneumoniae* B5055. The detection of protein in the whole cell lysates was determined by western-blot, using the αhexahistidine-HRP conjugate for BamL-H6 or Precision Protein™ StrepTactin-HRP conjugate for LupV-ST. As shown in Figure 4.5.2, test expression conditions varying the concentration of inducing agent were used to evaluate relative yields of LupV-ST and BamL-H6. Log-phase *K. pneumoniae* B5055 cultured in LB

media induced with a final concentration of 0.05 % arabinose, incubated for 3 hours at 37 °C was selected for further downstream studies based on expressed protein levels, cell viability and processing duration.
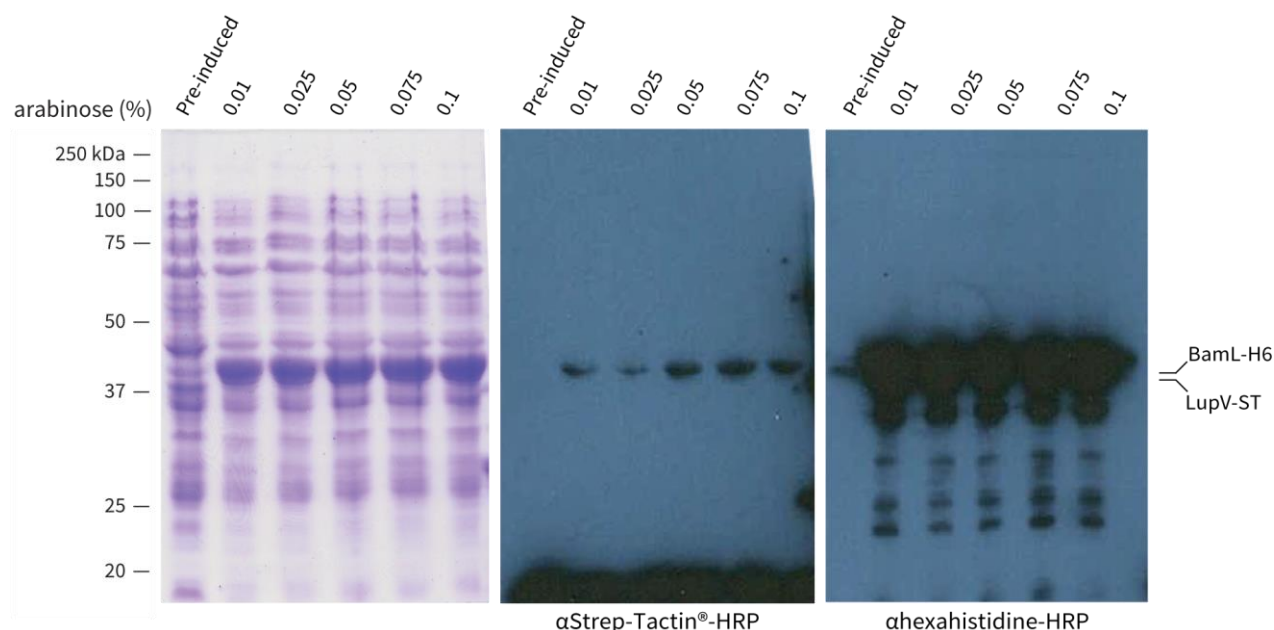
Figure 4.5.2: Test expression of tagged *bamL* locus proteins in *K. pneumoniae* B5055. *K. pneumoniae* B5055harbouring pBAD-bcLV were grown to mid-log phase then induced with arabinose at different concentrations (0.01 – 0.1% (w/v)) for 3hr. Cells were harvested and LupV and BamL were visualised by analysing the whole cell lysates by SDS-PAGE (12%) then by immunoblotting with the indicated HRP-conjugate antibodies (1:2000 dilution). It was found that 0.01-0.05% (w/v) of arabinose induction was just as effective as 0.1% (w/v). Expected molecular weights, LupV-ST (39.4 kDA) and BamL-H6 (42.2 kDA), determined by ExPASy-ProtParam tool (Gasteiger *et al.*, 2003).

To investigate if LupV was an effector/exoprotein, I analysed the culture supernatant for its secretion. Recombinant *bamL* locus proteins were expressed under the same conditions as previously, but instead the resulting culture supernatant was immunoblotted for secretion of LupV-ST using Precision Protein™ StrepTactin:HRP conjugate antibody. Under the tested growth conditions, recombinant LupV was not detected in the culture supernatants even after protein sample enrichment using centrifugation filters or TCA precipitation (Figure 4.5.3). These results suggest that LupV is not secreted into the supernatant and is most likely situated in a subcellular localisation as indicated by the previous immunoblot detecting whole cell lysates expressing LupV-ST and BamL-H6 (Figure 4.5.2).
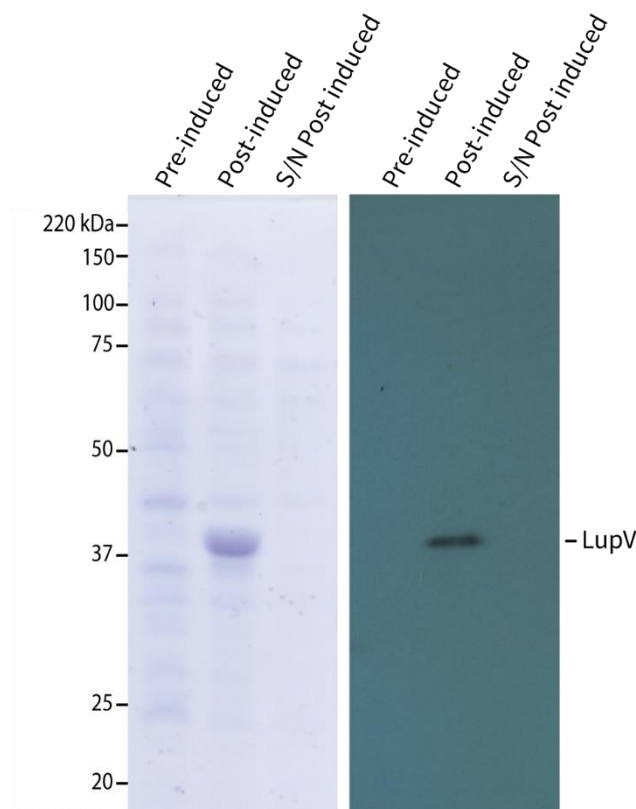
Figure 4.5.3: LupV is not secreted into the culture supernatant of *K. pneumoniae* B5055. *K. pneumoniae* B5055 harbouring pBAD-bcLV were grown to mid-log phase then induced with arabinose 0.05% (w/v) for 3hr. Whole cell lysates for the pre-induced, post-induced and post-induced culture supernatants were analysed by SDS-PAGE (12%) followed by immunoblot detection with Precision Protein™ StrepTactin:HRP conjugate. As the secretion of exoproteins/recombinant proteins could be relatively low, hence culture supernatants were enriched for proteins via TCA-precipitation or using centrifugation filters (molecular weight cut off 5kDa). The above Coomassie staining and immunoblot is a representative result for culture supernatant enriched by TCA precipitation.

Isopycnic sucrose gradient centrifugation is a technique that can reproducibly separate the phospholipid bilayer membrane components of Gram-negative bacteria into defined sucrose density fractions based on their buoyant densities (Dunstan *et al.*, 2017). Utilising the expression conditions of recombinant LupV-ST and BamL-H6 (Figure 4.5.2), total membranes were isolated from cells expressing LupV-ST and BamL-H6 and were subjected to isopycnic sucrose gradient centrifugation to separate the inner and outer membranes. Included in the assay are known membrane marker proteins F1β (inner membrane) and BamA (outer membrane). As shown in Figure 4.5.4 A-B, SDS-PAGE analysis followed by Coomassie staining shows separation of major OMPs (e.g. OmpA, OmpK35 and OmpK36) into heavier, fractions which correspond to fractions of higher sucrose density. Total membranes isolated from *K. pneumoniae* B5055 harbouring pBAD24-bcLV (Figure 4.5.4 B) contained distinct protein bands in the heavy sucrose fractions which were not see in the control sample (Figure 4.5.4 A). To investigate if these protein bands were LupV-ST or BamL-H6 I performed immunoblot analysis on the sucrose gradient analysed fractions show in Figure 4.5.4 B.

Immunoblotting for LupV-ST (Strep-Tactin®-HRP) and BamL-H6 (αhexahistdine-HRP) found that both targets localised to the outer membrane based on their co-elution in fractions with the BamA OM control (Figure 5.3.2).
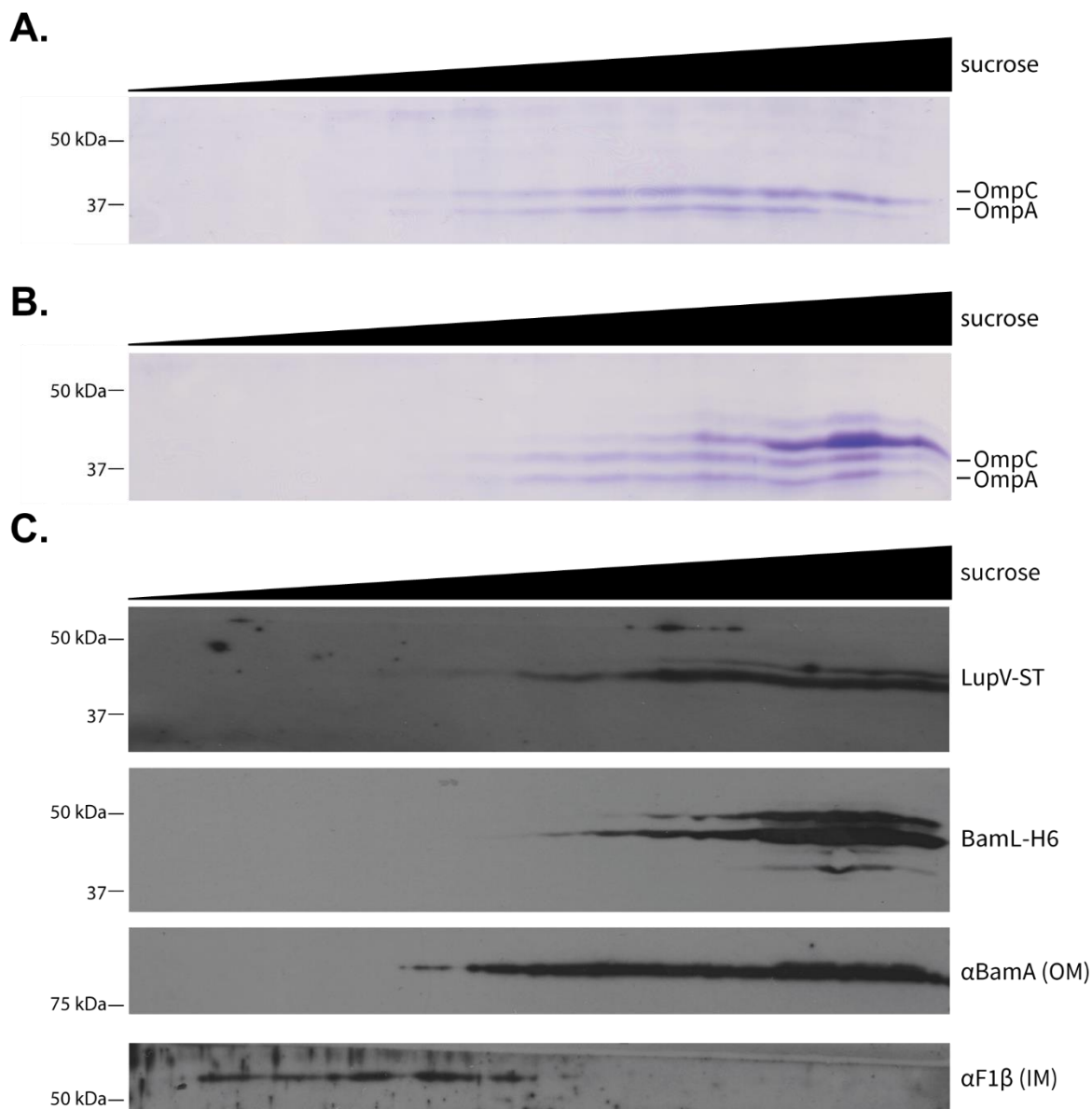
Figure 4.5.4: Subcellular localisation studies via isopycnic sucrose gradient centrifugation of tagged *bamL* locus proteins in *K. pneumoniae* B5055. (A) SDS-PAGE (12%) analysis of the sucrose gradient elution profile separating the inner and outer membranes from prepared *K. pneumoniae* B5055 harbouring empty pBAD24 as a control (B) SDS-PAGE (12%) analysis of the sucrose gradient elution profile separating the inner and outer membranes from prepared *K. pneumoniae* B5055 harbouring pBAD-bcLV. Migrating above the abundant porin bands was a species not observed in the control fractionation. (C) The same fractions from (B) were immunoblotted with the indicated HRP-conjugate antibodies. The recombinant proteins BamL-H6 and LupV-ST comigrated with OM protein marker - BamA.

To further characterise BamL and LupV, I assessed their association levels and correct assembly to the OM by performing a urea extraction wash on total membranes which contained LupV-ST and BamL-H6. To differentiate between integral and peripherally associated membrane proteins, 5M urea can be used to diminish relatively weak and exposed hydrophobic interactions and thereby removes them from membranes (Okamoto *et al.*, 2001). Correctly assembled integral membrane proteins are usually highly resistant to extraction by urea (Voulhoux *et al.*, 2003, Collin *et al.*, 2007). As shown in Figure 4.5.5, even after urea extraction treatment both LupV-ST and BamL-H6 were still found in the membrane pellet (P) fractions. The peripherally bound inner membrane F1β was extracted into the supernatant (S) fractions, which was included as a control. The resistance to urea suggested that both proteins were correctly assembled and integrally associated to the OM of *K. pneumoniae* B5055.
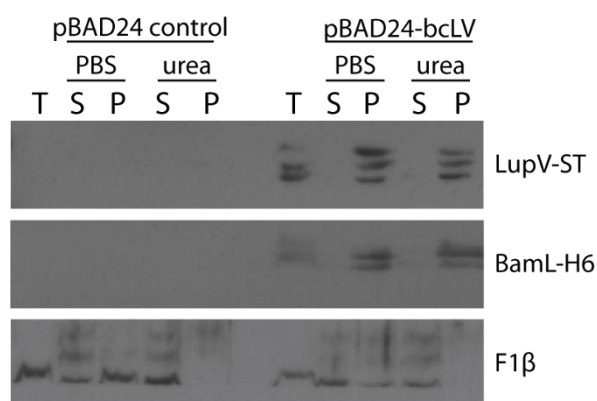
Figure 4.5.5: Urea extraction of tagged *bamL* locus proteins from *K. pneumoniae* B5055 membranes. Total membranes from *K. pneumoniae* B5055 harbouring either pBAD24 (empty vector) or pBAD-bcLV were treated with 5M urea or PBS and analysed by SDS-PAGE (12%) for subsequent immunoblotting with the indicated HRP-conjugate antibodies. 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).The doublet/triplet bands could represent partially denatured forms of LupV-ST or BamL-H6.

The BamL protein is a member of the Omp85/TpsB superfamily and may play a role in assembly and insertion of its cognate locus partner LupV into the OM and hence an interaction was hypothesised. To test this, I performed affinity purification to assess if interactions occur between LupV-ST and BamL-H6. Assays were performed using total membranes isolated from *K. pneumoniae* B5055 cells expressing recombinant LupV-ST and BamL-H6 as bait. The bait proteins and any ensuing binding (prey) proteins were purified using affinity resins: Strep-tactin® and Ni-NTA beads. In the affinity purification experiments which utilised Strep-tactin® beads, affinity binding of recombinant LupV- ST was found predominantly in the unbound elution fractions (Figure 4.5.7, A). This suggests binding did not occur between LupV-ST and the Strep-tactin® beads under the tested conditions. In affinity purification experiments which utilised Ni-NTA resin, the binding of the cognate recombinant protein

BamL-H6 did occur but co-elution of expectant LupV-ST was not observed. For Ni-NTA resin-based affinity purifications, recombinant BamL bait protein from prepared total membranes was able to bind and elute from the resin under the tested conditions however LupV was not detected in the final elution (Figure 4.5.7, B). Results from these affinity purification experiments indicate binding does not occur under the tested conditions or could be simply unobservable using this system if the protein interaction is of a weak/transient affinity. Other reasons could be the conformation of recombinant LupV could be forcing the C-terminal tag to be inaccessible wherein it is buried within the folds of the structure or the length of the affinity is too short preventing binding. Further optimisation of the binding buffer is another factor to consider as pH, temperature, salt concentration, presence of metal cofactors are all important for specific protein conformations and hence binding studies.
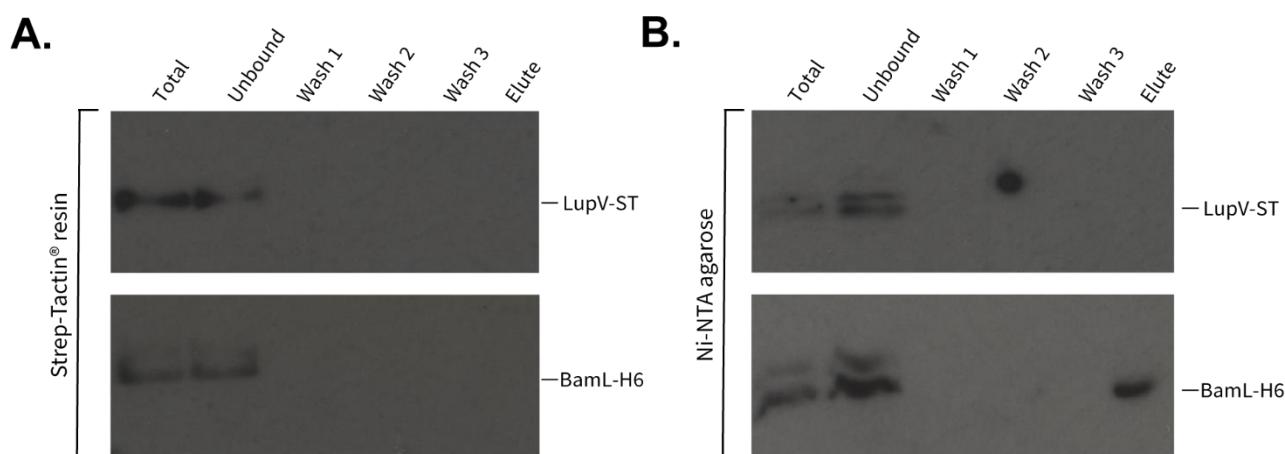
Figure 4.5.6: Affinity purification experiments against total membranes expressing LupV-ST and BamL-H6. (A) Total membrane extracts prepared from *K. pneumoniae* B5055 expressing recombinant LupV-Strep Tag II and BamL-His$_6$ was pulled down with Strep-Tactin® resin, followed by detection with the indicated antibodies after SDS–PAGE (12%). It was found that target bait protein, LupV-Strep Tag II, did not bind to the resin under the tested conditions (B) Total membrane extracts prepared from *K. pneumoniae* B5055 expressing recombinant LupV-Strep Tag II and BamL- His$_6$ was pulled down with Ni-NTA agarose followed by detection with the indicated antibodies after SDS–PAGE (12%). The target binding of BamL- His$_6$ was able to bind to the agarose beads but co-elution of LupV-Strep Tag II was not observed.

4.6 Preliminary crystallisation studies for BamL

Using commercial protein crystallisation screens JCSG-plus™ and PACT, a total of 192 crystallisation conditions were screened by the hanging-drop, vapor-diffusion method using either BamL (6 mg/mL) samples in a DDM or C_8E_4 micelle. From the 192 crystallisations conditions tested, a majority resulted in clear drops (Figure 4.6.1 A) suggesting the protein did not reach saturation and thus could not nucleate into favourable crystal lattice structures (Luft *et al.*, 2011). In other drops, precipitation of the protein ranging from brown (Figure 4.6.1 B), white and amorphous (Figure 4.6.1 C) blobs was observed or phase separation of the sample from precipitants (Figure 4.6.1 D). When comparing the commercial crystallisation screens (JCSG-plus™ and PACT), more precipitant-protein reactions occurred in the JCSG-plus™ whereas the PACT screen contained only clear drops in either DDM or C_8E_4 based buffers. Altogether, these data indicate that the purified BamL samples, from Section 4.5 at the utilised protein concentration and crystallisation conditions, did not result in positive crystal hits and thus requires further optimisation.
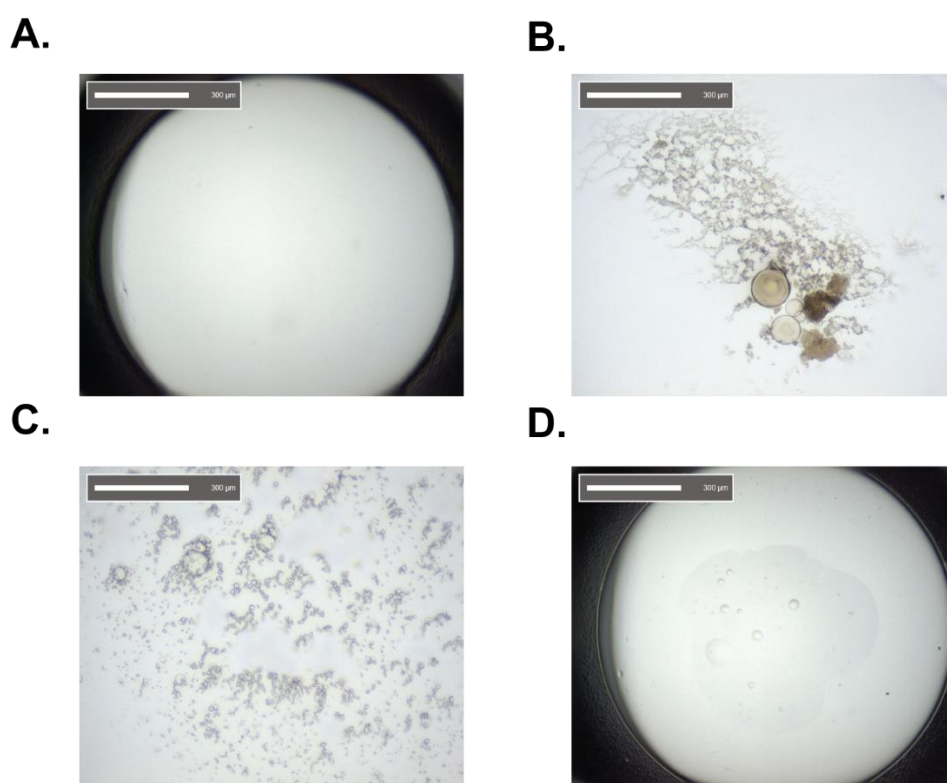
Figure 4.6.1: Selection of crystal drops results of BamL. Depicted here are select crystallisation drop results of BamL (100 nL BamL protein plus 100 nL reservoir) prepared by using the Mosquito pipetting robot (TTP LabTech). Crystal trays were incubated at room temperature over a 90-day screening period. (A) Clear drop: PACT condition A1, BamL (DDM micelle) (protein), 0.1 M SPG 4 pH (Buffer) and 25 % w/v PEG 1500 (Precipitant). (B) Brown protein precipitation: JSCG-plus™ condition E1, BamL (DDM micelle) (protein), 0.1 M Na Cacod. pH 6.5 (Buffer) and 1M Na_3 citrate (Salt). (C) White amorphous protein precipitation JSCG-plus™ condition G11, BamL (C_8E_4 micelle) (protein), 0.1 M BIS-TRIS pH 5.5 (Buffer) and 2 M $(NH_4)_2SO_4$ (Salt). (D) Phase separation: JSCG-plus™ condition E12, BamL (C_8E_4 micelle) (protein), 0.1 M imidazole pH 8 (Buffer) and 10 % w/v PEG 8K (Precipitant).

4.7 Discussion

The diversity of the Omp85 superfamily in pathogenic bacteria is widespread, with many implicated in essential cellular processes and virulence (Heinz & Lithgow, 2014). Using *K. pneumoniae* B5055 as a model, I investigated BamL, described previously as part of the "noNterm" phylogenetic branch of the Omp85 superfamily (Torres *et al.*, 2018). In this chapter, I show that linked genes *bamL* and *lupV* encode integral proteins in the outer membrane that do not form a classical secretion system in spite of its TPS-like genetic arrangement.

Although many genes encoding potential two partner secretion systems components have been identified through analysis of microbial genomes, only a limited number of these putative genes have been experimentally validated as TPS systems. Of those few characterised TPS systems, many have been shown to play a role in cell adhesion (Serra *et al.*, 2011), cytolysis (Palmer & Munson, 1995, Elsen *et al.*, 2014, Basso *et al.*, 2017) or iron acquisition (Cope *et al.*, 1994, Létoffé *et al.*, 1998) with newer functions linked to "contact-dependent growth inhibition" (CDI) systems where inter-bacterial interactions modulate bacterial growth (Aoki, 2005). The *bamL* locus of *K. pneumoniae* B5055 consists of two genetic elements, *bamL* and *lupV* (Figure 4.1.1), and shares a genetic arrangement akin to a canonical TPS format (TpsA-TpsB)(Jacob-Dubuisson *et al.*, 2013) (Guérin *et al.*, 2017). Sequence analysis of the *bamL* locus genes encoded for two proteins, both with extensive β-strand structures, potentially forming transmembrane β-barrel domains. Interestingly, BamL also shared structural homology to prototype translocase FhaC; a TpsB member protein necessary for extracellular localisation of virulence associated protein FHA (Clantin *et al.*, 2007, Delattre *et al.*, 2011).

Homology models provided further structural clues about BamL and LupV. The predicted BamL homology model was a single β-barrel templated from solved structures Omp85/TpsB superfamily (i.e. TamA and FhaC) but excluded any POTRA domains (Figure 4.1.1 B). The lack of POTRA domains does not rule out that BamL could still be involved in an aspect of OM biogenesis, as it has been shown mitochondrial Omp85 homolog, Sam50, has been shown to retain functionality even with removal of its single POTRA domain (Stroud *et al.*, 2011). Structure prediction and sequence analysis of LupV was difficult as its N-terminal region has no recognised sequence motifs and is predicted as an intrinsically disordered ~140amino acid stretch. While no evidence has yet been gathered, this short segment of LupV might interact with BamL.

Distribution of genes and their conservation rates can give contextual clues of function (Blin *et al.*, 2017, Cooper & Brown, 2008). Thus, to better understand how ubiquitous the *bamL* locus was in *K. pneumoniae*, I looked at its distribution, nucleotide conservation and prevalence

rates in a globally diverse genomic dataset of isolates (Figure 4.1.2). Analysis of *K. pneumoniae* sequences showed that the nucleotide conservation of the *bamL* locus proteins was very high (>90%). The lack of accumulated mutations in both *lupV* and *bamL* in the analysed genome sequences argues against the idea that they are pseudogenes. Furthermore, this observed high conservation could also highlight that the predicted domains of BamL or LupV all have functional value and have a similar mode of activity. When looking at the gene prevalence rates of *bamL* and *lupV* in the same global dataset, I found that the two genes are found together in ~80% (561 out of 705 genomes). These data suggest that two genes most likely encode functional proteins and due to their co-prevlance rates further highlights a likely functional association.

K. pneumoniae is a causative agent of nosocomial pneumonia and part of its success can be attributed to its evolution in evading and survival from macrophage killing (Cano *et al.*, 2015, Doorduijn *et al.*, 2016). I therefore tested whether the locus plays a role in protection against uptake by macrophages. Illustrated in Figure 4.2.4-4.4.5, internalisation by macrophage of the mutant strain showed only a mild difference compared to the isogenic wildtype *K. pneumoniae* that are not reflective of an immunologically relevant level of attenuation. Perhaps, under the tested growth conditions there is little to no expression of the *bamL* locus or is it not functioning to protect against macrophage phagocytosis.

Immunoblotting studies did not detect LupV secretion in concentrated supernatants (Figure 4.5.3) arguing against the operon encoded for a classical Type 5 Secretion system (i.e. T5SSb). Subsequent subcellular localisation and urea extraction analyses indicated that recombinant BamL and LupV were both integrally associated in the outer membrane of *K. pneumoniae* B5055. These analyses also suggest that it was unlikely that the C-terminal tags impacted on the correct folding of LupV and BamL into the OM. While these data suggest that LupV and BamK are not involved in classical Type 5 secretion, one could speculate that they assist the secretion of other unidentified substrates. An alternative speculation, inspired by a recent study (Hsueh *et al.*, 2018), is that BamL could be involved in an aspect of metabolite-nutrient transport. The authors found that deletion of P39, a noNterm Omp85 in *Arabidopsis thaliana*, is non-lethal but resulted in noticeable differences in metabolite content in mature plant leaves and speculate its function as an assembly factor for OMPs involved in nutrient transport.

The purification of recombinant BamL for antibody production and crystallisation trials proved promising but requires further optimisation. Crystallisation of membrane proteins are inherently difficult (Carpenter *et al.*, 2008). One major consideration is the choice of detergent for protein solubilisation. In this study the use of DDM was used due to its applicability for

downstream studies and C_8E_4 was chosen as it has been successfully used previously for the crystallisation of several Omp85/TpsB superfamily proteins (Noinaj *et al.*, 2013, Ni *et al.*, 2014b, Fairman *et al.*, 2011). In the aims of structure determination of BamL, commercially available protein crystallisation screens were tested but did not result in any positive hits (Figure 4.6.1). These specific solution and precipitant conditions are empirically derived based on known or published crystallisation conditions of various proteins in the past, to sample as large a range of buffer, pH, additive and precipitant components as possible, using minute amounts of the protein target (Parker & Newstead, 2016). The lack of success in crystallising this Bam could be due to one or more aspects which include sample homogeneity, lack of inherent structural protein stability and starting protein concentration.

This study sets the groundwork for future investigations to fully characterise the newly-identified proteins BamL and LupV. As discussed, genetic analysis investigated the conservation and prevlance rates of the *bamL* locus. Future analyses could statistically address whether specific pathogenic lineages of *K. pneumoniae* have this locus in high frequency and whether the *bamL* locus and other genes show co-prevalence through specific lineages. As an example, in *K. pneumoniae* a hypervirulent phenotype can be credited to its propensity to harbour RmpA and MagA, known transcriptional activators of the capsule biosynthesis. Dysregulation results of *rmpA* and *rmpA2* results in a hypermucoid phenotype associated with many virulent nosocomial isolates (Shon *et al.*, 2013, Lai *et al.*, 2003, Shankar *et al.*, 2018). Future studies could also focus on whether BamL directly assembles LupV using pulse chase assays as developed for the assembly of other OMPs (Stubenrauch *et al.*, 2016). Finally, if we are to assume BamL functions as a conduit for LupV assembly/insertion/translation, the question still remains what functional role in virulence or metabolism the latter has. Future studies looking at the *bamL* locus in the context of nutrient uptake could also be explored, as some OMP β-barrels functions as receptors, such as the TonB-dependent receptor family for iron and vitamin B12 acquisition (Garcia *et al.*, 2011, Shultis *et al.*, 2006). While the precise role of the *bamL* locus in OM biogenesis is still unclear, this locus is an enticing target for further investigation given its high co-prevalence rate in *K. pneumoniae* genomes and its shared homology to established virulence-associated molecules like TamA and FhaC.

5

Results: Identification and functional characterisation of BamK

5.0 Introduction

In Gram-negative bacteria, the correct folding and insertion of β-barrel proteins into the outer membrane is facilitated by the BAM complex (Knowles *et al.*, 2009, Hagan *et al.*, 2011). The BAM complex is a hetero-pentameric machine that catalyses this important function by a not fully understood mechanism but is believed to involve interactions with nascent protein substrates, periplasmic chaperones and the outer membrane itself (Konovalova *et al.*, 2017, Noinaj *et al.*, 2017). The central component of the BAM complex is BamA, that has been shown to be essential for cell viability (Malinverni *et al.*, 2006, Misra *et al.*, 2015).

BamA is the best studied member of the Omp85 superfamily of proteins in terms of structure and function. In a comprehensive phylogenetic Omp85 study (Heinz & Lithgow, 2014), it was reported that at least 8 sub-families exist with a myriad of different domain architectures, with BamA representing the archetypal sub-family. In recent years the structure of BamA proteins from various organisms have been solved by crystallography and cryo-electron microscopy with all showing a highly conserved 16 stranded β-barrel and five periplasmic domains termed POTRAs (Albrecht *et al.*, 2014, Gu *et al.*, 2016, Han *et al.*, 2016, Iadanza *et al.*, 2016).

In *K. pneumoniae* genomes, there is a *bamA* located in a locus conserved in Enterobacteriaceae but also a near identical uncharacterised gene copy, I term *bamK*. This *bamK* is present at distinct loci and is also conserved across all known *K. pnuemoniae* genomes. Bioinformatic analyses showed that *bamK* formed a divergent monophyletic branch distinguishable from *bamA* found primarily in *Klebsiella* spp. but also in other closely related Enterobacteriaceae family members which include: *Raoultella*, *Kosakonia* and *Mangrovicbacter* spp. (Torres *et al.*, 2018). In this chapter, I present results on the initial bioinformatic analyses, phenotypic assays and biochemical approaches undertaken to identify and characterise the function of this novel Omp85 member.

5.1 Genomic context, sequence and prevalence of *bamK*

Previous bioinformatic analysis of *K. pneumoniae* genomes revealed that two genes belonging to the Omp85 protein family are encoded with predicted domain architectures homologous to BamA (Heinz & Lithgow, 2014). Many studies have shown essentiality of BamA, a member of the Omp85 superfamily involved in assembly and insertion of virtually all β-barrel proteins for cell viability (Fairman *et al.*, 2011, Walther *et al.*, 2009). Therefore, I wanted to further investigate why *Klebsiella* spp. would retain a duplicate gene copy paralogous to BamA. In the genome of model strain - *K. pneumoniae* B5055, gene locus tag BN49_4147 is annotated as *yaeT* but is designated as *bamA* based on its matching genomic context and sequence similarity to that of the *E. coli bamA* locus (Figure 5.1.1). The second gene locus tag BN49_4981, annotated as *yaet2* I designate as *bamK* (*Klebsiella*) and is not found in *E. coli* genomes. Unlike the *bamA* gene, neighbouring genes in proximity to *bamK* do not seem to have functions related to outer membrane biogenesis.
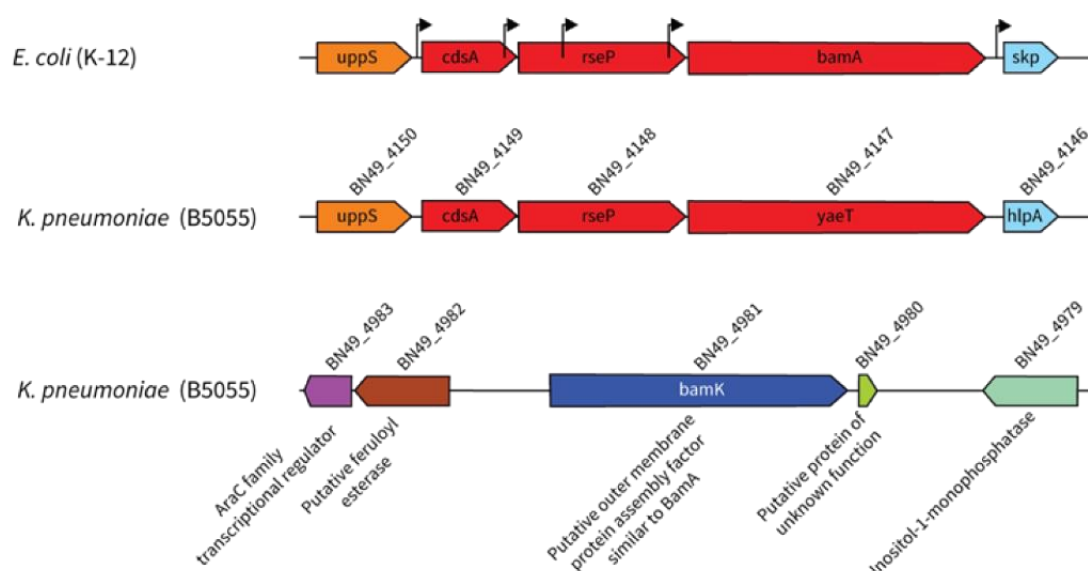
Figure 5.1.1: Genomic context of bamA and *bamK* in *E. coli* and *K. pneumoniae*. The genetic synteny at the *bamA* locus was found to be identical between highly characterised lab strain *E. coli* K-12 and model strain *K. pneumoniae* B5055. The *bamA* gene locus is colour coded in the same fashion to show conserved genes between the two species even where annotated gene names differ (e.g. *yaeT* a previous alias of *bamA* in *E. coli*). On the other hand, the *bamK* locus shown is unique to *Klebsiella* spp. and is not found in a gene neighbourhood with functions known to be involved in outer membrane biogenesis (i.e. *rseP* - regulator of σ^E, involved in envelope stress response; and *skp* a known periplasmic chaperone involved in transit of nascent OMPs to the outer membrane for proper folding and assembly) (Sklar *et al.*, 2007b).

Using the Phyre2 protein prediction server, BamK was modelled to have a high degree of structural homology to available BamA crystal structures, reflecting its moderately high amino acid identity of ~64% with >95% confidence (Figure 5.1.2). A closer inspection of the

generated model of BamK showed five POTRA domains followed by prototypical Omp85 β-barrel paralleling an identical structure to that of BamA.

A.
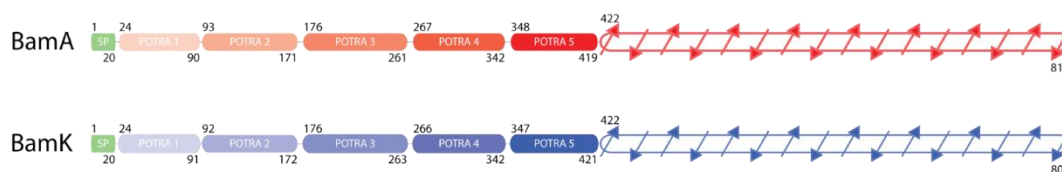
B.
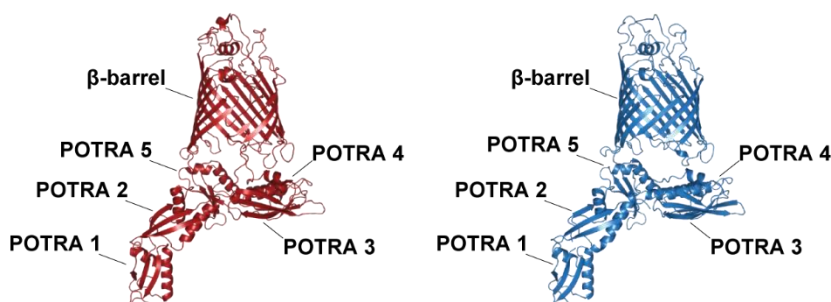
Figure 5.1.2: Predicted domain architecture of BamK. (A) The domain architecture of the two homologous proteins are predicted to be identical. Domains from (N-terminus to C-terminus): (i) outer membrane signal peptide (SP), (ii) 5 POTRA domains and (iii) β-barrel. (B) Crystal structure of BamA (PDB: 5EQK) compared to a homology structure BamK generated by PHYRE2 (Kelley *et al.*, 2015).

To dissect the differences between the Omp85 members at the protein domain and amino acid levels, a multiple sequence alignment of select *K. pneumoniae* genomes were compared (Figure 5.1.3A-B). From the multiple sequence alignment, it was observed that the POTRA domains are typically conserved and most sequence differences between the BamA and BamK sequences were situated in the C-terminal β-barrel domain, especially at positions in the extracellular loops (Figure 5.1.3B). Despite this, the essential VRGF BamA motif found in extracellular loop 6 (L6) is still conserved in BamK sequences. This could suggest a tight barrel association with L6 of BamK is achieved analogously to the L6 function of BamA, as any alternations in this motif compromises the folded state of the Omp85 β-barrel domain obtruding proper function.

To identify how prevalent Omp85 *bamK* was in *K. pneumoniae* in comparison to BamA, a bioinformatic analysis on a dataset of globally diverse *Klebsiella* spp. isolate genomes (Torres *et al.*, 2018) was previously undertaken. The analysis demonstrated that Omp85 genes, *bamA* and *bamK*, are co-conserved in the genome of all four distinct *K. pneumoniae* subspecies (Figure 5.1.4).

```
                                        |Start Potra 1
MGH78578_BamA    MAMKKLLIASLLFSSATVYGAEGFVVKDIHFEGLQRVAVGAALLSMPVRPGDTVTDDDIS  60
NTUH-K2044_BamA  MAMKKLLIASLLFSSATVYGAEGFVVKDIHFEGLQRVAVGAALLSMPVRPGDTVTDDDIS  60
342_BamA         MAMKKLLIASLLFSSATVYGAEGFVVKDIHFEGLQRVAVGAALLSMPVRPGDTVTDDDIS  60
B5055_BamA       MAMKKLLIASLLFSSATVYGAEGFVVKDIHFEGLQRVAVGAALLSMPVRPGDTVTDDDIS  60
E_coli BamA      MAMKKLLIASLLFSSATVYGAEGFVVKDIHFEGLQRVAVGAALLSMPVRTGDTVNDEDIS  60
MGH78578_BamK    MLKKTHIISGLLITPLTLYAATSYQVDDIRFEGLQRVTVGAALLSMPLHAGDAVTPEDVS  60
NTUH-K2044_BamK  MLKKTHIISGLLIAPLTLYAATSYQVDDIRFEGLQRVTVGAALLSMPLHAGDAVTPEDVS  60
342_BamK         MLKKTHIISGLLIAPLTLYAATSYQVDDIRFEGLQRVTVGAALLSMPLHAGDAVTPEDVS  60
B5055_BamK       MLKKTHIISGLLITPLTLYAATSYQVDDIRFEGLQRVTVGAALLSMPLHAGDAVTPEDVS  60
                 *  *. :*:.**::. *:*.* .: *.**:******:*********::.**:*. :*:*

                              End Potra 1||Start Potra 2
MGH78578_BamA    NTIRALFATGNFEDVRVLRDGDTLLVQVKERPTIASITFSGNKSVKDDMLKQNLEASGVR 120
NTUH-K2044_BamA  NTIRALFATGNFEDVRVLRDGDTLLVQVKERPTIASITFSGNKSVKDDMLKQNLEASGVR 120
342_BamA         NTIRALFATGNFEDVRVLRDGDTLLVQVKERPTIASITFSGNKSVKDDMLKQNLEASGVR 120
B5055_BamA       NTIRALFATGNFEDVRVLRDGDTLLVQVKERPTIASITFSGNKSVKDDMLKQNLEASGVR 120
E_coli BamA      NTIRALFATGNFEDVRVLRDGDTLLVQVKERPTIASITFSGNKSVKDDMLKQNLEASGVR 120
MGH78578_BamK    EAVRALYASGNFENVQILRDGKTLVVQVKERPTIASVSFSGNKAVKDDALKENLTASGIS 120
NTUH-K2044_BamK  EAVRALYASGNFENVQILRDGKTLVVQVKERPTIASVSFSGNKAVKDDALKENLTASGIS 120
342_BamK         EAVRALYASGNFENVQILRDEKTLVVQVKERPTIASVSFSGNKAVKDDALKENLTASGIS 120
B5055_BamK       EAVRALYASGNFENVQILRDGKTLVVQVKERPTIASVSFSGNKAVKDDALKENLTASGIS 120
                 ::.:***:*:****:*::*** .**:**********:*****:**** **:** ***:

                                          End Potra 2|  |Start Potra 3
MGH78578_BamA    VGESLDRTTIADIEKGLEDFYYSVGKYSASVKAVVTPLPRNRVDLKLVFQEGVSAKIQQI 180
NTUH-K2044_BamA  VGESLDRTTIADIEKGLEDFYYSVGKYSASVKAVVTPLPRNRVDLKLVFQEGVSAKIQQI 180
342_BamA         VGESLDRTTIADIEKGLEDFYYSVGKYSASVKAVVTPLPRNRVDLKLVFQEGVSAKIQQI 180
B5055_BamA       VGESLDRTTIADIEKGLEDFYYSVGKYSASVKAVVTPLPRNRVDLKLVFQEGVSAKIQQI 180
E_coli BamA      VGESLDRTTIADIEKGLEDFYYSVGKYSASVKAVVTPLPRNRVDLKLVFQEGVSAEIQQI 180
MGH78578_BamK    AGSALDRNSLSEIEKGLQDFYYSAGKYSAQVHAVVTPLPRNRVDLTFVFQEGISAKIAQI 180
NTUH-K2044_BamK  AGSALDRNSLSEIEKGLQDFYYSAGKYSAQVHAVVTPLPRNRVDLTFVFQEGISAKIAQI 180
342_BamK         AGSALDRNSLSEIEKGLQDFYYSAGKYSAQVHAVVTPLPRNRVDLTFVFQEGISAKIAQI 180
B5055_BamK       AGSALDRNSLSEIEKGLQDFYYSAGKYSAQVHAVVTPLPRNRVDLTFVFQEGISAKIAQI 180
                 .*.:***.::::*****:*****.*****.*:************.:*****:**:* **

MGH78578_BamA    NIVGNHAFSTDELISHFQLRDEVPWWNVVGDRKYQKQKLAGDLETLRSYYLDRGYARFNI 240
NTUH-K2044_BamA  NIVGNHAFSTDELISHFQLRDEVPWWNVVGDRKYQKQKLAGDLETLRSYYLDRGYARFNI 240
342_BamA         NIVGNHAFSTDELISHFQLRDEVPWWNVVGDRKYQKQKLAGDLETLRSYYLDRGYARFNI 240
B5055_BamA       NIVGNHAFSTDELISHFQLRDEVPWWNVVGDRKYQKQKLAGDLETLRSYYLDRGYARFNI 240
E_coli BamA      NIVGNHAFTTDELISHFQLRDEVPWWNVVGDRKYQKQKLAGDLETLRSYYLDRGYARFNI 240
MGH78578_BamK    NIIGNQAFREETLLDQLQLRDNVPWWNVVADKKYQKQKLEADLETLRSFYLDRGYARFAI 240
NTUH-K2044_BamK  NIIGNQAFREETLLDQLQLRDNVPWWNVVADKKYQKQKLEADLETLRSFYLDRGYARFAI 240
342_BamK         NIIGNQAFREETLLDQLQLRDNVPWWNVVADKKYQKQKLEADLETLRSFYLDRGYARFAI 240
B5055_BamK       NIIGNQAFREETLLDQLQLRDNVPWWNVVADKKYQKQKLEADLETLRSFYLDRGYARFAI 240
                 **:**:**  :  *:.::****:*******.*:******  .*******:********* *

                              End Potra 3|  |Start Potra 4
MGH78578_BamA    DSTQVSLTPDKKGIYITVNITEGDQYKFSGVQVTGNLAGHSAEIEALTKVEPGELYNGAK 300
NTUH-K2044_BamA  DSTQVSLTPDKKGIYITVNITEGDQYKFSGVQVTGNLAGHSAEIEALTKVEPGELYNGAK 300
342_BamA         DSTQVSLTPDKKGIYITVNITEGDQYKISGVQVTGDLAGHSAEIEALTKVEPGELYNGAK 300
B5055_BamA       DSTQVSLTPDKKGIYITVNITEGDQYKFSGVQVTGNLAGHSAEIEALTKVEPGELYNGAK 300
E_coli BamA      DSTQVSLTPDKKGIYVTVNITEGDQYKLSGVEVSGNLAGHSAEIEQLTKIEPGELYNGTK 300
MGH78578_BamK    ESTQVSMTPDKKSLYITIALNEGERYRVDRTQVTGDLAQHGPEIEALAQPLAGAWYSGAQ 300
NTUH-K2044_BamK  ESTQVSMTPDKKSLYITIALNEGERYRVDRTQVTGDLAQHGPEIEALAQPLAGAWYSGAQ 300
342_BamK         ESTLVSMTPDKKSLYITIALNEGERYRVDRTQVTGDLAQHGPEIEALAQPLAGAWYSGAQ 300
B5055_BamK       ESTQVSMTPDKKSLYITIALNEGERYRVDRTQVTGDLAQHGPEIEALAQPLAGAWYSGAQ 300
                 :** **:*****.:*:*: :.**::*:.. .:*:*:** *..*** *:: .* *.*::

                                          End Potra 4|  |Start Potra 5
MGH78578_BamA    VTRMENDIKKLLGRYGYAYPRVQSQPEINDSDKTVKLHVNVDAGNRYYVRKIRFEGNDTS 360
NTUH-K2044_BamA  VTKMENDIKKLLGRYGYAYPRVQSQPEINDSDKTVKLHVNVDAGNRYYVRKIRFEGNDTS 360
342_BamA         VTKMENDIKKLLGRYGYAYPRVQSQPEINDSDKTVKLHVNVDAGNRYYVRKIRFEGNDTS 360
B5055_BamA       VTRMENDIKKLLGRYGYAYPRVQSQPEINDSDKTVKLHVNVDAGNRYYVRKIRFEGNDTS 360
E_coli BamA      VTKMEDDIKKLLGRYGYAYPRVQSMPEINDADKTVKLRVNVDAGNRFYVRKIRFEGNDTS 360
MGH78578_BamK    VTTVENEIKKHFGKYGYAWPQVTSTPEIDDAHHRVVLHIQVNAGRRYSVRQIRFSGNDTS 360
NTUH-K2044_BamK  VTTVENEIKKHFGKYGYAWPQVTTTPEIDDAHHRVVLHIQVNAGRRYSVRQIRFSGNDTS 360
342_BamK         VTTVENEIKKRFGKYGYAWPQVTTTPEIDDAHHRVALHIQVNAGRRYSVRQIRFSGNDTS 360
B5055_BamK       VTTVENEIKKHFGKYGYAWPQVTSTPEIDDAHHRVVLHIQVNAGRRYSVRQIRFSGNDTS 360
                 ** :*::*** :*:*****:*:* : ***:*:.: * *:::*:**.*: **:***.*****

                                              End Potra 5|
MGH78578_BamA    KDAVLRREMRQMEGAWLGSDLVDQGKDRLNRLGFFETVDTDTQRVPGSPDQVDVVYKVKE 420
NTUH-K2044_BamA  KDAVLRREMRQMEGAWLGSDLVDQGKDRLNRLGFFETVDTDTQRVPGSPDQVDVVYKVKE 420
342_BamA         KDAVLRREMRQMEGAWLGSDLVDQGKDRLNRLGFFETVDTDTQRVPGSPDQVDVVYKVKE 420
B5055_BamA       KDAVLRREMRQMEGAWLGSDLVDQGKDRLNRLGFFETVDTDTQRVPGSPDQVDVVYKVKE 420
E_coli BamA      KDAVLRREMRQMEGAWLGSDLVDQGKERLNRLGFFETVDTDTQRVPGSPDQVDVVYKVKE 420
MGH78578_BamK    RDAVLRREMRQMEGAWLNNEKVDQGKVRLDRTGFFENVEQQIVPVNGTADQVDVVYKVKE 420
NTUH-K2044_BamK  RDAVLRREMRQMEGAWLNNEKVDQGKVRLDRTGFFENVEQQIVPVNGTADQVDVVYKVKE 420
342_BamK         RDAVLRREMRQMEGAWLNNEKVEQGKVRLDRTGFFENVSGTADQVDVVYKVKE 420
B5055_BamK       RDAVLRREMRQMEGAWLNNEKVDQGKVRLDRTGFFENVEQQIVPVNGTADQVDVVYKVKE 420
                 :****************..: *:*** **:* ****.*: :   * *:.***********
```

Figure 5.1.3A: Conserved sequence features in POTRA domains of BamA and BamK. Multiple sequence alignment of the POTRA domains of a select set of BamA and BamK amino acid sequences. Residues are coloured according to side-chain properties to guide assessment of conservative and non-conservative substitutions, (*) indicates sequence identity (:) indicates highly conserved (.) indicates some degree of conservation. The boundaries for the five POTRA domains are indicated, based on the crystal structure of BamA from *E. coli* (PDB: 5EKQ). The positions of the transmembrane β-strands in the β-barrel domain of BamA from *E. coli* are indicated by arrows and are numbered (β1, β2 etc). Red dashes indicate the extracellular loop regions. Blue dots indicate the short periplasmic turns. The multiple sequence alignment continues in Figure 5.1.3B.

116

```
                    ↓β1      β1↓--↓β2      β2↓●●●●●●●●↓β3   β3↓--↓β4    β4↓●●●●
MGH78578_BamA     RNTGSFNFGIGYGTESGVSFQAGVQQDNWLGTGYAVGINGTKNDYQTYTELSVTNPYFTV 480
NTUH-K2044_BamA   RNTGSFNFGIGYGTESGVSFQAGVQQDNWLGTGYAVGINGTKNDYQTYTELSVTNPYFTV 480
342_BamA          RNTGSFNFGIGYGTESGVSFQAGVQQDNWLGTGYAVGINGTKNDYQTYTELSVTNPYFTV 480
B5055_BamA        RNTGSFNFGIGYGTESGVSFQAGVQQDNWLGTGYAVGINGTKNDYQTYTELSVTNPYFTV 480
E_coli_BamA       RNTGSFNFGIGYGTESGVSFQAGVQQDNWLGTGYAVGINGTKNDYQTYAELSVTNPYFTV 480
MGH78578_BamK     RNTGSFNVGLGFGTDSGVSYQLGVTQDNWLGTGNSVSFNGTRNSYQSYLELGATNPWFTV 480
NTUH-K2044_BamK   RNTGSFNVGLGFGTDSGVSYQLGVTQDNWLGTGNSVSFNGTRNSYQSYLELGATNPWFTV 480
342_BamK          RNTGSFNVGLGFGTDSGVSYQLGVTQDNWLGTGNSVSFNGTRNSYQSYLELGATNPWFTV 480
B5055_BamK        RNTGSFNVGLGFGTDSGVSYQLGVTQDNWLGTGNSVSFNGTRNSYQSYLELGATNPWFTV 480
                  *******.*:*:**:****:* **.******** :*.:***:*.**:* **.***:***

                  ●●●↓β5      β5↓---------β6●●●●●●●●●●●●●●↓β7     β7↓-----
MGH78578_BamA     DGVSLGGRVFYNDFDANDADLSDYTNKSYGTDITLGFPVNEYNTLRAGVGYVHNSLSNMQ 540
NTUH-K2044_BamA   DGVSLGGRVFYNDFDANDADLSDYTNKSYGTDITLGFPVNEYNTLRAGVGYVHNSLSNMQ 540
342_BamA          DGVSLGGRIFYNDFDANDADLSDYTNKSYGTDVTLGFPVNEYNTLRAGVGYVHNSLSNMQ 540
B5055_BamA        DGVSLGGRVFYNDFDANDADLSDYTNKSYGTDITLGFPVNEYNTLRAGVGYVHNSLSNMQ 540
E_coli_BamA       DGVSLGGRLFYNDFQADDADLSDYTNKSYGTDVTLGFPINEYNSLRAGLGYVHNSLSNMQ 540
MGH78578_BamK     DGISLGGKIFYNSYDASDADAGSYNQQSYGLGSTLGFPISENNSLNLGLDYVHNRLTNMD 540
NTUH-K2044_BamK   DGISLGGKIFYNSYDASDADAGSYNQQSYGLGSTLGFPISENNSLNLGLDYVHNRLTNMD 540
342_BamK          DGISLGGKIFYNSYDASDADAGSYNQQSYGLGSTLGFPISENNSLNLGLDYVHNRLTNMD 540
B5055_BamK        DGISLGGKIFYNSYDASDADAGSYNQQSYGLGSTLGFPISENNSLNLGLDYVHNRLTNMD 540
                  **:****::***.::*.*** ..*.::*** . *****:.* *:*. *:.**** *:**:

                  ----------------------↓β8          β8↓●●●●●●●●●●●●●●↓β9
MGH78578_BamA     PQVAMWRYLNSMGQYPDNTNDRNS----FSANDFTFNYGWTYNKLDRGFFPTEGSRVNLN 596
NTUH-K2044_BamA   PQVAMWRYLNSMGQYPDNTNDRNS----FSANDFTFNYGWTYNKLDRGFFPTEGSRVNLN 596
342_BamA          PQVAMWRYLNSMGQYPDNTNDRNS----FSANDFTFNYGWTYNKLDRGFFPTEGSRVNLN 596
B5055_BamA        PQVAMWRYLNSMGQYPDNTNDRNS----FSANDFTFNYGWTYNKLDRGFFPTEGSRVNLN 596
E_coli_BamA       PQVAMWRYLYSMGEHPSTSDQDNS----FKTDDFTFNYGWTYNKLDRGYFPTDGSRVNLT 596
MGH78578_BamK     PELTTWRYLSSRGIEPSVVTKDGDSGAKYSANDYFVSLGWGYNDLDRGFFPRAGNKSSLS 600
NTUH-K2044_BamK   PELTTWRYLSSRGIEPSVVTKDGDSGAKYSANDYFVSLGWGYNDLDRGFFPRAGNKSSLS 600
342_BamK          PELTTWRYLSSRGIEPSVVTKDGDSGAKYSANDYFVSLGWGYNDLDRGFFPRAGNKSSLS 600
B5055_BamK        PELTTWRYLSSRGIEPSVVTKDGDSGAKYSANDYFVSLGWGYNDLDRGFFPRAGNKSSLS 600
                  *:::: **** * * *.  . .. :.::*:.. ** **.****:** *.: .*.

                   β9↓--------↓β10   β10↓●●●●●●●●↓β11   β11↓----------------
MGH78578_BamA     GKVTIPGSDNEYYKATLDTATYVPIDNDHQWVVLGRTRFGYGDGIGGKEMPFYENFYAGG 656
NTUH-K2044_BamA   GKVTIPGSDNEYYKATLDTATYVPIDNDHQWVVLGRTRFGYGDGIGGKEMPFYENFYAGG 656
342_BamA          GKVTIPGSDNEYYKATLDTATYVPIDNDHQWVVLGRTRFGYGDGIGGKEMPFYENFYAGG 656
B5055_BamA        GKVTIPGSDNEYYKATLDTATYVPIDNDHQWVVLGRTRFGYGDGIGGKEMPFYENFYAGG 656
E_coli_BamA       GKVTIPGSDNEYYKVTLDTATYVPIDDDHKWVVLGRTRWGYGDGLGGKEMPFYENFYAGG 656
MGH78578_BamK     GKVTLPGSDNSYYKLSFDTAQYLPLSENKRWVWMERLRAGYAGGLDGKSVPFYDNFYAGG 660
NTUH-K2044_BamK   GKVTLPGSDNSYYKLSFDTAQYLPLSENKRWVWMERLRAGYAGGLDGKSVPFYDNFYAGG 660
342_BamK          GKVTLPGSDNSYYKLSFDTAQYLPLSENKRWVWMERLRAGYAGGLDGKSVPFYDNFYAGG 660
B5055_BamK        GKVTLPGSDNSYYKLSFDTAQYLPLSENKRWVWMERLRAGYAGGLDGKSVPFYDNFYAGG 660
                  ****:*****.*** ::*** *:*:.::::*  :  *  * **..*:.**.:***:*****

                  -------------------------------------------------↓β12
MGH78578_BamA     SSTVRGFQSNTIGPKAVYFPASSRHDDDDSYDNECKSTESAP--CKSDDAVGGNAMAVAS 714
NTUH-K2044_BamA   SSTVRGFQSNTIGPKAVYFPASSRHDDDDSYDNECKSTESAP--CKSDDAVGGNAMAVAS 714
342_BamA          SSTVRGFQSNTIGPKAVYFPSSSRHDDGDSGYTNDCKSTESAP--CKSDDAVGGNAMAVAS 714
B5055_BamA        SSTVRGFQSNTIGPKAVYFPASSRHDDDDSYDNECKSTESAP--CKSDDAVGGNAMAVAS 714
E_coli_BamA       SSTVRGFQSNTIGPKAVYFPHQASN-YDPDYDYECATQDGAKDLCKSDDAVGGNAMAVAS 715
MGH78578_BamK     SSSVRGFSSNTIGPKAAYYRCNGSESSY---SACPLDASS-------DAVGGNAMAVLN 709
NTUH-K2044_BamK   SSSVRGFSSNTIGPKAAYYRCNGSESSY---SACPLDASS-------DAVGGNAMAVLN 709
342_BamK          SSSVRGFSSNTIGPKAAYYRCNGSESSY---SACPLDASS-------DAVGGNAMAVLN 709
B5055_BamK        SSSVRGFSSNTIGPKAAYYRCNGSESSY---SACPLDASS-------DAVGGNAMAVLN 709
                  **:**** .********.*:  .. .        :. :         ********** .

                  β12↓●●●●●●●●●●●●●↓β13     β13↓-----------------------↓β14
MGH78578_BamA     LELITPTPFISDKYANSVRTSVFWDMGTVWDTHWDSSA---YAGYPDYSDPSNIRMSAGI 771
NTUH-K2044_BamA   LELITPTPFISDKYANSVRTSVFWDMGTVWDTHWDSSA---YAGYPDYSDPSNIRMSAGI 771
342_BamA          LELITPTPFISDKYANSVRTSVFWDMGTVWDTHWDSNA---YGGYPDYSDPSNIRMSAGI 771
B5055_BamA        LELITPTPFISDKYANSVRTSVFWDMGTVWDTHWDSSA---YAGYPDYSDPSNIRMSAGI 771
E_coli_BamA       LEFITPTPFISDKYANSVRTSFFWDMGTVWDTNWDSSQ---YSGYPDYSDPSNIRMSAGI 772
MGH78578_BamK     SEFIIPTPFVNDKYADSLRTSLFVDAGTVWSTSWHNTAQTLAAGIPDYGDPSHIRLSAGI 769
NTUH-K2044_BamK   SEFIIPTPFVNDKYADSLRTSLFVDAGTVWSTSWHNTAQTLAAGIPDYGDPSHIRLSAGI 769
342_BamK          SEFIIPTPFVNDKYADSLRTSLFVDAGTVWSTSWQNTAQTLAAGIPDYGDPGHIRLSAGI 769
B5055_BamK        SEFIIPTPFVNDKYADSLRTSLFVDAGTVWSTSWHNTAQTLAAGIPDYGDPSHIRLSAGI 769
                  *:* ****:.****:*:***.*  ****.*.*...    .* ***.**.:**:****

                  β14↓●↓β15   β15↓-------------↓β16   ↓
MGH78578_BamA     AVQWMSPLGPLVFSYAQPFKKYDGDKAEQFQFNIGKTW 809
NTUH-K2044_BamA   AVQWMSPLGPLVFSYAQPFKKYDGDKAEQFQFNIGKTW 809
342_BamA          AVQWMSPLGPLVFSYAQPFKKYDGDKAEQFQFNIGKTW 809
B5055_BamA        AVQWMSPLGPLVFSYAQPFKKYDGDKAEQFQFNIGKTW 809
E_coli_BamA       ALQWMSPLGPLVFSYAQPFKKYDGDKAEQFQFNIGKTW 810
MGH78578_BamK     AVQWMSPLGPLVFSWAEPFKKYDGDKAEQFQFNIGKTW 807
NTUH-K2044_BamK   AVQWMSPLGPLVFSWAEPFKKYDGDKAEQFQFNIGKTW 807
342_BamK          AVQWMSPLGPLVFSWAEPFKKYDGDKAEQFQFNIGKTW 807
B5055_BamK        AVQWMSPLGPLVFSWAEPFKKYDGDKAEQFQFNIGKTW 807
                  *:************:*:*********************
```

Figure 5.1.3B: Conserved sequence features in the β-barrel domains of BamA and BamK. Multiple sequence alignment of the β-barrel domain of a select set of BamA and BamK amino acid sequences. This multiple sequence alignment continues on from Figure 5.1.3A. Highlighted in yellow is the conserved Omp85 VRGF motif which has been shown to make contacts to conserved residues in the β-barrel wall for stability (Noinaj *et al.*, 2017, Thoma *et al.*, 2018). Residues are coloured according to side-chain properties to guide assessment of conservative and non-conservative substitutions, (*) indicates sequence identity (:) indicates highly conserved (.) indicates some degree of conservation. The boundaries for the five POTRA domains are indicated, based on the crystal structure of BamA from *E. coli* (PDB: 5EKQ). The positions of the transmembrane β-strands in the β-barrel domain of BamA from *E. coli* are indicated by arrows and are numbered (β1, β2 etc). Red dashes indicate the extracellular loop regions. Blue dots indicate the short periplasmic turns.
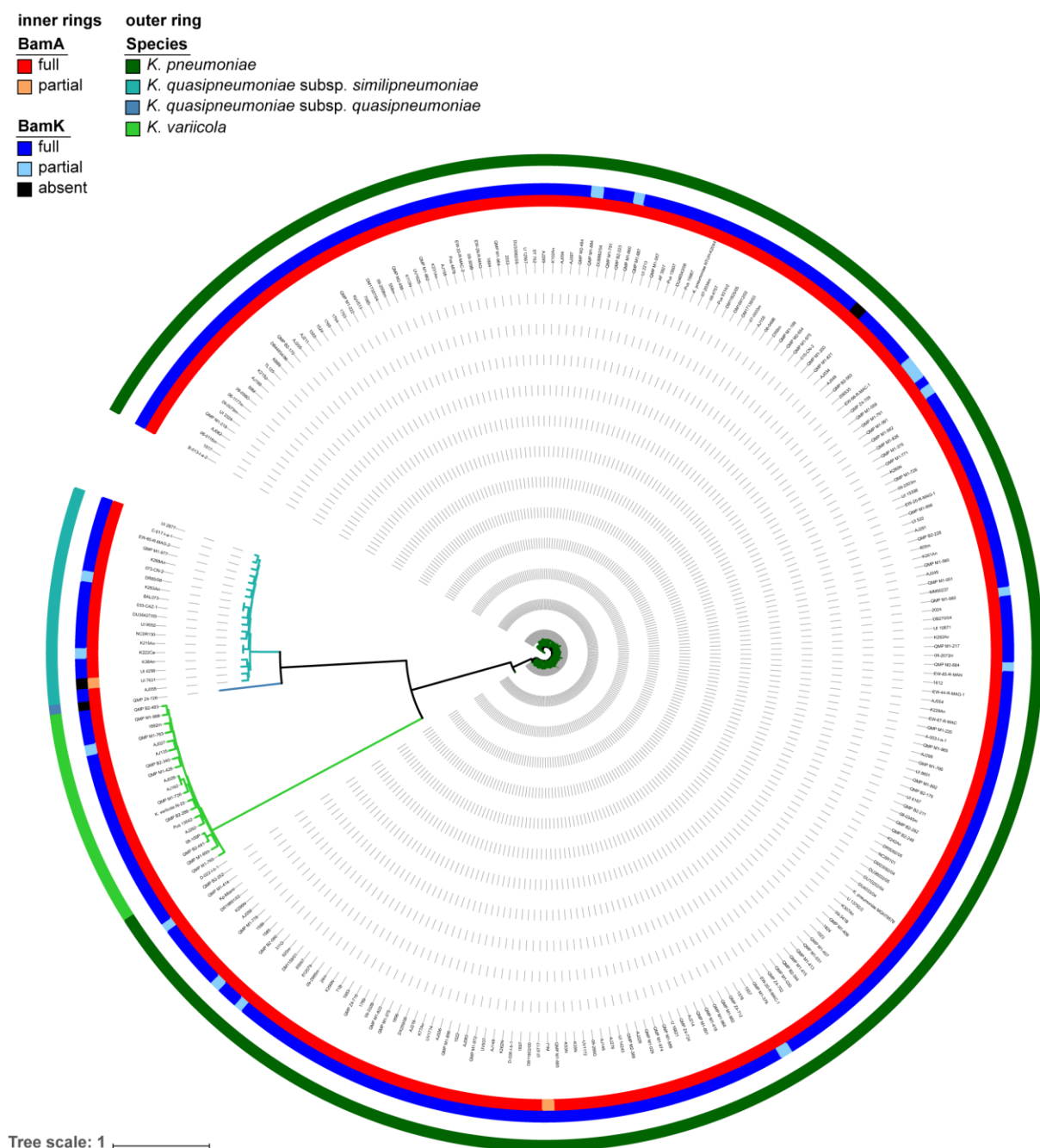
Figure 5.1.4: BamK is encoded in the core genome of *K. pneumoniae*. Sequence distribution of *bamA* and *bamK* genes from a globally diverse genome dataset of *K. pneumoniae* isolates (Holt *et al.*, 2015b) represented as four distinct *K. pneumoniae* species (outer rings). The *bamA* gene is predominantly conserved among all the isolates (red inner ring) and *bamK* shares a co-presence in distribution in virtually all corresponding genomes (blue inner ring). Each genome (branch) is represented by a unique genome identifier text tag. It should be noted that the partial presence or absence of either of the genes shown in the circular phylogram were found to be in genomes with incomplete assemblies and/or poor sequence coverage in these specific regions, and hence may not actually reflect actual gene truncation or full absence. Figure with caption taken directly from reference (Torres *et al.*, 2018). Analysis performed by Dr. Eva Heinz (The Wellcome Sanger Institute, United Kingdom).

5.2 Phenotypic analysis of *bamK* and its role in a host infection model

The function of *bamA* is known to be essential in *E. coli* (Misra *et al.*, 2015, Pfitzner *et al.*, 2016). Hence, I hypothesised that deletion of the highly similar *bamK* may impact processes fundamental for bacterial cell physiology. To address this idea, a *bamK* deletion strain was constructed using the "gene gorging" method for comparison to its isogenic parent strain - *K. pneumoniae* B5055 (constructed by Dr. Abigail Clements, Imperial College London). The resultant deletion strain had no major defects in growth kinetics when compared to the wildtype, when grown in rich (LB), minimal (M9) growth media or in DMEM which is often used as a proxy for "in host" growth conditions (Figure 5.2.1).
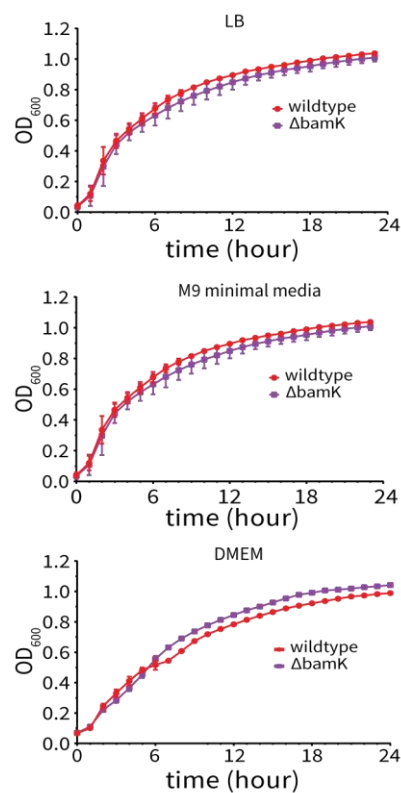
Figure 5.2.1: Deletion of *bamK* in K. *pneumoniae* B5055 and growth kinetics in different media. The planktonic growth of B5055 wildtype and B5055 Δ*bamK* was observed over 24 hours with reading at hourly intervals in rich media (n=3, error bars represent SD) (LB), minimal media (M9) and "host-like" media (DMEM).

A direct assessment of the virulence phenotype of B5055 in a mouse model of infection showed that colonisation of *K. pneumoniae* B5055, with or without BamK, were equivalent in lung and liver samples of C57BL6 mice (Figure 5.2.2). For complementation, a plasmid harbouring *bamK* under its native promoter was transformed into wildtype *K. pneumoniae* B5055 (Δ*bamK* C').
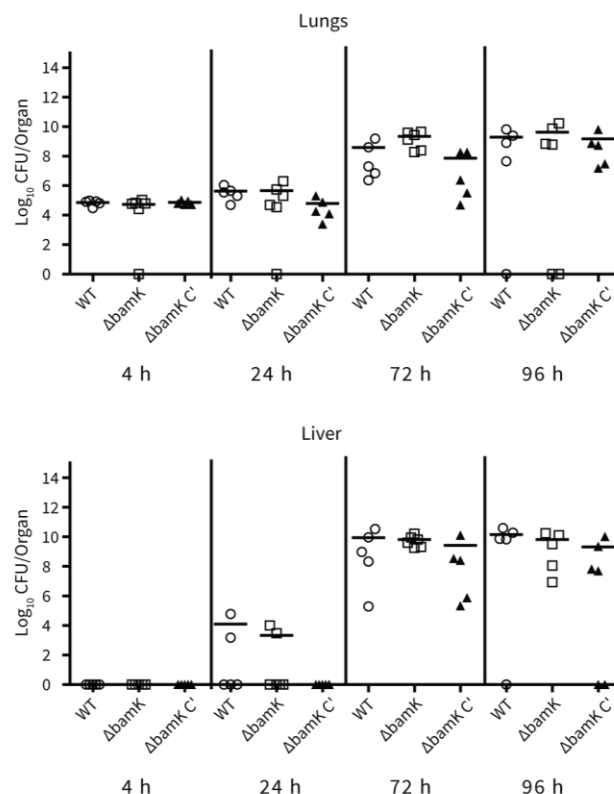
Figure 5.2.2: Colonisation rates of a *bamK* mutant in a mouse infection model. Using an intranasal murine infection model to measure virulence, mice were infected with 8-9 × 10³ CFU of *K. pneumoniae* B5055 wildtype, or isogenic Δ*bamK* mutant strain, or a complemented mutant (B5055 Δ*bamK* harbouring a plasmid with *bamK* under its native promoter). At time points 4, 24, 72 and 96 hours post-infection, bacteria in the lungs and liver were enumerated. the colonization peaked in all mice by 72 hours. The geometric mean is shown for each treatment group with include at least 5 mice each. (performed by Dr. Hanwei Cao, The University of Melbourne, Australia)

As Omp85 proteins are involved in bacterial outer membrane biogenesis and integrity, I sought to assess the integrity of the outer membrane of the Δ*bamK* mutant. Various methods can assess the integrity and permeability barrier function of the Gram-negative outer membrane (Nikaido, 1989, Delcour, 2009, Ieva, 2017). Usually the outer membrane of Gram-negatives quite effectively excludes several antibiotics from entry into the cell; however, previous studies have reported the loss of *bamB* and *bamE* in *E. coli* leads to an increase in the susceptibility against various antibiotics as a result of increased permeability of the outer membrane (Sklar *et al.*, 2007b). To assess if BamK also played a role in outer membrane integrity, I investigated the susceptibility of Δ*bamK* deletion strains and the parent strain to the antibiotics vancomycin, rifampicin, ampicillin, cefipeme, imipenem and meropenem using an agar disk diffusion assay (Table 5.2.3). The diameter of the zone of inhibition will determine the effectiveness of an antibiotic at its given concentration, therefore the larger the diameter the greater sensitivity of the bacterium to its antibiotic. However, if outer membrane integrity was compromised one would expect the diameter of the zone of inhibition to become larger and thus more susceptible. The outer membrane integrity of the Δ*bamK* deletion strains did not reflect

a compromised outer membrane integrity when compared to its parent strain as no marked increase in the zone of inhibitions was observed for any of the tested antibiotics.

Table 5.2.3: Antibiotic sensitivity of B5055 parent strain compared to the ΔbamK strain using the (Kirby-Bauer) disk diffusion method. The strains were tested for susceptbility using commercially available antibiotic impregnated filter discs (BD BBL™ Sensi-disc™, U.S.), (n=9 with± indicating SEM). The non-measurable zone of inhibition of vancomycin, rifampin and ampicillin was expected at the available concentration(s) from previously performed MIC panels on *K. pneumoniae* B5055. The differences in the zone of inibition for the carbapenems (cefipeme, iminipenem and meropenem) were neglibile between the wildtype and ΔbamK strain.

	B5055 wildtype zone of inhibition (mm)	B5055 ΔbamK zone of inhibition (mm)
Vancomycin (30 μg)	Not measurable*	Not measurable*
Rifampin (5 μg)	Not measurable*	Not measurable*
Ampicillin (10 μg)	Not measurable*	Not measurable*
Cefipeme (30 μg)	32.7 ± 1.3	30.1 ± 0.5
Imipenem (10 μg)	30.2 ± 0.4	30.2 ± 0.5
Meropenem (10 μg)	28.6 ± 0.7	29.6 ± 0.4

* indicates zone of inhibition was absent (presumably due to resistance) and therefore not measurable

5.3 Functionally replacing BamA with BamK in the BAM complex in *K. pneumoniae* B5055

In Chapter 3.0, the observed native expression levels of *bamK* was very low and therefore difficult to functionally characterise. Because of the high homology between BamA and BamK (Figure 5.1.2), I therefore wanted to ask if BamK could functionally replace BamA. To do this, a functional gene replacement strain was constructed using the previously described "gene-gorging" method (Herring *et al.*, 2003). The coding region of essential gene *bamA* was precisely replaced with the *bamK* coding region resulting in the *K. pneumoniae* B5055 strain with genotype Δ*bamA*::*bamK* (Figure 5.3.1). The functional gene replacement and following excision of the kanamycin cassette was confirmed by PCR analysis and sequencing using primers flanking the targeted regions (Figure 5.3.2). The resultant strain thus had two copies of *bamK*, one under the constitutively expressed *bamA* promoter and the other under its native *bamK* promoter with cryptic regulation.
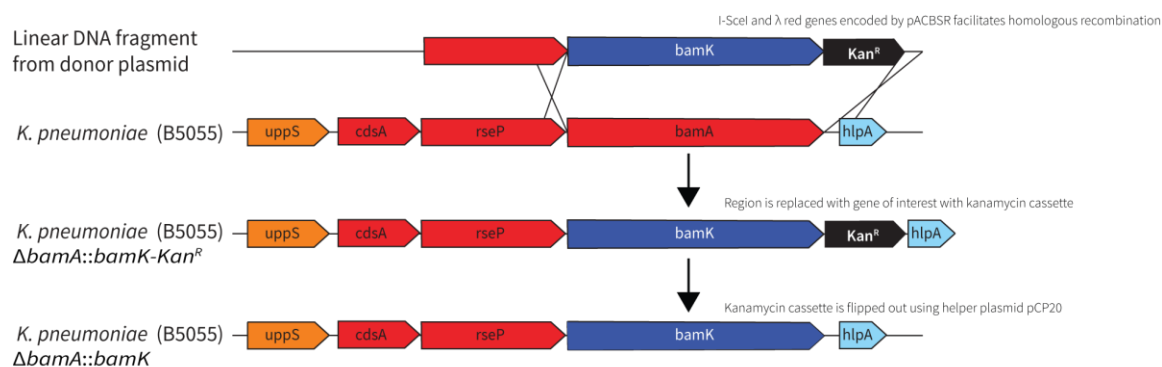
Figure 5.3.1: Strategy for genomic gene replacement of the *bamA* with *bamK* in *K. pneumoniae* B5055. Schematic showing the replacement of the *bamA* coding sequence with the *bamK* coding sequence to produce gene replacement strain *K. pneumoniae* B5055 Δ*bamA*::*bamK*. The resulting *bamK* gene replacement strain was under the *bamA* promoter.
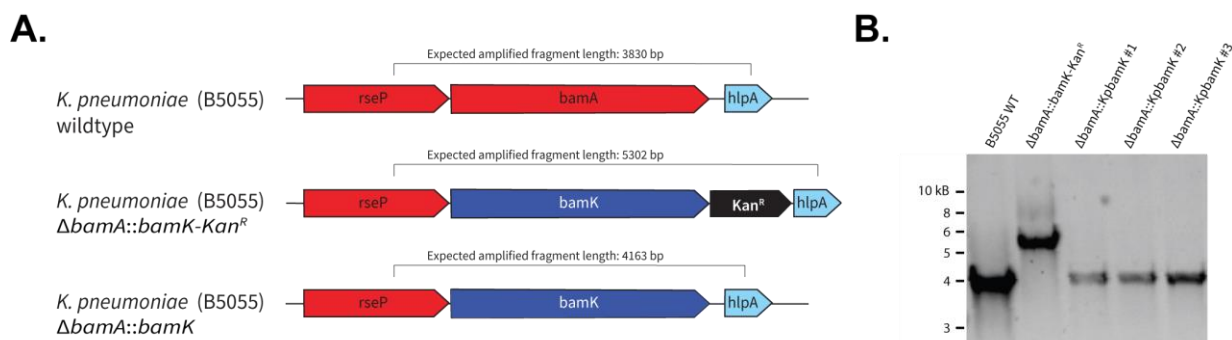
Figure 5.3.2: PCR confirmation of *bamA* replacement with *bamK* in *K. pneumoniae* B5055. (A) Cartoon showing the expected fragment lengths of *K. pneumoniae* B5055 at different stages during the gene knock-in process (refer to Figure 5.3.1). (B) Assaying the intermediate gene knock-in strains by PCR amplification using primers, VVT063 and VVT064, which lie outside the region of genomic integration. PCR products were analysed by 1 % TAE agarose gel and sequenced to confirm the gene replacement.

5.4 Phenotypically characterising the BamA to BamK replacement strain in *K. pneumoniae* B5055

The depletion of the essential *bamA* gene results in abnormal membrane protein profiles due to defective assembly and insertion of integral β-barrel proteins (Misra *et al.*, 2015, Hoang *et al.*, 2011). In Section 5.3, I constructed a gene replacement strain and thus wanted to test its viability and if BamK had any differences in total membrane protein profiles as the resultant strain presumably lacked BamA. Thus, to test this I performed comparative analysis looking for differences in the total membrane protein profiles and growth kinetics between the gene replacement strain (Δ*bamA*::*bamK*) and its isogenic parent strain. Total membranes were prepared from *K. pneumoniae* B5055 Δ*bamA*::*bamK* and wildtype strains and profiles were compared by analysing them through SDS-PAGE and Coomassie blue staining. When comparing the two isogenic strains, no major differences were observed between the two isogenic strains (Figure 5.4.1 A). Furthermore, I also found that the engineered *K. pneumoniae* B5055 Δ*bamA*::*bamK* had growth kinetics similar to the parental B5055 strain when grown in nutrient rich LB-media (Figure 5.4.1 B).
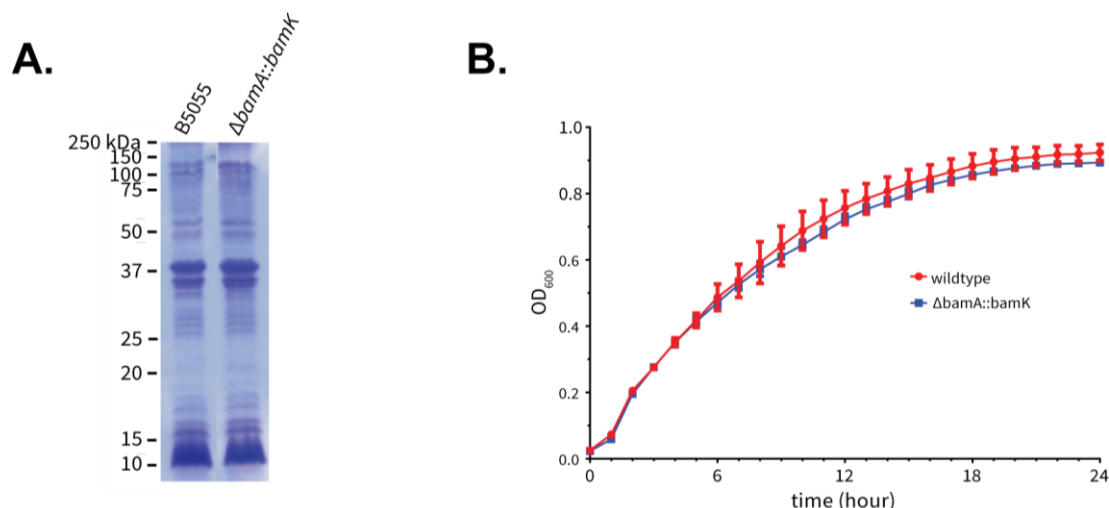
Figure 5.4.1: Comparison between B5055 wildtype and B5055 Δ*bamA*::*bamK*: total membrane protein profiles and growth kinetics. (A) Outer membranes isolated from B5055 wildtype and B5055 Δ*bamA*::*bamK* were analysed by SDS-PAGE (15%) and Coomassie stained (performed by Dr. Rhys Dunstan, Monash University), no major differences in major porins or other OMP levels were seen. (B) Growth curves comparing the parent B5055 strain to the engineered B5055 Δ*bamA*::*bamK* gene replacement isogenic strain. (n=3, error bars represent SD)

Given the wild-type nature of the genomic replacement strains depicted in Figure 5.4.1, BamK was presumably moonlighting as the core of the BAM complex in place of BamA. To directly establish this point, blue native polyacrylamide gel electrophoresis (BN-PAGE) and immunoblotting were employed to assess whether BamK is acting as substitute for the core of the BAM complex. Immunoblot analysis showed that in the replacement strain (B5055 Δ*bamA*::*bamK*), BamK comigrated with all other BAM complex subunits (BamB through BamE), confirming that BamK is functionally replacing BamA in this strain. Immunoblotting for BamA revealed presence in only the wildtype strain comigrating with BAM complex subunits with greater mobility when compared to the replacement strain (Figure 5.4.2).
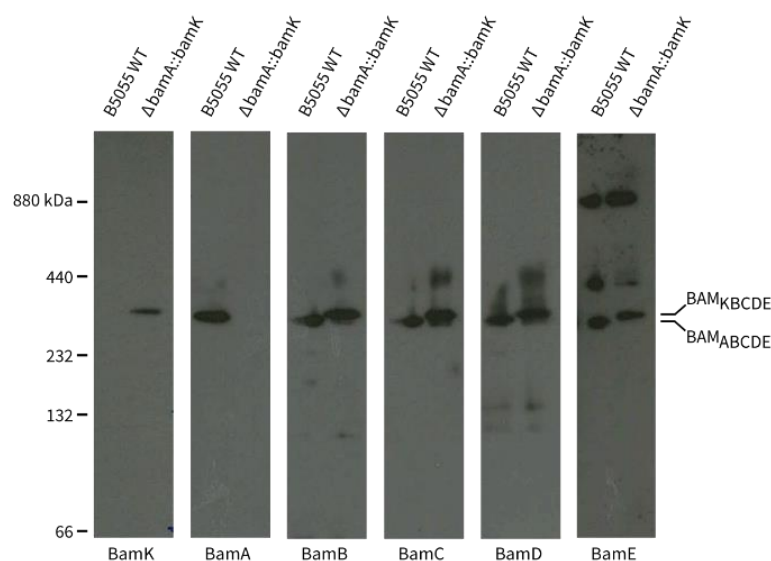
Figure 5.4.2: BN-PAGE of total membranes isolated from B5055 wildtype and B5055 Δ*bamA*::*bamK*. (C) Outer membranes were isolated from B5055 wildtype and B5055 Δ*bamA*::*bamK* and then analysed by BN-PAGE (5-16%) then immunoblotted for BamK and BAM complex proteins (BamA through BamE).

5.5 Functionally replacing BamA with BamK in the BAM complex in *E. coli* BL21 Star™ (DE3)

The functional replacement of *bamA* with *bamK* was achieved in a *K. pneumoniae* model supporting the idea that *bamK* encodes for a functional Omp85 insertase. In a previous study by Volokhina *et al.* (2013), the authors demonstrated that BamA homologs from more diverse Proteobacteria: *Neisseria meningitidis*, *Neisseria gonorrhoeae*, *Bordetella pertussis*, *Burkholderia mallei*, could not functionally replace *E. coli* BamA. To further explore the species-specificity of BamA and whether closely related Proteobacteria member *K. pneumoniae* could functionally complement an *E. coli* Δ*bamA* mutant with *bamK* or Kp*bamA* (*bamA* from *K. pneumoniae* B5055). Using the protein expression strain *E. coli* BL21 Star™ (DE3), the essential *bamA* gene was replaced by either *bamA* - BN49_4147 (resultant strain: *E. coli* BL21 Star™ (DE3) Δ*bamA*::Kp*bamA*) or *bamK* - BN49_4981 (resultant strain: *E. coli* BL21 Star™ (DE3) Δ*bamA*::Kp*bamK*) using a similar "gene gorging" strategy as performed in *K. pneumoniae* B5055 (Figure 5.5.1). The resultant *E. coli* gene replacement strains which now harboured *K. pneumoniae* B5055 BamA (Δ*bamA*::Kp*bamA*) or BamK (Δ*bamA*::Kp*bamK*) were viable, had similar total membrane protein profiles and growth kinetics matching the parental isogenic strain (Figure 5.5.2, A & B).
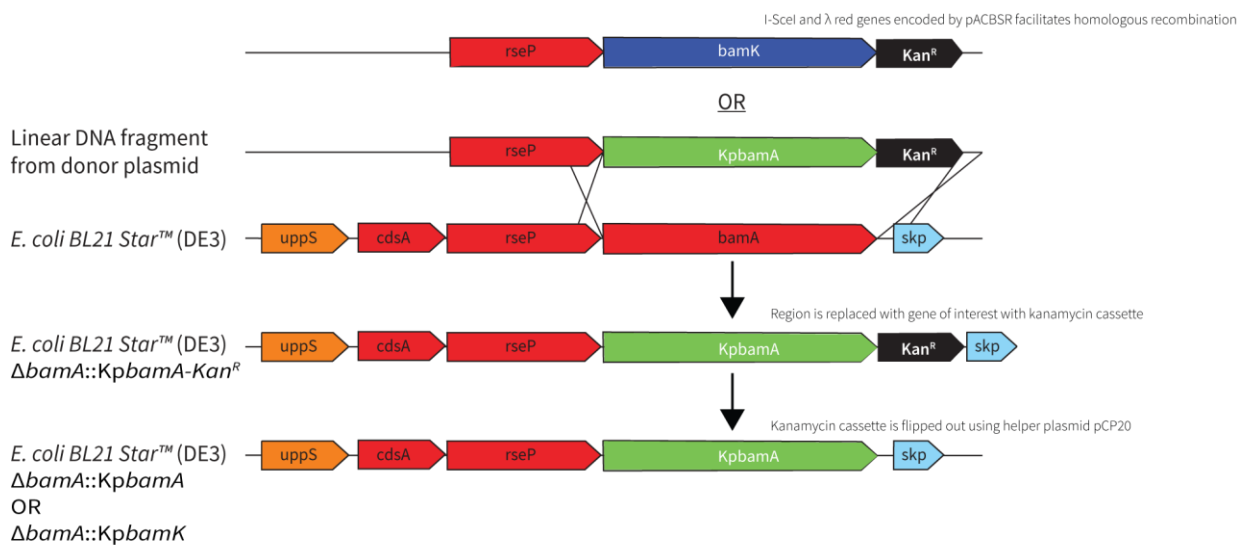
Figure 5.5.1: Strategy for genomic gene replacement of the *bamA* with Kp*bamA* or Kp*bamK* in *E. coli* BL21 Star™. Schematic showing the replacement of the *bamA* coding sequence with the *bamK* coding sequence to produce gene replacement strain *E. coli* BL21 Star™ (DE3) Δ*bamA*::Kp*bamA*. For comparative studies comparing the functionality between BamA and BamK, *E. coli* BL21 Star™ (DE3) Δ*bamA*::Kp*bamK* was also constructed in a similar manner. The resulting gene replacement strains either had Kp*bamA* or Kp*bamK* under the transcriptiona control of the *E. coli* BL21 Star™ (DE3) *bamA* promoter.
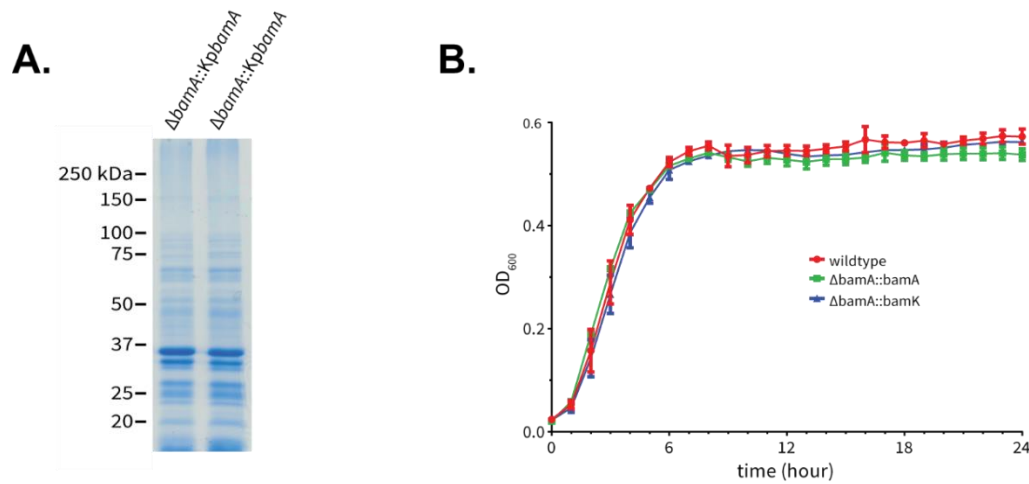
Figure 5.5.2: Comparison between *E. coli* BL21 Star™ (DE3) Δ*bamA*::Kp*bamA* and *E. coli* BL21 Star™ (DE3) Δ*bamA*::Kp*bamK*: Total membrane profiles and growth. (A) Total membranes isolated from *E. coli* BL21 Star™ (DE3) Δ*bamA*::*KpbamA* and *E. coli* BL21 Star™ (DE3) Δ*bamA*::*KpbamK* were analysed by SDS-PAGE (4-16%) and Coomassie stained, no major differences in major porins or other OMP levels were seen. (B) Growth curves comparing the parent *E. coli* BL21 Star™ (DE3) strain to the engineered *E. coli* BL21 Star™ (DE3) (Δ*bamA*::*KpbamA* and Δ*bamA*::*KpbamK*) gene replacement isogenic strains. (n=3, error bars represent SD)

To determine if BamK was comigrating with the BAM subunits (BamB through BamE), isolated outer membranes from the *E. coli* replacement strains analysed by BN-PAGE and immunoblotted with the indicated antibodies revealed a comigration pattern of BamK with the other *E. coli* BAM complex subunits in both replacement strains (Figure 5.5.3). KpBamA also reconstituted as the core of the *E. coli* BAM complex as it comigrated with the BAM subunits despite being from a non-native species (i.e. *K. pneumoniae* B5055).
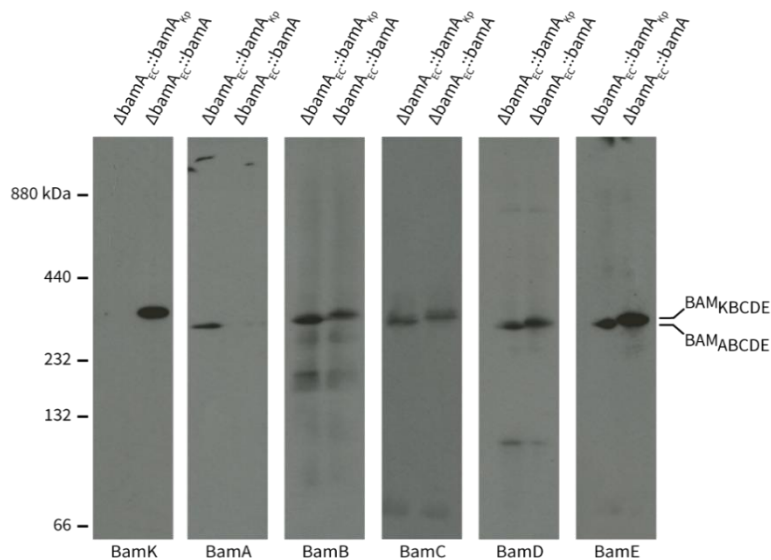
Figure 5.5.3: BN-PAGE analysis on total membranes of *E. coli* BL21 Star™ (DE3) Δ*bamA*::Kp*bamA* and *E. coli* BL21 Star™ (DE3) Δ*bamA*::Kp*bamK*. Outer membranes were isolated from B5055 wildtype and B5055 Δ*bamA*::*bamK*, analysed by BN-PAGE (5-16%) then immunoblotted for BamK and BAM complex proteins (BamA through BamE).

To confirm if BamK was actually interacting with the BAM subunits and not just coincidentally comigrating in *E. coli* BL21 Star™ (DE3), subsequent gene deletions (*bamB* and *bamC*) were

constructed in parent strain *E. coli* BL21 Star™ (DE3) Δ*bamA*::*KpbamK*, analysed by BN-PAGE analysis then immunoblotted with the indicated antibodies. The mobility of the suspected BAM_{KBDE} and BAM_{KCDE} complexes migrated to faster positions in both *ΔbamA*::*KpbamK,bamB* and *ΔbamA*::*KpbamK,bamC* validating interactions between BamK and BAM lipoproteins, BamB and BamC (Figure 5.4.4). An even smaller sub-complex was also observed which may correspond to a BAM_{KB} architecture. A comparable modular architecture represented by small sub-complexes of the BAM was previously reported in *E. coli* (Sklar *et al.*, 2007a, Kim *et al.*, 2007, Vuong *et al.*, 2008, Hagan *et al.*, 2010, Noinaj *et al.*, 2011, Webb *et al.*, 2012b).
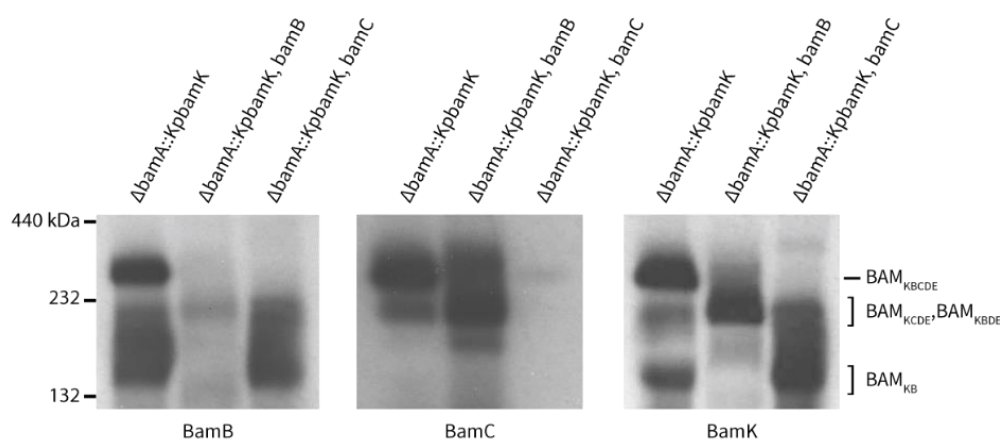
Figure 5.5.4: BN-PAGE analysis on total mebranes of *E. coli* BL21 Star™ (DE3) Δ*bamA*::Kp*bamA*, *bamB* and *E. coli* BL21 Star™ (DE3) Δ*bamA*::Kp*bamK*, *bamC*. Outer membranes isolated from *E. coli* BL21 Star™ (DE3) Δ*bamA*::*KpbamA, bamB* and *E. coli* BL21 Star™ (DE3) Δ*bamA*::*KpbamK, bamC* analysed by BN-PAGE (5-16%) then immunoblotted with the indicated antibodies. Performed by Dr. Chris Stubenrauch, Monash University.

5.6 An *in vitro* assay to measure BamK function

Because it was clear BamK could functionally replace BamA in both *K. pneumoniae and E. coli* models as demonstrated by similar membrane protein profiles, growth kinetics and homologous binding protein partners; I therefore wanted to test whether protein assembly rates of BAM substrates are affected. The assembly of outer membrane protein complexes can be observed *in vivo* and measured using a pulse-chase assay and semi-native PAGE where monomeric units assemble into mature outer membrane protein complexes (Heinz *et al.*, 2016, Stubenrauch *et al.*, 2016). To determine whether outer membrane protein assembly rates of canonical BAM substrates between BamA or BamK differed, the previously described replacement strains in *E. coli* BL21 Star™ Δ*bamA*::Kp*bamA* and *E. coli* BL21 Star™ (DE3) Δ*bamA*::Kp*bamK* were used as models due to their controllable protein expression systems with IPTG-inducible T7 promoters (Studier, 2005, Alexander *et al.*, 1992). To determine and compare BAM complex substrate assembly rates between the two replacement strains, two canonical BAM complex substrates were chosen. PhoE is a homotrimeric porin and LptDE which is a heterodimeric complex (LptD-LptE) (Figure 5.5.1). In the replacement *E. coli* strains, the respective membrane proteins were introduced via IPTG-inducible plasmids that

encode either PhoE or LptD and LptE. These strains were assayed cccording to a previously established standard lab assay (Chapter 2.0, Section 2.3) (Stubenrauch *et al.*, 2016). Samples were taken at the indicated timepoints then analysed by semi-native PAGE for subsequent detection via phosphor storage screens. It should be noted that observation of folding states and conformational stability (i.e. trimeric or dimeric) of β-barrel proteins can be further differentiated due to their unique property of being heat modifiable hence a heat treatment of select duplicate samples were also taken (Noinaj *et al.*, 2015).
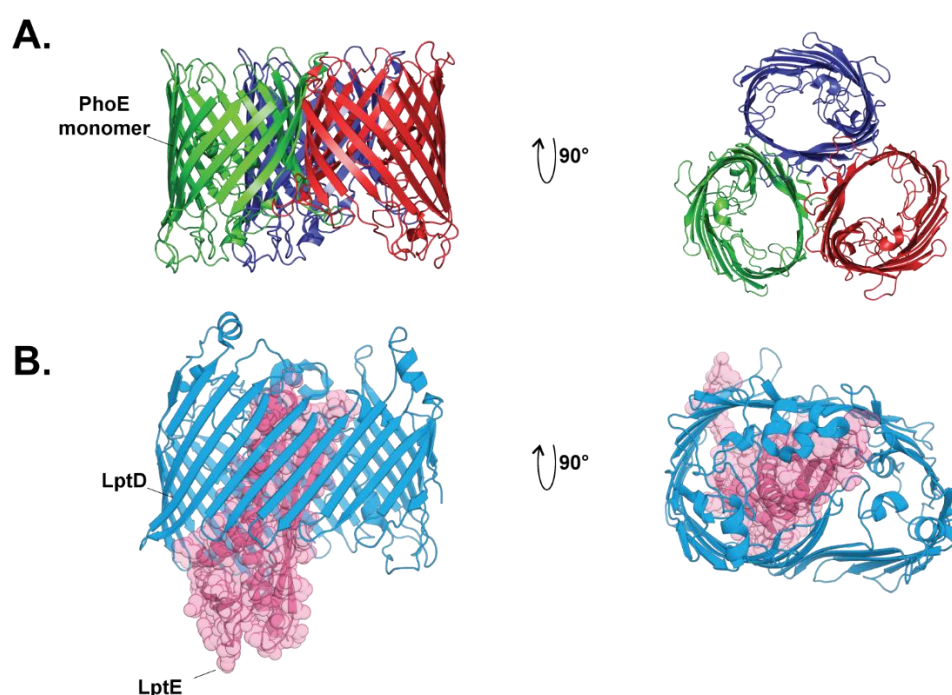
Figure 5.6.1: Biological assemblies of PhoE trimer and LptDE dimer. (A) Structure of the PhoE trimer (1PHO). The mature porin is a homotrimer of PhoE monomers (red, blue and green), a top down view shows the "open" conformation of the channel. (B) Structure of LptDE dimer complex (PDB:4RHB). In the correctly assembled LptDE, the lipoprotein subunit LptE (pink) sits in the lumen as a core stabilising the β-barrel protein LptD (blue).

PhoE is an osmotically regulated homotrimeric porin in *E. coli* with its expression normally induced by phosphate deprivation (Meyer *et al.*, 1990, Vilain *et al.*, 2002). The quaternary structure adopted by PhoE is the grouping of three β-barrel monomers arranged at the vertices of a triangular formation (Cowan *et al.*, 1992). As seen in panels of Figure 5.6.2 A-B, the observation of a discrete band which solely represents a homotrimeric species was difficult to visualise due to a smear presumably representing a mix of oligomeric PhoE species (dimeric and trimeric). A similar smear or oligomeric "ladder" feature has been observed for trimeric porin, LamB (Misra *et al.*, 1991, Wu *et al.*, 2005, Arunmanee *et al.*, 2016), which the authors speculated its correspondence to different oligomeric intermediates with varying stable conformations or binding of LPS molecules. The dimeric and trimeric species were differentiated from monomeric PhoE bands by including a duplicate heated

sample where the dimeric/trimeric interfaces are disrupted due to the heat denaturation and SDS present in the loading buffer. Hence, for densitometry analysis, I monitored the decrease in monomeric PhoE band over time which should correlate to an increase in oligomeric species. As seen in panels of Figure 5.5.2 C-D, the observed rate constants calculated from densitometry time-point data fitted with first order curves is summarised in a histogram. The assembly of PhoE, into oligomers over time was achieved in both replacement strains ($\Delta bamA$::Kp$bamA$ and $\Delta bamA$::Kp$bamK$) but the difference in observed rate constants was calculated to not be significant under the tested conditions.
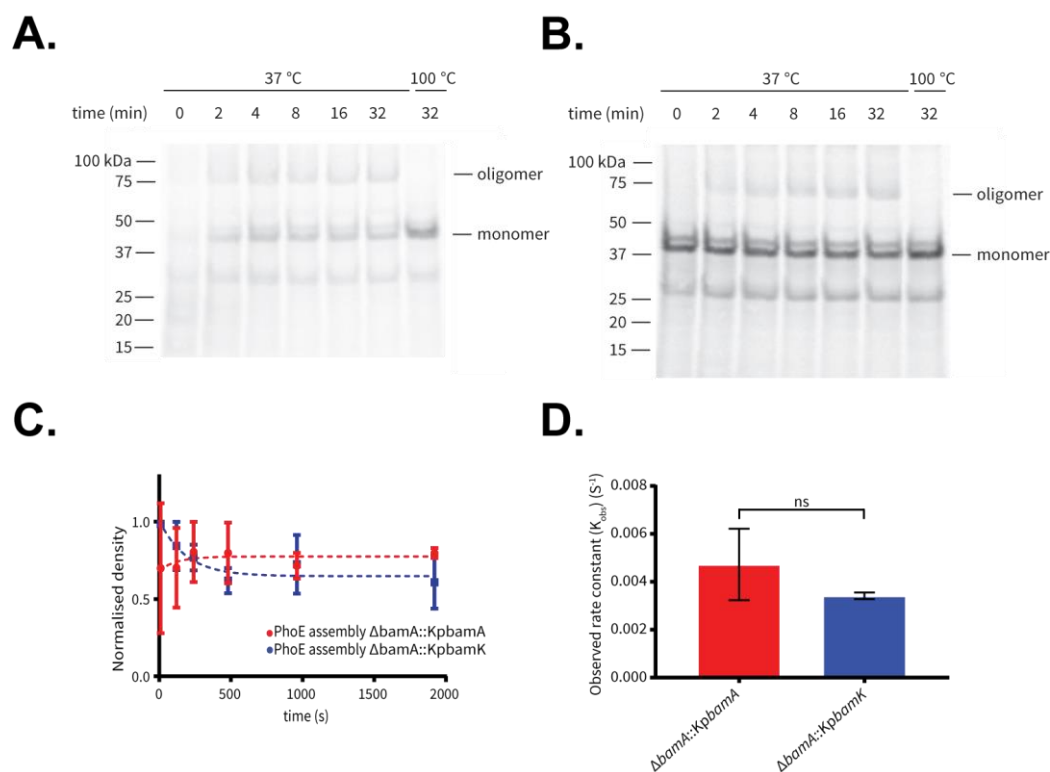
Figure 5.6.2: Assembly of PhoE oligomers in E. coli BL21 Star™ ΔbamA::KpbamA and ΔbamA::KpbamK. (A) The rate of assembly of [^{35}S]-PhoE into PhoE oligomers was assessed in the *E. coli* strains: $\Delta bamA$::Kp$bamA$ or (B) $\Delta bamA$::Kp$bamK$, using a pulse chase assay as described previously (Stubenrauch *et al.*, 2016) and in the methods section (Chapter 2.0). (C) PhoE trimer assembly was calculated by plotting the normalised density of the corresponding monomeric band over time fitted to a curve following a decay equation. The decrease of the monomeric band over time was assumed to be in proportion to the oligomerisation of PhoE monomers into functional trimers. (D) Histogram of the observed rate constants for appearance of the PhoE trimer. Error bars correspond to SEM of fit (n=3). The statistical significance was determined by an unpaired two-tailed Student's t test (ns, not significant). Panels (A) and (B) are representatives of three repeated experiments.

The incorporation of LPS onto the outer membrane outer leaflet of Gram-negative is an essential process for cell viability and is enabled by crucial steps by the dimeric LptDE complex (Bowyer *et al.*, 2011, Botos *et al.*, 2016). LptDE has a novel quaternary structure comprised of a 26-stranded β-barrel, LptD, with the lipoprotein LptE docked into the lumen of the LptD β-barrel (Dong *et al.*, 2014, Qiao *et al.*, 2014). Initial studies showed that LptDE assembly can be quantified directly (Torres *et al.*, 2018). Thus, unlike the densitometry analysis performed for PhoE I instead quantified the increasing bands corresponding to the LptDE dimer complex over time (Figure 5.6.3 A-B). The

LptDE oligomeric species was differentiated from its respective subunit species (i.e. LptD and LptE) by including a duplicate heated sample were the LptDE dimer is destabilised due to extreme heat and SDS. It should be noted that the LptDE dimer migrates faster than LptD, alone as LptE stabilises the LptD β-barrel forming a compact architecture. When plotted, a first order curve to the normalised intensities was fitted and summarised in a histogram (Figure 5.6.3 C-D). The assembly of LptDE into a dimer complex from monomers (LptD and LptE) over time was achieved in both replacement strains (ΔbamA::KpbamA and ΔbamA::KpbamK) but the difference in observed rate constants was calculated to not be significant under the tested conditions. Altogether, these data suggest BamK can functionally replace BamA as the core of the BAM complex at a seemingly analogous substrate assembly level which results in viable cells with similar total membrane protein profiles.
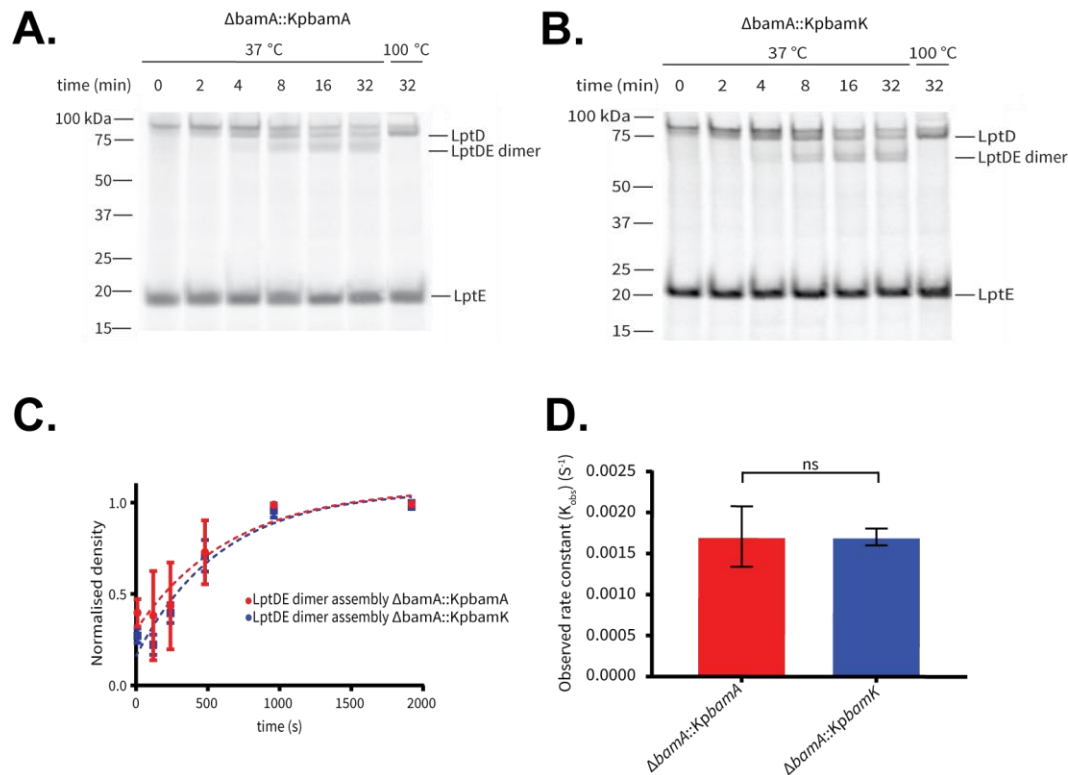
Figure 5.6.3: Assembly of LptDE dimers in *E. coli* BL21 Star™ Δ*bamA*::Kp*bamA* and Δ*bamA*::Kp*bamK*. (A) The rate of assembly of [^{35}S]-LptD and [^{35}S]-LptE into the LptDE complex was assessed in the *E. coli* BL21 Star™ strains: Δ*bamA*::Kp*bamA* or (B) Δ*bamA*::Kp*bamK*, using a pulse chase assay as described previously (Stubenrauch *et al.*, 2016) and in the methods section (Chapter 2.0). (C) A representative fitted curve demonstrating LptDE dimer assembly was calculated by plotting the normalised density of each dimeric band over time to an exponential growth equation. The accumulation of the dimeric band over time was assumed to be in proportion to the oligomerisation of LptD and LptE monomers into a functional dimer. (D) Histogram of the observed rate constants for appearance of the LptDE dimer. Error bars correspond to SEM of fit (n=3). The statistical significance was determined by an unpaired two-tailed Student's t test (ns, not significant). Panels (A) and (B) are representatives of three repeated experiments.

5.7 Differences in the extracellular loops of BamA and BamK

Given the observations that BamK is a functional isoform of BamA, I revisited areas of sequence polymorphisms between the paralogous proteins for contextual clues of its specific function. Previous structural studies have shown that BamA extracellular loop 4 (L4), loop 6 (L6) and loop 7 (L7) contribute to a cell surface dome which stabilises the β-barrel domain (Browning *et al.*, 2013, Bakelar *et al.*, 2016, Gu *et al.*, 2016, Han *et al.*, 2016). Comparative sequence analysis between BamA and BamK focusing on the sequences of these loops highlighted some distinct differences. In extracellular loops: L4, L6 and L7, notable truncations are observed resulting in presumably "shorter" loops compared to BamA (Figure 5.7.1 A). More specifically in relation to L6, the essential VRGF BamA motif found in the analogous extracellular L6 is also conserved in BamK sequences. Mutagenesis studies have shown any alterations to the essential VRGF motif found in L6 of BamA compromises the folded state of the Omp85 β-barrel domain thereby impeding proper function (Leonard-Rivera & Misra, 2012, Noinaj *et al.*, 2013, Wzorek *et al.*, 2017, Thoma *et al.*, 2018). Altogether, these polymorphisms could suggest the loops (L4, L6 and L7) of BamK function in a similar manner to BamA as a capping "dome" (Figure 5.7.2 B) but does not reveal a mechanism by which it provides a selective advantage.

Differences in the lengths of these same three loops of BamA have been noted before in a previous study which identified and isolated Shiga toxin-encoding (Stx) phage-resistant *E. coli* strains (Smith *et al.*, 2007). Under pressure from Shiga toxin-encoding (Stx) phage infection, suppressor mutants with polymorphisms mapping to *bamA* were strikingly similar to the extracellular loops of *Klebsiella* BamK (Figure 5.7.1 A). Structural modelling of BamA(Φ) and BamK, suggests substantive changes in the dome of the β-barrel compared to a solved crystal structure of BamA (PDB: 5EKQ) mostly due to the different lengths of L4, L6 and L7 (Figure 5.7.1 C-D). In all three structures, the loops form a cap but more obvious revisions of L6 are apparent in the BamA(Φ) and BamK homology models when compared to structural template BamA (PDB: 5EKQ). In L4 and L7, more subtle differences are also noted which together with L6 many completely change binding interfaces but also conformation of the Omp85 monomer. These conserved mutations in the loop regions of BamK may point to an adaptive evolutionary response in response to hostile phage or other loop binding agents in the environment.

A.

```
                              loop 4                      β-strand 8              β-strand 9
KpBamA  PQVAMWRYLNSMGQYPDNTNDRNS----FSANDFTFNYGWTYNKLDRGFFPTEGSRVNLN  596
EcBamA  PQVAMWRYLYSMGEHPSTSDQDNS----FKTDDFTFNYGWTYNKLDRGYFPTDGSRVNLT  596
EcBamA(Φ) PQVAMWRYLYSMGEHPSVVGENVKSSADFKTDDFTFNYGWTYNKLDRGYFPTDGSRVNLT  596
KpBamK  PELTTWRYLSSRGIEPSVVTKDGDSGAKYSANDYFVSLGWGYNDLDRGFFPRAGNKSSLS  600
```

```
                                         loop 6             β-strand 12
KpBamA  SSTVRGFQSNTIGPKAVYFPASSRHDDDDSYDNECKSTESAP--CKSDDAVGGNAMAVAS  714
EcBamA  SSTVRGFQSNTIGPKAVYFPHQASN-YDPDYDYECATQDGAKDLCKSDDAVGGNAMAVAS  715
EcBamA(Φ) SSTVRGFQSNTIGPKAVYLYKDGSP-----------KKESP-------DAVGGNAMAVAS  715
KpBamK  SSSVRGFSSNTIGPKAAYYRCNGSESSY---SACPLDASS--------DAVGGNAMAVLN  709
```

```
        β-strand 12              β-strand 13        loop 7             β-strand 14
KpBamA  LELITPTPFISDKYANSVRTSVFWDMGTVWDTHWDSSA---YAGYPDYSDPSNIRMSAGI  771
EcBamA  LEFITPTPFISDKYANSVRTSFFWDMGTVWDTNWDSSQ---YSGYPDYSDPSNIRMSAGI  772
EcBamA(Φ) LEFITPTPFISDKYANSVRTSFFWDMGTVWDTNWENTAETLKAGYPDYSDPSNIRMSAGI  772
KpBamK  SEFIIPTPFVNDKYANSLRTSLFVDAGTVWSTSWHNTAQTLAAGIPDYGDPSHIRLSAGI  769
```

B.

C.

D.

Figure 5.7.1: Comparative sequence and structural analysis between *E. coli* BamA, *E. coli* BamA(Φ) and *K. pneumoniae* BamK. (A) Multiple sequence alignment with BamA from *E. coli* (*Ec*BamA) highlights the similarities and differences in *Kp*BamA and *Kp*BamK, with comparison to the modification sites found in phage-resistant *Ec*BamA(Φ) (Smith *et al.*, 2007). (B) Structure of the *E. coli* BamA (PDB: 5EKQ) highlighting the extracellular strands which contribute to a dome which covers the lumen of the barrel (Noinaj *et al.*, 2013, Ni *et al.*, 2014b) (C) Cartoon representations of homology models of the β-barrel domains of BamA from phage-resistant *E. coli* MC1061- MRL1(*Ec*BamA(Φ)) and BamK from *K. pneumoniae* B5055 (*Kp*BamK) alongside the crystal structure BamA from (*Ec*BamA). (D) Structural overlay of homology models, *Ec*BamA(Φ) and BamK, with crystal structure of BamA (PDB:5EKQ). Residues composing surface-exposed loops 4, 6, and 7 are coloured magenta, cyan, and yellow, respectively Homology models were generated using the Phyre2 prediction server (Kelley *et al.*, 2015).

5.8 Discussion

In Gram-negative bacteria, the archetypal Omp85 BamA is involved in the efficient assembly and insertion of β-barrel proteins into the outer membrane (Ricci *et al.*, 2012, McCabe *et al.*, 2017, Hussain & Bernstein, 2018). In previous comprehensive phylogenetic analyses on the Omp85 superfamily and members, it was hypothesised that *tamA* evolved from an early gene duplication event of *bamA* (Heinz & Lithgow, 2014, Heinz *et al.*, 2015). Following gene duplication, it is usually assumed a redundant copy becomes a non-functional pseudogene over time due to the accumulation of mutations; for example, the protein coding region may contain a premature stop codon, or a frameshift mutation, or an internal deletion or insertion relative to the native sequence (Lerat & Ochman, 2005, Gil & Latorre, 2012). In other cases, these acquired mutations can give to rise to a new functional role (neofunctionalisation), or a specialised functional role of its ancestral paralogue (subfunctionalisation) (Shakhnovich & Koonin, 2006, Espinosa-Cantú *et al.*, 2015). Of the two processes, subfunctionalisation is thought to apply to another conserved Omp85 protein TamA, where its function has been shown to be involved in the efficient assembly of at least some virulence associated β-barrel proteins in concert with BamA (Stubenrauch *et al.*, 2016, Heinz *et al.*, 2016). In a similar vein, *bamK* represents another gene duplication event of *bamA* and may reflect a novel subfunctionalisation or even neofunctional role in OMP biogenesis. However, from the available data I report *bamK* is in fact a functional isoform of *bamA* and we will discuss what implications in *K. pneumoniae* outer membrane biogenesis this may have.

The construction of the gene replacement strain *K. pneumoniae* B5055 *ΔbamA*::Kp*bamK* (Figure 5.3.1) allowed investigation of the functional role and constraints of *bamK* in the context of substituting essential *bamA*. As described in Chapter 3.0, *bamK* was expressed at very low levels under standard laboratory conditions making studying the protein difficult and taken together with Section 5.2, may be reflective of a gene with cryptic regulation. Previous efforts by Dr. Abigail Clements (Imperial College London, UK) to generate viable *bamA* deletions in *K. pneumoniae* B5055 was not achieved suggesting that the *bamA* gene in Gram-negative bacterium *K. pneumoniae* B5055 is essential in cell viability a phenomenon also seen in other phylogenetically related organisms such as *E. coli* and *Salmonella* spp. (Baba *et al.*, 2006, Goodall *et al.*, 2018, Rowley *et al.*, 2011). The attempted deletion of *bamA* in *K. pneumoniae* B5055 was lethal presumably due to the low expression levels of *bamK*, which could not rescue BamA function. Upon changing the upstream promoter elements of *bamK* with the constitutively active promoter of *bamA*, the ensuing BamK product was expressed and found to functionally replace essential BamA as shown through the various protein-protein interaction assays (Section 5.4-5.5) and substrate assembly studies (Section 5.6).

The functional gene replacement strains of BamA in both *E. coli* and *K. pneumoniae* is fascinating as previous studies have reported high degrees of species specificity in the functioning of the BamA component for efficient folding and insertion of OMPs (Volokhina *et al.*, 2013, Browning *et al.*, 2015). In these aforementioned studies, the authors investigated if chimeric BamA proteins, carrying either the β-barrel or POTRA domains from various BamA orthologues, can functionally replace *E. coli* BamA. Results from these studies demonstrated the C-terminal β-barrel domains are functionally interchangeable between bacterial species but the fusion to non-native/orthologous N-terminal POTRA domains from a different bacterial species generally resulted in OMP assembly defects. The authors suspected this was due to disruption of coordinated interactions between the β-barrel domain and cognate POTRA domains of the BamA protein for proper function. Results presented in this chapter extends on this by showing that the full-length BamA and BamK from *K. pneumoniae*, while still in the same major phylum of Gram-negative bacteria (Figure 5.8.1), was able to rescue BamA function in *E. coli and K. pneumoniae* despite sequence differences in the POTRA and β-barrel domains.
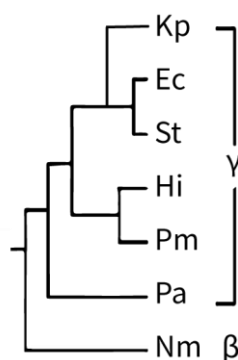
Figure 5.8.1: Cladistic analysis of Gram-negative bacteria based on the 16s rRNA gene. Depicted here is the genetic relation of select proteobacteria (subclasses: Gamma and Beta) used in this study and the other BamA species-specificity studies (Volokhina *et al.*, 2013, Browning *et al.*, 2015). In the study by Volokhina *et al.* (2013), functional gene replacement between *Neisseria meningitis* (Nm) and *E. coli* (Ec) could not be achieved. In a later study by Browning *et al.* (2015), protein chimeras keeping the native *E. coli* (Ec) BamA β-barrel domain but with different POTRA domains from: *Salmonella enterica* serovar Typhimurium (St), *Haemophilus influenzae* (Hi), *Pasteurella multocida* (Pm), *Pseudomonas aeruginosa* (Pa) and *Neisseria meningitidis* (Nm) were assayed for their ability to rescue BamA depletion. Only *Salmonella enterica* serovar Typhimurium BamA POTRA chimera was able to support growth in an *E. coli* JWD3 BamA depletion strain. This study shows that full-length BamA and BamK from *Klebsiella pneumoniae* (Kp) was able to functionally replace BamA in *E. coli*.

From a structural standpoint, the protein domain architectures between BamA and predicted BamK are identical with five N-terminal periplasmic POTRA domains followed by a C-terminal β-barrel OMP based on protein prediction algorithms and sequence similarity (Figure 5.1.2). As seen in the multiple sequence alignments (Figure 5.1.3), the amino acid sequences between BamA and BamK has high conservation in their POTRA domains where critical contacts for protein-protein interactions are known to occur between BamA to cognate partners: BamB, BamC, BamD and BamE (Gu *et al.*, 2016, Han *et al.*, 2016, Bakelar *et al.*, 2016, Iadanza *et al.*, 2016). This was confirmed through BN-

PAGE analysis on total membranes prepared from functional gene replacement strain *K. pneumoniae* B5055 *ΔbamA*::Kp*bamK* where immunoblotting for BamK showed a comigration pattern with the other BAM lipoproteins suggesting a BAM complex (BAM$_{KBCDE}$) with BamK serving at its core (Figure 5.4.6). Interestingly, a similar comigration pattern was also observed in functional gene replacement strain *E. coli* BL21 Star™ *ΔbamA*::Kp*bamK* suggesting that the high genetic similarity between the two organisms allowed cross species interaction between the non-native Omp85 BamK and the equivalent *E. coli* BL21 Star™ BAM lipoproteins (Figure 5.5.3). To further confirm if BamK was indeed interacting with the BAM components and not just merely comigrating, double gene deletion strains in *E. coli* BL21 Star™. (*ΔbamA*::Kp*bamK, bamB* and *ΔbamA*::Kp*bamK, bamC* BL21 Star™) were engineered for additional BN-PAGE analysis. The mobility of the suspected BAM$_{KBDE}$ and BAM$_{KCDE}$ complexes migrated to lower positions in both *ΔbamA*::Kp*bamK, bamB* and *ΔbamA*::Kp*bamK, bamC* validating interactions between BamK and BAM lipoproteins, BamB and BamC (Figure 5.5.4). Similar modular architectures of the BAM complex have been reported previously in *E. coli* (Sklar *et al.*, 2007a, Kim *et al.*, 2007, Vuong *et al.*, 2008, Hagan *et al.*, 2010, Noinaj *et al.*, 2011, Webb *et al.*, 2012b) further supporting the idea that BamK is a functional isoform to BamA as illustrated by its same binding capacities to cognate BAM partners.

Aiming to further dissect if the functional gene replacement strain impacted outer membrane biogenesis at the membrane protein level, I performed a comparative analysis of isolated total membranes from *K. pneumoniae* B5055 *ΔbamA*::*bamK* (Figure 5.4.5 A) and *E. coli* BL21 Star™ *ΔbamA*::Kp*bamK* (Figure 5.5.2 A) compared to their isogenic parent strains. No marked difference was observed when comparing the profiles suggesting when BamK serves as the core of the BAM complex the OMP substrates for assembly and insertion are most likely one and the same but at potentially different efficiencies. Hence, to elucidate if any assembly rate differences exist between a BamA or BamK, canonical OMP substrates PhoE and LptDE were subjected to biochemical assembly assays as previously described (Stubenrauch *et al.*, 2016). PhoE is a class of porin in which the functional form is a trimer (Meyer *et al.*, 1990, Vilain *et al.*, 2002, Arunmanee *et al.*, 2016), whilst LptDE is a unique hetero-dimeric protein complex comprised of the LptD β-barrel with lumen situated lipoprotein plug LptE (Bowyer *et al.*, 2011, Botos *et al.*, 2016). Through assembly assays performed in functional gene replacement *E. coli* BL21 Star™ strains (*ΔbamA*::Kp*bamA* and *ΔbamA*::Kp*bamK*), the assembly rates between BamA and BamK are identical (Figure 5.6.1 D, Figure 5.6.2 D). These data still suggest that *bamK* encodes for a functional protein able to serve as a novel component in BAM complexes in environmental scenarios that have yet to be identified.

The essential BamA plays an important functional role in OMP biogenesis, but it has been implicated from a number of studies that the extracellular loop regions act as a receptor for ligand binding which can adversely affect bacterial growth (Smith *et al.*, 2007, Aoki *et al.*, 2008, Urfer *et al.*, 2016, Ghequire *et al.*, 2018). As previously mentioned, it was found that sites within the extracellular loops (4, 6 and 7) of BamA (formerly called YaeT or Omp85) of *E. coli* were the predominant elements recognised by Stx phage for adsorption and subsequent infection (Smith *et al.*, 2007). In a study by Aoki *et al.* (2008), it was found that contact-dependent growth inhibition (CDI) as an intra-species competitive strategy was regulated in *E. coli* cultures by the CdiA/CdiB two partner secretion system identifying BamA as the likely outer membrane receptor. In a different study, it was found that lectin-like bacteriocin LlpA from various *Pseudomonas* strains kills related species (or even their own) as an interspecies control strategy which again acknowledged the surface exposed loop 6 of BamA as the most likely receptor (Ghequire *et al.*, 2018). Finally, it was also shown through photolabeling and fluorescence microscopy experiments, defensin-like peptidomimetic compound JB-95 potentially bound to BamA on its extracellular side again inhibiting cellular growth (Urfer *et al.*, 2016).

When compared to BamA, a closer inspection of the predicted BamK sequence reveals some differences which could highlight a neofunctional role of the protein in bacterial survival. In *E. coli*, select cell surface exposed loops (4, 6 and 7) of BamA are recognized by Stx phage as a receptor for subsequent adsorption and cell lysis facilitated through the conserved tail spike protein of the bacteriophage (Smith *et al.*, 2007). In the same study, the authors recovered phage-resistant *E. coli* mutants which were attributed to changes in the extracellular receptor loops of BamA. Curiously, it was found through a multiple sequence alignment, truncations and insertions in the loops of the phage-resistant BamA variant mirrored quite closely to the *K. pneumoniae* B5055 BamK sequence (Figure 5.7.1 A). One could imagine these evolutionary divergent changes in BamK provides a selection advantage where Stx phage recognition via extracellular loop ligand binding is abrogated as polymorphism in the loops have altered the binding interfaces found within the cell surface dome of normal BamA (Figure 5.7.1 B-C). Whether these BamK polymorphisms actually provide a mechanism of Stx phage-resistance remains to be investigated, but this idea runs in parallel with other occurrences where the extracellular loops of BamA acts as a receptor in inter- and intraspecies competition for bacterial growth as previously discussed (Aoki *et al.*, 2008, Ghequire *et al.*, 2018). Here, the idea is put forward that BamK is a functional Omp85 which is expressed as a phage-resistant mechanism or possibly a part of a decoy receptor strategy to improve bacterial fitness. As previously discussed, antagonistic interactions can result in signalling cascades which places restrictions on intraspecies growth and spread (Aoki *et al.*, 2008). Contrarily, decoy receptor mechanisms are another

strategy of some bacterial species where isoforms of a receptor are both present and expressed (Paulus & van der Hoorn, 2018), but one may be of higher affinity or abundance and therefore acts as a sponge for adverse agents such as phage and toxins (Figure 5.8.2). It should be noted the precedence of additional BamA-like paralogues encoded within genomes in not only widespread in *K. pneumoniae* but also other closely related *Enterobactericiae* family members (Heinz & Lithgow, 2014, Torres *et al.*, 2018). Therefore, this hypothesised resistance/decoy mechanism extends on the concept that additional Omp85 gene copies with general insertase functions are conserved as they provide a selective advantage. In any case, the function of BamK (or BamA-like) as a moonlighting paralogue is curious and warrants further investigation.
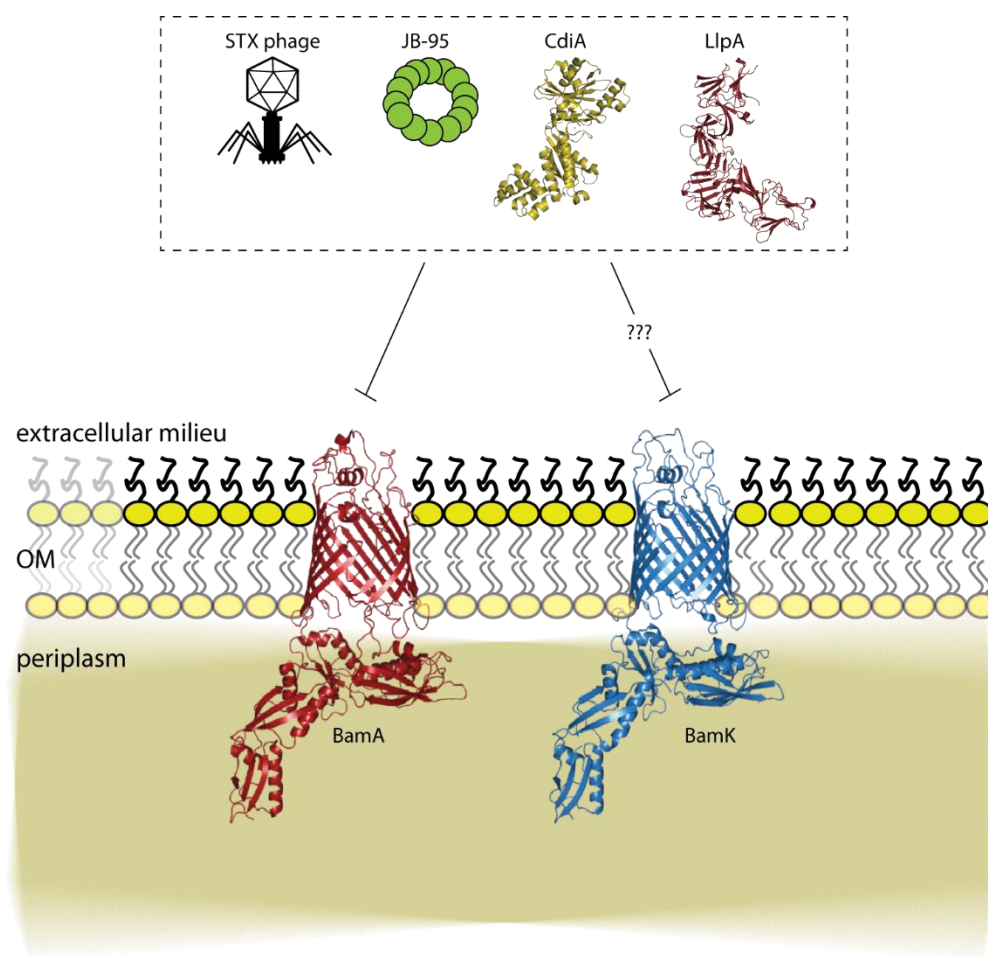
Figure 5.8.2: Schematic showing the exploitation of BamA as a receptor. Depicted here is a schematic showing agents reported to inhibit bacterial growth by binding cell surface moieties of BamA. STX phage have been reported to bind *E. coli* for adsorption mediated by extracellular loops (L4, L6 and L7) (Smith *et al.*, 2007). The macrocyclic peptide JB-95 has displayed potent antimicrobial activity against *E. coli* by binding OMPs which include BamA (Urfer *et al.*, 2016). The CdiA toxin, part of the CDI secretion system from *E. coli* EC93 recognises BamA as a cell surface receptor (Aoki *et al.*, 2008) (PDB: 4G6U). Lectin-like bacteriocins (LlpAs) are secreted proteins by proteobacteria which selectively kill strains of their own and closely related species as an intra-species competition strategy. The susceptibility of *Pseudaomonas* spp. to LlpAs has been linked to bind selectively to extracellular loops L4, L6 and L7 (Ghequire *et al.*, 2018) (PDB: 3M7H). If BamK can also be targeted by any of these reported agents remains or similar ligands remains to be investigated. OM – outer membrane. IM – inner membrane

6

Conclusion

6.0 Defining the roles of Omp85 protein family members in *K. pneumoniae*

In Gram-negative bacteria, OMPs are transmembrane protein that feature a β-barrel domain and are localised to the OM (Koebnik *et al.*, 2000). The role of OMPs are many and serve essential functions that include nutrient uptake, cell adhesion, cell signalling, virulence, membrane structural integrity and even OMP biogenesis (Wimley, 2003, Fairman *et al.*, 2011). Outer membrane protein assembly factor BamA is a member of the Omp85 superfamily and is the core protein of the β-barrel assembly machinery involved in the assembly and insertion of Gram-negative OMPs into the OM (Noinaj *et al.*, 2015, Rollauer *et al.*, 2015). Several members of the Omp85/TpsB family are known to catalyse the efficient assembly and insertion of other nascent OMPs and are unified by shared domain architectures and sequence motifs (Heinz & Lithgow, 2014). The best studied Omp85s in Gram-negative bacteria are BamA and TamA; however, it is worth mentioning other Gram-negative species harbour additional Omp85 paralogues many with unknown functions (Heinz & Lithgow, 2014, Torres *et al.*, 2018). In clinically relevant pathogen *K. pneumoniae*, additional Omp85 members, BamL and BamK (divergent to BamA and TamA), are yet to be characterised and may represent novel components in *K. pneumoniae* OMP biogenesis. In this thesis, I aimed to characterise the function and mechanism of these *K. pneumoniae* specific Omp85s through different approaches that included genetics, biochemistry, phylogenomics and molecular modeling. Although the work presented in this study does not present holistic answers as to their function and mechanism, I will briefly describe our current state of knowledge of these Omp85 members and outline some future directions.

The expression of *bamL* and *bamK* is cryptic

In Chapter 3, results presented demonstrate that transcript levels of *K. pneumoniae* Omp85 members, *bamL* and *bamK*, are expressed at very low or near background levels. This observation could reflect a cryptic genetic regulatory system for expression as transcripts were not observed under a range of standard laboratory conditions. In many cases cryptic regulation in prokaryotes usually applies to pseudogenes but as both Omp85 genes are prevalent in *K. pneumoniae*, this idea runs in stark contrast with a pseudogene scenario as the high sequence conservation and ubiquitous distribution suggests they are most likely expressed in as yet unidentified conditions. In this *K. pneumoniae* model, the relative expression levels of *bamA* were more abundant than the other three Omp85s (i.e. *bamL, bamK* and *tamA*). This suggests that OMP biogenesis in *K. pneumoniae* is likely to be similar to *E. coli*, where BamA is the major pathway for OMP assembly. Our current understanding of OMP biogenesis in a *K. pneumoniae* model posits that BamA mostly deals with the flux of substrates for OMP biogenesis and the other three Omp85s may play more specialised roles like *E. coli* TamA. While this project selected to focus on BamK, I was unable to identify conditions or genetic regulatory factors controlling its expression (i.e. transcriptional repressors, activators or regulatory sRNAs). The transcriptional regulation of *bamK* is mysterious, and it is assumed that its expression would be in relation to its isomorphic parent *bamA*. Whilst BamA is the major OMP pathway, it is tempting to speculate there could be selective pressures that would activate genetic regulatory switches that favour expression of *bamK* over *bamA*. In other Gram-negative bacteria regulatory systems that control OMP biosynthesis can be very intricate as several checks and balances tightly regulate the expression and translation of OMPs. This could highlight that a more complex regulatory network is tightly regulating the expression of these *K. pneumoniae* specific Omp85s. Future studies could further dissect the promoter regions of the *bamL* and *bamK* and examine if transcriptional activators could also be involved in the expression of these Omp85s.

The *bamL* locus encodes for outer membrane localised proteins, novel nanomachine?

In Chapter 4, I demonstrated that proteins of the *bamL* operon encoded for two distinct integrally associated OM localised proteins. The novel Omp85 member BamL of *K. pneumoniae* was suspected to represent a novel translocase OMP similar to a two-partner secretion system due to its genetic organisation with partner protein LupV. Furthermore, BamL and LupV seemed to be co-regulated and expressed as their coding regions overlapped. Although the function and mechanism were not uncovered for the genes encoded by the *bamL* locus, the BamL Omp85 protein lacked POTRA domains and therefore its proposed mechanism for outer membrane protein assembly may

be different to other bacterial Omp85s such as BamA or TamA. LupV was similarly interesting as it was predicted to have a transmembrane C-terminal OmpA-like domain with an unstructured N-terminal domain and might represent a substrate for BamL. The unstructured N-terminal region of LupV could be further investigated as its topology and protein fold were not elucidated during this study. Additional future studies should look at whether the *bamL* locus is involved in other virulence aspects such as cell adherence, cytolysis and iron acquisition. However, based on more recent studies of noNterm Omp85s, this protein member may function in a more metabolic scope, forming a channel porin in itself for diffusion of nutrients or functioning as an assembly complex which assembles and inserts OMP substrate into the OM involved in nutrient transport. Finally, all characterised Omp85s in Gram-negative bacteria have their functions linked to assembly and insertion of OMP substrates. One could speculate because BamL and LupV are presumably co- regulated and expressed, LupV may require a partner Omp85 for efficient assembly and insertion. Pulse chase assays determining if LupV requires BamL can be undertaken to ascertain this hypothesised biogenesis pathway.
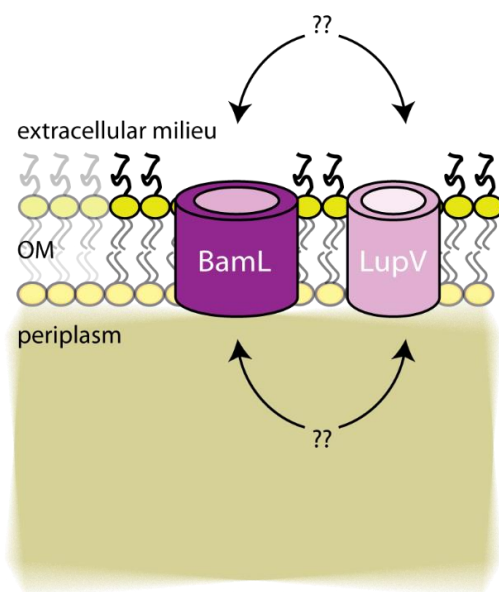
Figure 6.0.1: The roles of the BamL locus proteins are still unknown. The BamL locus proteins are integrally bound proteins localised to the OM. The function of both proteins is still unknown. Does BamL function as a translocon OMP for substrate passage? Do BamL and LupV proteins interact? Does LupV require BamL for efficient assembly and insertion?

BamK can functionally replace BamA in the BAM complex for OMP biogenesis

In Chapter 5, I presented results that demonstrate the novel Omp85 BamK of *K. pneumoniae* can functionally replace BamA in the BAM complex in *K. pneumoniae* and *E. coli*. This functional gene replacement was further evaluated and found the gene replacement strain had comparable substrate assembly rates and OMP profiles to when BamA is at the core of the BAM complex. These

finding suggests that *K. pneumoniae* genomes can encode for a functional isoform of BamA but does not reveal under what environmental conditions BamK expression would be favourable. Comparative sequence and molecular modelling analysis revealed major differences between the extracellular loops of BamK and BamA and may contribute to a mechanism for resistance against bacteriophage infections or as a decoy receptor for other lethal agents (e.g. BamA specific toxins and antimicrobial peptides). Future studies will aim to identify inducive conditions for BamK expression and determine if known growth inhibiting BamA binding factors are able to bind to BamK. Additional structural studies and molecular dynamics analysis could aid in better understanding if the nuanced differences of BamK display the same mechanistic features of BamA for OMP assembly and insertion.
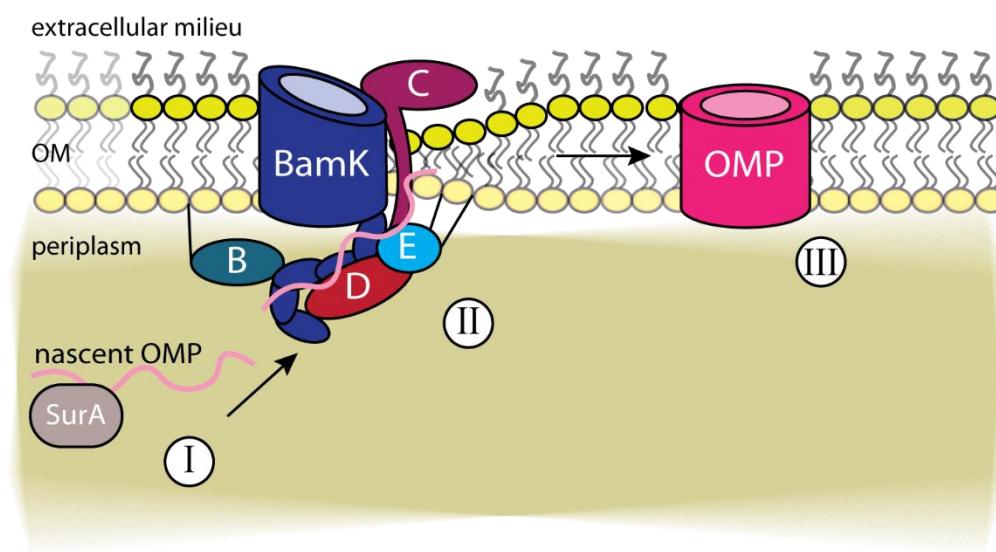
Figure 6.0.2: Schematic of OMP assembly with BamK as the core component of the BAM Complex. The *K. pneumoniae* specific Op85 can functionally replace BamA in the BAM complex for assembly and insertion of OMPs into the OM (see Figure 1.3.7). As BamK expression is presumably tightly regulated, are there certain environments/growth conditions where its expression would be more favourable than BamA? Is the mechanism for OMP assembly and insertion identical between BamA and BamK, or does BamK favour certain conformations and substrates?

The pathogenic success of *K. pneumoniae* can be largely attributed to outer surface features such as LPS, CPS, porins and many are yet to be characterised. The Omp85 featuring BAM complex impacts on the assembly and insertion of these formerly mentioned features of the OM, and the discovery of BamL and BamK may represent evolutionary adaptations of *K. pneumoniae* to new environmental niches. A better understanding of how these divergent Omp85 proteins are regulated and how each contribute to OM biogenesis would prove valuable in understanding how some pathogenic bacteria are more virulent or can better adapt to certain environmental pressures.

References

Alba, B.M., J.A. Leeds, C. Onufryk, C.Z. Lu & C.A. Gross, (2002) DegS and YaeL participate sequentially in the cleavage of RseA to activate the sigma(E)-dependent extracytoplasmic stress response. *Genes & development* **16**: 2156-2168.

Albenne, C. & R. Ieva, (2017) Job contenders: roles of the beta-barrel assembly machinery and the translocation and assembly module in autotransporter secretion. *Mol Microbiol* **106**: 505-517.

Albrecht, R., M. Sch??tz, P. Oberhettinger, M. Faulstich, I. Bermejo, T. Rudel, K. Diederichs & K. Zeth, (2014) Structure of BamA, an essential factor in outer membrane protein biogenesis. *Acta Crystallographica Section D: Biological Crystallography* **70**: 1779-1789.

Albrecht, R. & K. Zeth, (2010) Crystallization and preliminary X-ray data collection of the Escherichia coli lipoproteins BamC, BamD and BamE. *Acta Crystallographica Section F: Structural Biology and Crystallization Communications*.

Alexander, W.A., B. Moss & T.R. Fuerst, (1992) Regulated expression of foreign genes in vaccinia virus under the control of bacteriophage T7 RNA polymerase and the Escherichia coli lac repressor. *Journal of virology* **66**: 2934-2942.

Alsteens, D., N. Martinez, M. Jamin & F. Jacob-Dubuisson, (2013) Sequential Unfolding of Beta Helical Protein by Single-Molecule Atomic Force Microscopy. *PLoS ONE* **8**: 1-10.

Amit, R., A.B. Oppenheim & J. Stavans, (2003) Increased bending rigidity of single DNA molecules by H-NS, a temperature and osmolarity sensor. *Biophysical Journal* **84**: 2467-2473.

Anandan, A. & A. Vrielink, (2016) Detergents in Membrane Protein Purification and Crystallisation. In: The Next Generation in Membrane Protein Structure Determination. I. Moraes (ed). Cham: Springer International Publishing, pp. 13-28.

Andersson, H. & G. Von Heijne, (1993) Position-specific Asp-Lys pairing can affect signal sequence function and membrane protein topology. *J. Biol. Chem.* **268**: 21389-21393.

Anwari, K., C.T. Webb, S. Poggio, A.J. Perry, M. Belousoff, N. Celik, G. Ramm, A. Lovering, R.E. Sockett, J. Smit, C. Jacobs-Wagner & T. Lithgow, (2012) The evolution of new lipoprotein subunits of the bacterial outer membrane BAM complex. *Mol. Microbiol.* **84**: 832-844.

Aoki, S.K., (2005) Contact-Dependent Inhibition of Growth in Escherichia coli. *Science* **309**: 1245-1248.

Aoki, S.K., J.C. Malinverni, K. Jacoby, B. Thomas, R. Pamma, B.N. Trinh, S. Remers, J. Webb, B.A. Braaten, T.J. Silhavy & D.A. Low, (2008) Contact-dependent growth inhibition requires the essential outer membrane protein BamA (YaeT) as the receptor and the inner membrane transport protein AcrB. *Mol Microbiol* **70**: 323-340.

Arunmanee, W., M. Pathania, A.S. Solovyova, A.P. Le Brun, H. Ridley, A. Baslé, B. van den Berg & J.H. Lakey, (2016) Gram-negative trimeric porins have specific LPS binding sites that are essential for porin biogenesis. *Proceedings of the National Academy of Sciences* **113**.

Asmar, A.T., J.L. Ferreira, E.J. Cohen, S.H. Cho, M. Beeby, K.T. Hughes & J.F. Collet, (2017) Communication across the bacterial cell envelope depends on the size of the periplasm. *PLoS Biology* **15**.

Baba, T., T. Ara, M. Hasegawa, Y. Takai, Y. Okumura, M. Baba, K.A. Datsenko, M. Tomita, B.L. Wanner & H. Mori, (2006) Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: The Keio collection. *Molecular Systems Biology* **2**.

Bagley, S.T., (1985) Habitat association of Klebsiella species. *Infection control : IC* **6**: 52-58.

Bakelar, J., S.K. Buchanan & N. Noinaj, (2016) The structure of the beta-barrel assembly machinery complex. *Science* **351**: 180-186.

Bakelar, J., S.K. Buchanan & N. Noinaj, (2017) Structural snapshots of the β-barrel assembly machinery. *FEBS Journal* **284**: 1778-1786.

Baker, D., (2001) Protein Structure Prediction and Structural Genomics. *Science* **294**: 93-96.

Balakirev, E.S. & F.J. Ayala, (2003) Pseudogenes: Are They "Junk" or Functional DNA? *Annual Review of Genetics* **37**: 123-151.

Bamert, R.S., K. Lundquist, H. Hwang, C.T. Webb, T. Shiota, C.J. Stubenrauch, M.J. Belousoff, R.J.A. Goode, R.B. Schittenhelm, R. Zimmerman, M. Jung, J.C. Gumbart & T. Lithgow,

(2017) Structural basis for substrate selection by the translocation and assembly module of the β-barrel assembly machinery. *Mol. Microbiol.* **106**.

Barnhart, M.M. & M.R. Chapman, (2006) Curli Biogenesis and Function. *Annual Review of Microbiology* **60**: 131-147.

Bassler, J., B. Hernandez Alvarez, M.D. Hartmann & A.N. Lupas, (2015) A domain dictionary of trimeric autotransporter adhesins. *Int. J. Med. Microbiol.* **305**: 265-275.

Basso, P., M. Ragno, S. Elsen, E. Reboud, G. Golovkine, S. Bouillot, P. Huber, S. Lory, E. Faudry & I. Attrée, (2017) Pseudomonas aeruginosa Pore-Forming Exolysin and Type IV Pili Cooperate To Induce Host Cell Lysis. *mBio* **8**: 1-16.

Beacham, I.R., (1979) Periplasmic enzymes in gram-negative bacteria. *Int J Biochem* **10**: 877-883.

Bechtluft, P., A. Kedrov, D.J. Slotboom, N. Nouwen, S.J. Tans & A.J.M. Driessen, (2010) Tight hydrophobic contacts with the secB chaperone prevent folding of substrate proteins. *Biochemistry* **49**: 2380-2388.

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

Berg, B.L., C. Baron & V. Stewart, (1991) Nitrate-inducible formate dehydrogenase in Escherichia coli K-12 II. Evidence that a mRNA stem-loop structure is essential for decoding opal (UGA) as selenocysteine. *J. Biol. Chem.* **266**: 22386-22391.

Bialek-Davenet, S., A. Criscuolo, F. Ailloud, V. Passet, L. Jones, A.S. Delannoy-Vieillard, B. Garin, S.L. Hello, G. Arlet, M.H. Nicolas-Chanoine, D. Decré & S. Brisse, (2014) Genomic definition of hypervirulent and multidrug-resistant klebsiella pneumoniae clonal groups. *Emerg. Infect. Dis.* **20**: 1812-1820.

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

Blin, C., V. Passet, M. Touchon, E.P.C. Rocha & S. Brisse, (2017) Metabolic diversity of the emerging pathogenic lineages of Klebsiella pneumoniae. *Environmental Microbiology* **19**: 1881-1898.

Bohlin, J., V. Eldholm, J.H.O. Pettersson, O. Brynildsrud & L. Snipen, (2017) The nucleotide composition of microbial genomes indicates differential patterns of selection on core and accessory genomes. *BMC Genomics* **18**: 1-11.

Bölter, B., J. Soll, a. Schulz, S. Hinnah & R. Wagner, (1998) Origin of a chloroplast protein importer. *Proc. Natl. Acad. Sci. U. S. A.* **95**: 15831-15836.

Bonhivers, M., L. Plançon, A. Ghazi, P. Boulanger, M. Le Maire, O. Lambert, J.L. Rigaud & L. Letellier, (1998) FhuA, an Escherichia coli outer membrane protein with a dual function of transporter and channel which mediates the transport of phage DNA. *Biochimie* **80**: 363-369.

Bos, M.P., V. Robert & J. Tommassen, (2007) Biogenesis of the gram-negative bacterial outer membrane. In: Annual Review of Microbiology. pp. 191-214.

Bossi, L. & N. Figueroa-Bossi, (2007) A small RNA downregulates LamB maltoporin in Salmonella. *Mol. Microbiol.* **65**: 799-810.

Botos, I., N. Majdalani, S.J. Mayclin, J.G. McCarthy, K. Lundquist, D. Wojtowicz, T.J. Barnard, J.C. Gumbart & S.K. Buchanan, (2016) Structural and Functional Characterization of the LPS Transporter LptDE from Gram-Negative Pathogens. *Structure* **24**: 965-976.

Bowring, B.G., V.A. Fahy, A. Morris & A.M. Collins, (2017) An unusual culprit: Klebsiella pneumoniae causing septicaemia outbreaks in neonatal pigs? *Veterinary Microbiology* **203**: 267-270.

Bowyer, A., J. Baardsnes, E. Ajamian, L. Zhang & M. Cygler, (2011) Characterization of interactions between LPS transport proteins of the Lpt system. *Biochemical and Biophysical Research Communications* **404**: 1093-1098.

Brahmi, S., A. Touati, C. Dunyach-Remy, A. Sotto, A. Pantel & J.-P. Lavigne, (2018) High Prevalence of Extended-Spectrum β-Lactamase-Producing Enterobacteriaceae in Wild Fish from the Mediterranean Sea in Algeria. *Microbial Drug Resistance* **24**: 290-298.

Bratu, S., D. Landman, A. George, J. Salvani & J. Quale, (2009) Correlation of the expression of acrB and the regulatory genes marA, soxS and ramA with antimicrobial resistance in

clinical isolates of Klebsiella pneumoniae endemic to New York City. *Journal of Antimicrobial Chemotherapy* **64**: 278-283.

Brinkman, K.K. & R.A. Larsen, (2008) Interactions of the energy transducer TonB with noncognate energy-harvesting complexes. *J. Bacteriol.* **190**: 421-427.

Brinkworth, A.J., C.H. Hammer, L.R. Olano, S.D. Kobayashi, L. Chen, B.N. Kreiswirth & F.R. DeLeo, (2015) Identification of outer membrane and exoproteins of carbapenem-resistant multilocus sequence type 258 Klebsiella pneumoniae. *PLoS ONE* **10**: 1-16.

Brisse, S. & E. Van Duijkeren, (2005) Identification and antimicrobial susceptibility of 100 Klebsiella animal clinical isolates. *Veterinary Microbiology* **105**: 307-312.

Brombacher, E., A. Baratto, C. Dorel & P. Landini, (2006) Gene expression regulation by the curli activator CsgD protein: Modulation of cellulose biosynthesis and control of negative determinants for microbial adhesion. *J. Bacteriol.* **188**: 2027-2037.

Browning, D.F., V.N. Bavro, J.L. Mason, Y.R. Sevastsyanovich, A.E. Rossiter, M. Jeeves, T.J. Wells, T.J. Knowles, A.F. Cunningham, J.W. Donald, T. Palmer, M. Overduin & I.R. Henderson, (2015) Cross-species chimeras reveal BamA POTRA and β-barrel domains must be fine-tuned for efficient OMP insertion. *Mol. Microbiol.* **97**: 646-659.

Browning, D.F., S.A. Matthews, A.E. Rossiter, Y.R. Sevastsyanovich, M. Jeeves, J.L. Mason, T.J. Wells, C.A. Wardius, T.J. Knowles, A.F. Cunningham, V.N. Bavro, M. Overduin & I.R. Henderson, (2013) Mutational and topological analysis of the Escherichia coli BamA protein. *PLoS ONE* **8**.

Buchanan, S.K., (1999) Beta-barrel proteins from bacterial outer membranes: structure, function and refolding. *Current opinion in structural biology* **9**: 455-461.

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

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

Bushell, S.R., I.L. Mainprize, M.A. Wear, H. Lou, C. Whitfield & J.H. Naismith, (2013) Wzi Is an Outer Membrane Lectin that Underpins Group 1 Capsule Assembly in Escherichia coli. *Structure* **21**: 844-853.

Cano, V., C. March, J.L. Insua, N. Aguiló, E. Llobet, D. Moranta, V. Regueiro, G.P. Brennan, M.I. Millán-Lou, C. Martín, J. Garmendia & J.A. Bengoechea, (2015) Klebsiella pneumoniae survives within macrophages by avoiding delivery to lysosomes. *Cellular Microbiology* **17**: 1537-1560.

Carlson, J.H. & T.J. Silhavy, (1993) Signal sequence processing is required for the assembly of lamB trimers in the outer membrane of Escherichia coli. *J. Bacteriol.* **175**: 3327-3334.

Carpenter, E.P., K. Beis, A.D. Cameron & S. Iwata, (2008) Overcoming the challenges of membrane protein crystallography. *Current Opinion in Structural Biology* **18**: 581-586.

Carrier, T.A. & J.D. Keasling, (1997) Controlling messenger RNA stability in bacteria: Strategies for engineering gene expression. *Biotechnology Progress* **13**: 699-708.

Carsenti-Etesse, H., P.M. Roger, B. Dunais, S. Durgeat, G. Mancini, M. Bensoussan & P. Dellamonica, (1999) Gradient plate method to induce Streptococcus pyogenes resistance. *Journal of Antimicrobial Chemotherapy* **44**: 439-443.

Casasanta, M.A., C.C. Yoo, H.B. Smith, A.J. Duncan, K. Cochrane, A.C. Varano, E. Allen-Vercoe & D.J. Slade, (2017) A chemical and biological toolbox for Type Vd secretion: Characterization of the phospholipase A1 autotransporter FplA from Fusobacterium nucleatum. *J. Biol. Chem.* **292**: 20240-20254.

Chaba, R., B.M. Alba, M.S. Guo, J. Sohn, N. Ahuja, R.T. Sauer & C.A. Gross, (2011) Signal integration by DegS and RseB governs the σ E -mediated envelope stress response in Escherichia coli. *Proceedings of the National Academy of Sciences* **108**: 2106-2111.

Chaba, R., I.L. Grigorova, J.M. Flynn, T.a. Baker & C.a. Gross, (2007) Design principles of the proteolytic cascade governing the sigmaE-mediated envelope stress response in Escherichia coli: keys to graded, buffered, and rapid signal transduction. *Genes & development* **21**: 124-136.

Chandran, V., R. Fronzes, S. Duquerroy, N. Cronin, J. Navaza & G. Waksman, (2009) Structure of the outer membrane complex of a type IV secretion system. *Nature* **462**: 1011-1015.

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

Cheema, A.K., N.R. Choudhury & H.K. Das, (1999) A- and T-tract-mediated intrinsic curvature in native DNA between the binding site of the upstream activator NtrC and the nifLA promoter of Klebsiella pneumoniae facilitates transcription. *J. Bacteriol.* **181**: 5296-5302.

Cherepanov, P.P. & W. Wackernagel, (1995) Gene disruption in Escherichia coli: TcR and KmR cassettes with the option of Flp-catalyzed excision of the antibiotic-resistance determinant. *Gene* **158**: 9-14.

Chimalakonda, G., N. Ruiz, S.-S. Chng, R.A. Garner, D. Kahne & T.J. Silhavy, (2011) Lipoprotein LptE is required for the assembly of LptD by the -barrel assembly machine in the outer membrane of Escherichia coli. *Proceedings of the National Academy of Sciences* **108**: 2492-2497.

Chirwa, N.T. & M.B. Herrington, (2003) CsgD, a regulator of curli and cellulose synthesis, also regulates serine hydroxymethyltransferase synthesis in Escherichia coli K-12. *Microbiology* **149**: 525-535.

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

Clantin, B., H. Hodak, E. Willery, C. Locht, F. Jacob-Dubuisson & V. Villeret, (2004) The crystal structure of filamentous hemagglutinin secretion domain and its implications for the two-partner secretion pathway. *Proc. Natl. Acad. Sci. U. S. A.* **101**: 6194-6199.

Claret, L. & J. Rouviere-Yaniv, (1997) Variation in HU composition during growth of Escherichia coli: The heterodimer is required for long term survival. *J. Mol. Biol.* **273**: 93-104.

Cohen, E.J., J.L. Ferreira, M.S. Ladinsky, M. Beeby & K.T. Hughes, (2017) Nanoscale-length control of the flagellar driveshaft requires hitting the tethered outer membrane. *Science* **356**: 197-200.

Collin, S., I. Guilvout, M. Chami & A.P. Pugsley, (2007) YaeT-independent multimerization and outer membrane association of secretin PulD. *Mol. Microbiol.* **64**: 1350-1357.

Comandatore, F., D. Sassera, S.C. Bayliss, E. Scaltriti, S. Gaiarsa, X. Cao, A. Gales, R. Saito, S. Pongolini, S. Brisse, E. Feil & C. Bandi, (2018) Gene composition as a potential barrier to large recombinations in the bacterial pathogen Klebsiella pneumoniae. *bioRxiv*: 255349.

Cooper, G.M. & C.D. Brown, (2008) Qualifying the relationship between sequence conservation and molecular function. *Genome research* **18**: 201-205.

Cope, L.D., S.E. Thomas, J.L. Latimer, C.A. Slaughter, U. Müller-Eberhard & E.J. Hansen, (1994) The 100 kDa haem:haemopexin-binding protein of Haemophilus Influenzae: structure and localization. *Mol. Microbiol.* **13**: 863-873.

Costanzo, A., H. Nicoloff, S.E. Barchinger, A.B. Banta, R.L. Gourse & S.E. Ades, (2008) ppGpp and DksA likely regulate the activity of the extracytoplasmic stress factor sigmaE in Escherichia coli by both direct and indirect mechanisms. *Mol. Microbiol.* **67**: 619-632.

Costello, S.M., A.M. Plummer, P.J. Fleming & K.G. Fleming, (2016) Dynamic periplasmic chaperone reservoir facilitates biogenesis of outer membrane proteins. *Proceedings of the National Academy of Sciences* **113**: E4794-E4800.

Costerton, J.W., J.M. Ingram & K.J. Cheng, (1974) Structure and function of the cell envelope of gram-negative bacteria. *Bacteriol Rev* **38**: 87-110.

Cotter, P.A., M.H. Yuk, S. Mattoo, B.J. Akerley, J. Boschwitz, D.A. Relman & J.F. Miller, (1998) Filamentous hemagglutinin of Bordetella bronchiseptica is required for efficient establishment of tracheal colonization. *Infect. Immun.* **66**: 5921-5929.

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

Cress, B.F., J.A. Englaender, W. He, D. Kasper, R.J. Linhardt & M.A.G. Koffas, (2014) Masquerading microbial pathogens: Capsular polysaccharides mimic host-tissue molecules. *FEMS Microbiology Reviews* **38**: 660-697.

Crisafuli, F.A.P., E.B. Ramos & M.S. Rocha, (2014) Characterizing the interaction between DNA and GelRed fluorescent stain. *European Biophysics Journal* **44**: 1-7.

Dartigalongue, C., D. Missiakas & S. Raina, (2001) Characterization of the Escherichia coli sigma E regulon. *The Journal of biological chemistry* **276**: 20866-20875.

Datsenko, K.A. & B.L. Wanner, (2000) One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products. *Proc Natl Acad Sci U S A* **97**: 6640-6645.

Davis, G.S. & L.B. Price, (2016) Recent Research Examining Links Among Klebsiella pneumoniae from Food, Food Animals, and Human Extraintestinal Infections. *Current environmental health reports* **3**: 128-135.

De Las Peñas, A., L. Connolly & C.A. Gross, (1997) SigmaE is an essential sigma factor in Escherichia coli. *J. Bacteriol.* **179**: 6862-6864.

De Majumdar, S., J. Yu, M. Fookes, S.P. McAteer, E. Llobet, S. Finn, S. Spence, A. Monahan, A. Kissenpfennig, R.J. Ingram, J. Bengoechea, D.L. Gally, S. Fanning, J.S. Elborn & T. Schneiders, (2015) Elucidation of the RamA Regulon in Klebsiella pneumoniae Reveals a Role in LPS Regulation. *PLoS Pathog.* **11**: 1-22.

Decad, G.M. & H. Nikaido, (1976) Outer membrane of gram-negative bacteria. XII. Molecular-sieving function of cell wall. *J Bacteriol* **128**: 325-336.

Delattre, A.S., B. Clantin, N. Saint, C. Locht, V. Villeret & F. Jacob-Dubuisson, (2010) Functional importance of a conserved sequence motif in FhaC, a prototypic member of the TpsB/Omp85 superfamily. *FEBS Journal* **277**: 4755-4765.

Delattre, A.S., N. Saint, B. Clantin, E. Willery, G. Lippens, C. Locht, V. Villeret & F. Jacob-Dubuisson, (2011) Substrate recognition by the POTRA domains of TpsB transporter FhaC. *Mol. Microbiol.* **81**: 99-112.

Delcour, A.H., (2009) Outer membrane permeability and antibiotic resistance. *Biochimica et Biophysica Acta - Proteins and Proteomics* **1794**: 808-816.

Demchick, P. & a.L. Koch, (1996) The permeability of the wall fabric of Escherichia coli and Bacillus subtilis . The Permeability of the Wall Fabric of Escherichia coli and Bacillus subtilis. **178**: 768-773.

Deol, S.S., P.J. Bond, C. Domene & M.S.P. Sansom, (2004) Lipid-protein interactions of integral membrane proteins: A comparative simulation study. *Biophysical Journal* **87**: 3737-3749.

Dhingra, K.R., (2008) A Case of Complicated Urinary Tract Infection: Klebsiella pneumoniae Emphysematous Cystitis Presenting as Abdominal Pain in the Emergency Department. *The western journal of emergency medicine* **9**: 171-173.

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

Diwa, A., A.L. Bricker, C. Jain & J.G. Belasco, (2000) that directs feedback regulation of RNase E gene expression An evolutionarily conserved RNA stem – loop functions as a sensor that directs feedback regulation of RNase E gene expression. *Genes & Development*: 1249-1260.

Doerner, P.A. & M.C. Sousa, (2015) Small Angle X-ray Scattering (SAXS) Characterization of the POTRA Domains of BamA. *Methods in molecular biology (Clifton, N.J.)* **1329**: 149-155.

Doerner, P.A. & M.C. Sousa, (2017) Extreme Dynamics in the BamA β-Barrel Seam. *Biochemistry* **56**: 3142-3149.

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

Dong, C., H.F. Hou, X. Yang, Y.Q. Shen & Y.H. Dong, (2012) Structure of Escherichia coli BamD and its functional implications in outer membrane protein assembly. *Acta Crystallographica Section D: Biological Crystallography*.

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

Doorduijn, D.J., S.H.M. Rooijakkers, W. van Schaik & B.W. Bardoel, (2016) Complement resistance mechanisms of Klebsiella pneumoniae. In: Immunobiology. pp.

Dramsi, S., S. Magnet, S. Davison & M. Arthur, (2008) Covalent attachment of proteins to peptidoglycan. *FEMS Microbiology Reviews* **32**: 307-320.

Driessen, A.J.M. & N. Nouwen, (2008) Protein Translocation Across the Bacterial Cytoplasmic Membrane. *Annual Review of Biochemistry* **77**: 643-667.

Du, D., H.W. van Veen, S. Murakami, K.M. Pos & B.F. Luisi, (2015) Structure, mechanism and cooperation of bacterial multidrug transporters. *Current Opinion in Structural Biology* **33**: 76-91.

Dumon-Seignovert, L., G. Cariot & L. Vuillard, (2004) The toxicity of recombinant proteins in Escherichia coli: A comparison of overexpression in BL21(DE3), C41(DE3), and C43(DE3). *Protein Expression and Purification* **37**: 203-206.

Dunstan, R.A., I.D. Hay & T. Lithgow, (2017) Defining Membrane Protein Localization by Isopycnic Density Gradients. In: Bacterial Protein Secretion Systems: Methods and Protocols. L. Journet & E. Cascales (eds). New York, NY: Springer New York, pp. 81-86.

Dutta, R. & M. Inouye, (2000) GHKL, an emergent ATPase/kinase superfamily. *Trends in Biochemical Sciences* **25**: 24-28.

Dykhuizen, D.E. & A.M. Dean, (1990) Enzyme activity and fitness: Evolution in solution. *Trends in Ecology and Evolution* **5**: 257-262.

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

Elsen, S., P. Huber, S. Bouillot, Y. Couté, P. Fournier, Y. Dubois, J.F. Timsit, M. Maurin & I. Attrée, (2014) A type III secretion negative clinical strain of Pseudomonas aeruginosa employs a two-partner secreted exolysin to induce hemorrhagic pneumonia. *Cell Host and Microbe* **15**: 164-176.

Emsley, P., I.G. Charles, N.F. Fairweather & N.W. Isaacs, (1996) Structure of Bordetella pertussis virulence factor P.69 pertactin. In: Nature. pp. 90-92.

Endo, T., S. Kawano, K. Yamano, B. Clantin, T. Endo, K. Yamano, P. Gatzeva-Topalova, C. Hagan, S. Kim, T. Knowles, T. Knowles, G. Patel, D. Stroud & D. Walther, (2011) BamE structure: the assembly of β-barrel proteins in the outer membranes of bacteria and mitochondria. *EMBO reports* **12**: 94-95.

Erhardt, H., F. Dempwolff, M. Pfreundschuh, M. Riehle, C. Schäfer, T. Pohl, P. Graumann & T. Friedrich, (2014) Organization of the Escherichia coli aerobic enzyme complexes of oxidative phosphorylation in dynamic domains within the cytoplasmic membrane. *MicrobiologyOpen* **3**: 316-326.

Espinosa-Cantú, A., D. Ascencio, F. Barona-Gómez & A. De Luna, (2015) Gene duplication and the evolution of moonlighting proteins. *Frontiers in Genetics* **6**: 1-7.

Facey, S.J. & A. Kuhn, (2004) Membrane integration of E. coli model membrane proteins. *Biochimica et Biophysica Acta - Molecular Cell Research* **1694**: 55-66.

Fairman, J.W., N. Noinaj & S.K. Buchanan, (2011) The structural biology of β-barrel membrane proteins: a summary of recent reports. *Current opinion in structural biology* **21**: 523-531.

Fan, E., N. Chauhan, D.B.R.K.G. Udatha, J.C. Leo & D. Linke, (2016) Type V Secretion Systems in Bacteria. *Virulence Mechanisms of Bacterial Pathogens, Fifth Edition*: 305-335.

Fleming, K.G., (2015) A combined kinetic push and thermodynamic pull as driving forces for outer membrane protein sorting and folding in bacteria. *Philosophical Transactions of the Royal Society B: Biological Sciences* **370**.

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

Flores-Kim, J. & A.J. Darwin, (2014) Regulation of bacterial virulence gene expression by cell envelope stress responses. *Virulence* **5**: 835-851.

Garcia, E.C., A.R. Brumbaugh & H.L.T. Mobley, (2011) Redundancy and specificity of Escherichia coli iron acquisition systems during urinary tract infection. *Infect. Immun.* **79**: 1225-1235.

Gasteiger, E., A. Gattiker, C. Hoogland, I. Ivanyi, R.D. Appel & A. Bairoch, (2003) ExPASy: The proteomics server for in-depth protein knowledge and analysis. *Nucleic Acids Research* **31**: 3784-3788.

Gelis, I., A.M.J.J. Bonvin, D. Keramisanou, M. Koukaki, G. Gouridis, S. Karamanou, A. Economou & C.G. Kalodimos, (2007) Structural Basis for Signal-Sequence Recognition by the Translocase Motor SecA as Determined by NMR. *Cell* **131**: 756-769.

Gentle, I.E., L. Burri & T. Lithgow, (2005) Molecular architecture and function of the Omp85 family of proteins. *Mol. Microbiol.* **58**: 1216-1225.

George, A.M., R.M. Hall & H.W. Stokes, (1995) Multidrug resistance in Klebsiella pneumoniae: a novel gene, ramA, confers a multidrug resistance phenotype in Escherichia coli. *Microbiology* **141**: 1909-1920.

Gessmann, D., Y.H. Chung, E.J. Danoff, A.M. Plummer, C.W. Sandlin, N.R. Zaccai & K.G. Fleming, (2014) Outer membrane β-barrel protein folding is physically controlled by periplasmic lipid head groups and BamA. *Proc. Natl. Acad. Sci. U. S. A.* **111**: 5878-5883.

Ghequire, M.G.K., T. Swings, J. Michiels, S.K. Buchanan & R. De Mot, (2018) Hitting with a BAM: Selective Killing by Lectin-Like Bacteriocins. *mBio* **9**: e02138-02117.

Gil, R. & A. Latorre, (2012) Factors behind junk DNA in bacteria. *Genes* **3**: 634-650.

Girard, V., J.P. Côté, M.È. Charbonneau, M. Campos, F. Berthiaume, M.A. Hancock, N. Siddiqui & M. Mourez, (2010) Conformation change in a self-recognizing autotransporter modulates bacterial cell-cell interaction. *J. Biol. Chem.* **285**: 10616-10626.

Goemans, C., K. Denoncin & J.F. Collet, (2014) Folding mechanisms of periplasmic proteins. *Biochimica et Biophysica Acta - Molecular Cell Research* **1843**: 1517-1528.

Goodall, E.C.A., A. Robinson, I.G. Johnston, S. Jabbari, K.A. Turner, A.F. Cunningham, P.A. Lund, J.A. Cole, I.R. Henderson & K.A. Kline, (2018) The Essential Genome of Escherichia coli K-12. *mBio* **9**: e02096-02017.

Grabowicz, M. & T.J. Silhavy, (2017a) Envelope Stress Responses: An Interconnected Safety Net. In: Trends in Biochemical Sciences. pp.

Grabowicz, M. & T.J. Silhavy, (2017b) Redefining the essential trafficking pathway for outer membrane lipoproteins. *Proceedings of the National Academy of Sciences* **114**: 4769-4774.

Grant, G.A., (2003) Synthetic Peptides for Production of Antibodies that Recognize Intact Proteins. *Current Protocols in Immunology* **Chapter 9**: Unit 9.2.

Gruss, F., F. Zähringer, R.P. Jakob, B.M. Burmann, S. Hiller & T. Maier, (2013) The structural basis of autotransporter translocation by TamA. *Nature Structural & Molecular Biology* **20**: 1318-1320.

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

Guédin, S., E. Willery, J. Tommassen, E. Fort, H. Drobecq, C. Locht & F. Jacob-Dubuisson, (2000) Novel topological features of FhaC, the outer membrane transporter involved in the secretion of the Bordetella pertussis filamentous hemagglutinin. *J. Biol. Chem.* **275**: 30202-30210.

Guérin, J., S. Bigot, R. Schneider, S.K. Buchanan & F. Jacob-Dubuisson, (2017) Two-Partner Secretion: Combining Efficiency and Simplicity in the Secretion of Large Proteins for Bacteria-Host and Bacteria-Bacteria Interactions. *Frontiers in Cellular and Infection Microbiology* **7**: 1-23.

Guilhen, C., N. Charbonnel, N. Parisot, N. Gueguen, A. Iltis, C. Forestier & D. Balestrino, (2016) Transcriptional profiling of Klebsiella pneumoniae defines signatures for planktonic, sessile and biofilm-dispersed cells. *BMC Genomics* **17**.

Guimaraes, J.C., M. Rocha & A.P. Arkin, (2014) Transcript level and sequence determinants of protein abundance and noise in Escherichia coli. *Nucleic Acids Research* **42**: 4791-4799.

Gunasinghe, S.D., T. Shiota, C.J. Stubenrauch, K.E. Schulze, C.T. Webb, A.J. Fulcher, R.A. Dunstan, I.D. Hay, T. Naderer, D.R. Whelan, T.D.M. Bell, K.D. Elgass, R.A. Strugnell & T. Lithgow, (2018) The WD40 Protein BamB Mediates Coupling of BAM Complexes into Assembly Precincts in the Bacterial Outer Membrane. *Cell Reports* **23**.

Gupta, S.D. & H.C. Wu, (1991) Identification and subcellular localization of apolipoprotein N-acyltransferase in Escherichia coli. *FEMS Microbiology Letters* **62**: 37-41.

Guzman, L.-M., D. Belin, M.J. Carson & J. Beckwith, (1995) Tight Regulation, Modulation, and High-Level Expression by Vectors Containing the Arabinose P BAD Promoter. *J. Bacteriol.* **177**: 4121-4130.

Haarmann, R., M. Ibrahim, M. Stevanovic, R. Bredemeier & E. Schleiff, (2010) The properties of the outer membrane localized Lipid A transporter LptD. *Journal of Physics Condensed Matter* **22**.

Hagan, C.L., S. Kim & D. Kahne, (2010) Reconstitution of Outer Membrane Protein Assembly from Purified Components. *Science* **328**: 890-892.

Hagan, C.L., T.J. Silhavy & D. Kahne, (2011) β-barrel membrane protein assembly by the bam complex. In: Annual Review of Biochemistry. pp. 189-210.

Hagan, C.L., D.B. Westwood & D. Kahne, (2013) bam Lipoproteins Assemble BamA in vitro. *Biochemistry* **52**: 6108-6113.

Hagan, C.L., J.S. Wzorek & D. Kahne, (2015) Inhibition of the β-barrel assembly machine by a peptide that binds BamD. *Proceedings of the National Academy of Sciences* **112**: 2011-2016.

Hall, B.G., H. Acar, A. Nandipati & M. Barlow, (2014) Growth rates made easy. *Molecular Biology and Evolution* **31**: 232-238.

Han, L., J. Zheng, Y. Wang, X. Yang, Y. Liu, C. Sun, B. Cao, H. Zhou, D. Ni, J. Lou, Y. Zhao & Y. Huang, (2016) Structure of the BAM complex and its implications for biogenesis of outer-membrane proteins. *Nature Structural & Molecular Biology* **23**: 192-196.

Hashimoto, K. & A.R. Panchenko, (2010) Mechanisms of protein oligomerization, the critical role of insertions and deletions in maintaining different oligomeric states. *Proceedings of the National Academy of Sciences* **107**: 20352-20357.

Hasselblatt, H., R. Kurzbauer, C. Wilken, T. Krojer, J. Sawa, J. Kurt, R. Kirk, S. Hasenbein, M. Ehrmann & T. Clausen, (2007) Regulation of the sigmaE stress response by DegS: how the PDZ domain keeps the protease inactive in the resting state and allows integration of different OMP-derived stress signals upon folding stress. *Genes & development* **21**: 2659-2670.

Hay, I.D., M.J. Belousoff & T. Lithgow, (2017) Structural basis of type 2 secretion system engagement between the inner and outer bacterial membranes. *mBio* **8**: 1-6.

He, T., Y. Wang, L. Sun, M. Pang, L. Zhang & R. Wang, (2017) Occurrence and characterization of blaNDM-5-positive Klebsiella pneumoniae isolates from dairy cows in Jiangsu, China. *Journal of Antimicrobial Chemotherapy* **72**: 90-94.

Heinz, E. & T. Lithgow, (2014) A comprehensive analysis of the Omp85/TpsB protein superfamily structural diversity, taxonomic occurrence, and evolution. *Front. Microbiol.* **5**: 1-13.

Heinz, E., J. Selkrig, M.J. Belousoff & T. Lithgow, (2015) Evolution of the translocation and assembly module (TAM). *Genome Biology and Evolution* **7**.

Heinz, E., C.J. Stubenrauch, R. Grinter, N.P. Croft, A.W. Purcell, R.A. Strugnell, G. Dougan & T. Lithgow, (2016) Conserved Features in the Structure, Mechanism, and Biogenesis of the Inverse Autotransporter Protein Family. *Genome biology and evolution* **8**.

Hellman, L.M. & M.G. Fried, (2007) Electrophoretic mobility shift assay (EMSA) for detecting protein-nucleic acid interactions. *Nature Protocols* **2**: 1849-1861.

Henderson, I.R. & J.P. Nataro, (2001) Virulence functions of autotransporter proteins. *Infect. Immun.* **69**: 1231-1243.

Henderson, I.R., F. Navarro-Garcia, M. Desvaux, R.C. Fernandez & D. Ala'Aldeen, (2004) Type V protein secretion pathway: the autotransporter story. *Microbiology and molecular biology reviews : MMBR* **68**: 692-744.

Hendler, R.W. & A.H. Burgess, (1974) Fractionation of the electron-transport chain of Escherichia coli. *Biochim Biophys Acta* **357**: 215-230.

Henkin, T.M. & C. Yanofsky, (2002) Regulation by transcription attenuation in bacteria: How RNA provides instructions for transcription termination/antitermination decisions. *BioEssays* **24**: 700-707.

Hernandez-Alles, S., S. Alberti, D. Alvarez, A. Domenech-Sanchez, L. Martinez-Martinez, J. Gil, J.M. Tomas & V.J. Benedi, (1999) Porin expression in clinical isolates of Klebsiella pneumoniae. *Microbiology* **145**: 673-679.

Herring, C.D., J.D. Glasner & F.R. Blattner, (2003) Gene replacement without selection: regulated suppression of amber mutations in Escherichia coli. *Gene* **311**: 153-163.

Hertle, R., (2005) The family of Serratia type pore forming toxins. *Current protein & peptide science* **6**: 313-325.

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

Hoang, H.H., N.N. Nickerson, V.T. Lee, A. Kazimirova, M. Chami, A.P. Pugsley & S. Lory, (2011) Outer membrane targeting of Pseudomonas aeruginosa proteins shows variable dependence on the components of Bam and Lol machineries. *mBio* **2**: 1-8.

Höhr, A.I.C., C. Lindau, C. Wirth, J. Qiu, D.A. Stroud, S. Kutik, B. Guiard, C. Hunte, T. Becker, N. Pfanner & N. Wiedemann, (2018) Membrane protein insertion through a mitochondrial β-barrel gate. *Science* **359**.

Holt, K.E., H. Wertheim, R.N. Zadoks, S. Baker, C.A. Whitehouse, D. Dance, A. Jenney, T.R. Connor, L.Y. Hsu, J. Severin, S. Brisse, H. Cao, J. Wilksch, C. Gorrie, M.B. Schultz, D.J. Edwards, K.V. Nguyen, T.V. Nguyen, T.T. Dao, M. Mensink, V.L. Minh, N.T. Nhu, C. Schultsz, K. Kuntaman, P.N. Newton, C.E. Moore, R.A. Strugnell & N.R. Thomson, (2015a) Genomic analysis of diversity, population structure, virulence, and antimicrobial resistance in Klebsiella pneumoniae, an urgent threat to public health. *Proc Natl Acad Sci U S A* **112**: E3574-3581.

Holt, K.E., H. Wertheim, R.N. Zadoks, S. Baker, C.A. Whitehouse, D. Dance, A. Jenney, T.R. Connor, L.Y. Hsu, J. Severin, S. Brisse, H. Cao, J. Wilksch, C. Gorrie, M.B. Schultz, D.J. Edwards, K.V. Nguyen, T.V. Nguyen, T.T. Dao, M. Mensink, V.L. Minh, N.T.K. Nhu, C. Schultsz, K. Kuntaman, P.N. Newton, C.E. Moore, R.A. Strugnell & N.R. Thomson, (2015b) Genomic analysis of diversity, population structure, virulence, and antimicrobial resistance in Klebsiella pneumoniae , an urgent threat to public health. *Proceedings of the National Academy of Sciences* **112**.

Hong, H., S. Park, R.H.F. Jiménez, D. Rinehart & L.K. Tamm, (2007) Role of aromatic side chains in the folding and thermodynamic stability of integral membrane proteins. *Journal of the American Chemical Society* **129**: 8320-8327.

Hsieh, P.F., C.R. Hsu, C.T. Chen, T.L. Lin & J.T. Wang, (2016) The Klebsiella pneumoniae YfgL (BamB) lipoprotein contributes to outer membrane protein biogenesis, type-1 fimbriae expression, anti-phagocytosis, and in vivo virulence. *Virulence* **7**: 587-601.

Hsueh, Y.-C., K. Nicolaisen, L.E. Gross, J. Nöthen, N. Schauer, L. Vojta, F. Ertel, I. Koch, R. Ladig, H. Fulgosi, A.R. Fernie & E. Schleiff, (2018) The outer membrane Omp85-like protein P39 influences metabolic homeostasis in mature Arabidopsis thaliana. *Plant Biology*: 1-9.

Huang, C., P. Rossi, T. Saio & C.G. Kalodimos, (2016) Structural basis for the antifolding activity of a molecular chaperone. *Nature* **537**: 202-206.

Huerta, A.M. & J. Collado-Vides, (2003) Sigma70 promoters in Escherichia coli: Specific transcription in dense regions of overlapping promoter-like signals. *J. Mol. Biol.* **333**: 261-278.

Hufton, S.E., R.J. Ward, N.A. Bunce, J.T. Armstrong, A.J. Fletcher & R.E. Glass, (1995) Structure-function analysis of the vitamin B12 receptor of Escherichia coli by means of informational suppression. *Mol. Microbiol.* **15**: 381-393.

Hui, C.Y., Y. Guo, Q.S. He, L. Peng, S.C. Wu, H. Cao & S.H. Huang, (2010) Escherichia coli outer membrane protease OmpT confers resistance to urinary cationic peptides. *Microbiology and Immunology* **54**: 452-459.

Hussain, S. & H.D. Bernstein, (2018) The Bam complex catalyzes efficient insertion of bacterial outer membrane proteins into membrane vesicles of variable lipid composition. *J. Biol. Chem.* **293**: 2959-2973.

Iadanza, M.G., A.J. Higgins, B. Schiffrin, A.N. Calabrese, D.J. Brockwell, A.E. Ashcroft, S.E. Radford & N.A. Ranson, (2016) Lateral opening in the intact β-barrel assembly machinery captured by cryo-EM. *Nature Communications* **7**: 12865.

Ieva, R., (2017) Interfering with outer membrane biogenesis to fight Gram-negative bacterial pathogens. *Virulence* **8**: 1049-1052.

Ieva, R., P. Tian, J.H. Peterson & H.D. Bernstein, (2011) Sequential and spatially restricted interactions of assembly factors with an autotransporter beta domain. *Proc. Natl. Acad. Sci. U. S. A.* **108**: E383-391.

Ingledew, W.J. & R.K. Poole, (1984) The respiratory chains of Escherichia coli. *Microbiol Rev* **48**: 222-271.

Ishihama, A., (2010) Prokaryotic genome regulation: Multifactor promoters, multitarget regulators and hierarchic networks. *FEMS Microbiology Reviews* **34**: 628-645.

Ize, B., N.R. Stanley, G. Buchanan & T. Palmer, (2003) Role of the Escherichia coli Tat pathway in outer membrane integrity. *Mol. Microbiol.* **48**: 1183-1193.

J Simpson, R., (2006) *SDS-PAGE of proteins*.

Jacob-Dubuisson, F., R. Fernandez & L. Coutte, (2004) Protein secretion through autotransporter and two-partner pathways. *Biochimica et Biophysica Acta - Molecular Cell Research* **1694**: 235-257.

Jacob-Dubuisson, F., J. Guérin, S. Baelen & B. Clantin, (2013) Two-partner secretion: as simple as it sounds? *Research in Microbiology* **164**: 583-595.

Jacob-Dubuisson, F., V. Villeret, B. Clantin, A.S. Delattre & N. Saint, (2009) First structural insights into the TpsB/Omp85 superfamily. In: Biological Chemistry. pp.

Jagodnik, J., C. Chiaruttini & M. Guillier, (2017) Stem-Loop Structures within mRNA Coding Sequences Activate Translation Initiation and Mediate Control by Small Regulatory RNAs. *Molecular Cell* **68**: 158-170.e153.

Jansen, K.B., S.L. Baker & M.C. Sousa, (2015) Crystal structure of BamB bound to a periplasmic domain fragment of BamA, the central component of the β-barrel assembly machine. *J. Biol. Chem.* **290**.

Jarchow, S., C. Lück, A. Görg & A. Skerra, (2008) Identification of potential substrate proteins for the periplasmic Escherichia coli chaperone Skp. *Proteomics* **8**: 4987-4994.

Jeeves, M. & T.J. Knowles, (2015) A novel pathway for outer membrane protein biogenesis in Gram-negative bacteria. *Mol. Microbiol.* **97**.

Jeffery, C.J., (2016) Expression, solubilization, and purification of bacterial membrane proteins. *Current Protocols in Protein Science* **2016**: 29.15.21-29.15.15.

Jiang, D., H.W. Jarrett & W.E. Haskins, (2009) Methods for proteomic analysis of transcription factors. *Journal of Chromatography A* **1216**: 6881-6889.

Jiang, J., X. Zhang, Y. Chen, Y. Wu, Z.H. Zhou, Z. Chang & S.-F. Sui, (2008) Activation of DegP chaperone-protease via formation of large cage-like oligomers upon binding to substrate proteins. *Proceedings of the National Academy of Sciences* **105**: 11939-11944.

Julio, S.M. & P.A. Cotter, (2005) Characterization of the filamentous hemagglutinin-like protein FhaS in Bordetella bronchiseptica. *Infect. Immun.* **73**: 4960-4971.

Juncker, A. & H. Willenbrock, (2003) Prediction of lipoprotein signal peptides in Gram negative bacteria. *Protein science* **12**: 1652-1662.

Kajava, A.V. & A.C. Steven, (2006a) The turn of the screw: Variations of the abundant β-solenoid motif in passenger domains of Type V secretory proteins. *Journal of Structural Biology* **155**: 306-315.

Kajava, A.V. & A.C. Steven, (2006b) β-Rolls, β-Helices, and Other β-Solenoid Proteins. *Advances in Protein Chemistry* **73**: 55-96.

Katayama, Y., T. Ito & K. Hiramatsu, (2000) A new class of genetic element, staphylococcus cassette chromosome mec, encodes methicillin resistance in Staphylococcus aureus. *Antimicrob. Agents Chemother.* **44**: 1549-1555.

Kawakami, T., K. Kawamura, K. Fujimori, A. Koike & F. Amano, (2016) Influence of the culture medium on the production of nitric oxide and expression of inducible nitric oxide synthase by activated macrophages in vitro. *Biochemistry and Biophysics Reports* **5**: 328-334.

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

Kelley, L.A. & M.J.E. Sternberg, (2009) Protein structure prediction on the web: A case study using the phyre server. *Nature Protocols* **4**: 363-373.

Kendall, D.A., S.C. Bock & E.T. Kaiser, (1986) Idealization of the hydrophobic segment of the alkaline phosphatase signal peptide. *Nature* **321**: 706-708.

Khan, A., W.R. Miller & C.A. Arias, (2018) Mechanisms of antimicrobial resistance among hospital-associated pathogens. *Expert Review of Anti-Infective Therapy* **16**: 269-287.

Kim, D., J.S.-J. Hong, Y. Qiu, H. Nagarajan, J.-H. Seo, B.-K. Cho, S.-F. Tsai & B.Ø. Palsson, (2012) Comparative Analysis of Regulatory Elements between Escherichia coli and Klebsiella pneumoniae by Genome-Wide Transcription Start Site Profiling. *PLoS Genetics* **8**: e1002867.

Kim, K.H., S. Aulakh & M. Paetzel, (2011) Crystal structure of β-barrel assembly machinery BamCD protein complex. *The Journal of biological chemistry* **286**: 39116-39121.

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

Kleinschmidt, J.H., (2003) Membrane protein folding on the example of outer membrane protein A of Escherichia coli. *Cell. Mol. Life Sci.* **60**: 1547-1558.

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

Kleinschmidt, J.H., (2015) Folding of β-barrel membrane proteins in lipid bilayers - Unassisted and assisted folding and insertion. In: Biochim. Biophys. Acta Biomembr., pp.

Klemm, P., L. Hjerrild, M. Gjermansen & M.A. Schembri, (2004) Structure-function analysis of the self-recognizing Antigen 43 autotransporter protein from Escherichia coli. *Mol. Microbiol.* **51**: 283-296.

Knowles, T.J., M. Jeeves, S. Bobat, F. Dancea, D. McClelland, T. Palmer, M. Overduin & I.R. Henderson, (2008) Fold and function of polypeptide transport-associated domains responsible for delivering unfolded proteins to membranes. *Mol. Microbiol.* **68**: 1216-1227.

Knowles, T.J., A. Scott-Tucker, M. Overduin & I.R. Henderson, (2009) Membrane protein architects: the role of the BAM complex in outer membrane protein assembly. *Nat. Rev. Microbiol.* **7**: 206-214.

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

Koenig, T., B.H. Menze, M. Kirchner, F. Monigatti, K.C. Parker, T. Patterson, J.J. Steen, F.A. Hamprecht & H. Steen, (2008) Robust Prediction of the MASCOT Score for an Improved Quality Assessment in Mass Spectrometric Proteomics Robust Prediction of the MASCOT Score for an Improved Quality Assessment in Mass Spectrometric Proteomics. *Journal of Proteome Research*: 3708-3717.

Könninger, U.W., S. Hobbie, R. Benz & V. Braun, (1999) The haemolysin-secreting ShlB protein of the outer membrane of Serratia marcescens: Determination of surface-exposed residues and formation of ion-permeable pores by ShlB mutants in artificial lipid bilayer membranes. *Mol. Microbiol.* **32**: 1212-1225.

Konovalova, A., D.E. Kahne & T.J. Silhavy, (2017) Outer Membrane Biogenesis. *Annual Review of Microbiology* **71**.

Konovalova, A., A.M. Mitchell & T.J. Silhavy, (2016) A lipoprotein/β-barrel complex monitors lipopolysaccharide integrity transducing information across the outer membrane. *eLife* **5**.

Koronakis, V., J. Eswaran & C. Hughes, (2004) Structure and Function of TolC: The Bacterial Exit Duct for Proteins and Drugs. *Annual Review of Biochemistry* **73**: 467-489.

Kozak, M., (2005) Regulation of translation via mRNA structure in prokaryotes and eukaryotes. *Gene* **361**: 13-37.

Kroger, C., S.C. Dillon, A.D.S. Cameron, K. Papenfort, S.K. Sivasankaran, K. Hokamp, Y. Chao, A. Sittka, M. Hebrard, K. Handler, A. Colgan, P. Leekitcharoenphon, G.C. Langridge, A.J. Lohan, B. Loftus, S. Lucchini, D.W. Ussery, C.J. Dorman, N.R. Thomson, J. Vogel & J.C.D. Hinton, (2012) The transcriptional landscape and small RNAs of Salmonella enterica serovar Typhimurium. *Proceedings of the National Academy of Sciences* **109**: E1277-E1286.

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

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

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

Krylov, A.S., (2001) Massive parallel analysis of the binding specificity of histone-like protein HU to single- and double-stranded DNA with generic oligodeoxyribonucleotide microchips. *Nucleic Acids Research* **29**: 2654-2660.

Kudva, R., K. Denks, P. Kuhn, A. Vogt, M. Müller & H.G. Koch, (2013) Protein translocation across the inner membrane of Gram-negative bacteria: The Sec and Tat dependent protein transport pathways. *Research in Microbiology* **164**: 505-534.

Kuehn, M.J. & N.C. Kesty, (2005) Bacterial outer membrane vesicles and the host – pathogen interaction. *Genes & Development*: 2645-2655.

Kwon, Y.M. & S.C. Ricke, (2000) Efficient amplification of multiple transposon-flanking sequences. *Journal of Microbiological Methods* **41**: 195-199.

Lai, Y.-c., H.-l. Peng & H.-y. Chang, (2003) RmpA2, an Activator of Capsule Biosynthesis in. *Mbio* **185**: 788-800.

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

Lee, C.-R., J.H. Lee, K.S. Park, J.H. Jeon, Y.B. Kim, C.-J. Cha, B.C. Jeong & S.H. Lee, (2017) Antimicrobial Resistance of Hypervirulent Klebsiella pneumoniae: Epidemiology, Hypervirulence-Associated Determinants, and Resistance Mechanisms. *Frontiers in Cellular and Infection Microbiology* **7**: 483.

Lee, D.J., L.E. Bingle, K. Heurlier, M.J. Pallen, C.W. Penn, S.J. Busby & J.L. Hobman, (2009) Gene doctoring: a method for recombineering in laboratory and pathogenic Escherichia coli strains. *BMC Microbiol* **9**: 252.

Lee, J., H.A. Sutterlin, J.S. Wzorek, M.D. Mandler, C.L. Hagan, M. Grabowicz, D. Tomasek, M.D. May, E.M. Hart, T.J. Silhavy & D. Kahne, (2018) Substrate binding to BamD triggers a conformational change in BamA to control membrane insertion. *Proceedings of the National Academy of Sciences*: 201711727.

Lee, J., M. Xue, J.S. Wzorek, T. Wu, M. Grabowicz, L.S. Gronenberg, H.A. Sutterlin, R.M. Davis, N. Ruiz, T.J. Silhavy & D.E. Kahne, (2016) Characterization of a stalled complex on the β-barrel assembly machine. *Proceedings of the National Academy of Sciences* **113**: 8717-8722.

Leonard-Rivera, M. & R. Misra, (2012) Conserved residues of the putative L6 loop of escherichia coli BamA play a critical role in the assembly of ??-barrel outer membrane proteins, including that of BamA itself. *J. Bacteriol.* **194**: 4662-4668.

Lerat, E. & H. Ochman, (2005) Recognizing the pseudogenes in bacterial genomes. *Nucleic Acids Research* **33**: 3125-3132.

Létoffé, S., V. Redeker & C. Wandersman, (1998) Isolation and characterization of an extracellular haem-binding protein from Pseudomonas aeruginosa that shares function and sequence similarities with the Serratia marcescens HasA haemophore. *Mol. Microbiol.* **28**: 1223-1234.

Li, J., F. Liu, Q. Wang, P. Ge, P.C.Y. Woo, J. Yan, Y. Zhao, G.F. Gao, C.H. Liu & C. Liu, (2015a) Genomic and transcriptomic analysis of NDM-1 Klebsiella pneumoniae in spaceflight reveal mechanisms underlying environmental adaptability. *Scientific Reports* **4**: 6216.

Li, J., C.C. Overall, R.C. Johnson, M.B. Jones, J.E. McDermott, F. Heffron, J.N. Adkins & E.D. Cambronne, (2015b) ChIP-seq analysis of the σe regulon of salmonella enterica serovar typhimurium reveals new genes implicated in heat shock and oxidative stress response. *PLoS ONE* **10**: 1-15.

Linke, D., T. Riess, I.B. Autenrieth, A. Lupas & V.A.J. Kempf, (2006) Trimeric autotransporter adhesins: variable structure, common function. *Trends Microbiol.* **14**: 264-270.

Liu, Y., A. Beyer & R. Aebersold, (2016) On the Dependency of Cellular Protein Levels on mRNA Abundance. *Cell* **165**: 535-550.

Liu, Y., X. Fu, J. Shen, H. Zhang, W. Hong & Z. Chang, (2004) Periplasmic proteins of Escherichia coli are highly resistant to aggregation: Reappraisal for roles of molecular chaperones in periplasm. *Biochemical and Biophysical Research Communications* **316**: 795-801.

Loenen, W.A.M., D.T.F. Dryden, E.A. Raleigh & G.G. Wilson, (2014) Type I restriction enzymes and their relatives. *Nucleic acids research* **42**: 20-44.

Lucchini, S., G. Rowley, M.D. Goldberg, D. Hurd, M. Harrison & J.C.D. Hinton, (2006) H-NS mediates the silencing of laterally acquired genes in bacteria. *PLoS Pathog.* **2**: 0746-0752.

Luft, J.R., J.R. Wolfley & E.H. Snell, (2011) What's in a drop? Correlating observations and outcomes to guide macromolecular crystallization experiments. *Crystal Growth and Design* **11**: 651-663.

Lycklama a Nijeholt, J.A. & A.J.M. Driessen, (2012) The bacterial Sec-translocase: Structure and mechanism. *Philosophical Transactions of the Royal Society B: Biological Sciences* **367**: 1016-1028.

Lyu, Z.X. & X.S. Zhao, (2015) Periplasmic quality control in biogenesis of outer membrane proteins. *Biochemical Society transactions* **43**.

Mackie, A.M., K.A. Hassan, I.T. Paulsen & S.G. Tetu, (2014) Biolog Phenotype MicroArrays for Phenotypic Characterization of Microbial Cells. In: Environmental Microbiology: Methods and Protocols. I.T. Paulsen & A.J. Holmes (eds). Totowa, NJ: Humana Press, pp. 123-130.

Madan Babu, M., S. Balaji & L. Aravind, (2007) General trends in the evolution of prokaryotic transcriptional regulatory networks. *Genome Dyn* **3**: 66-80.

Madhusudhan, M.S., M.A. Marti-Renom, N. Eswar, B. John, U. Pieper, R. Karchin, M.-Y. Shen & A. Sali, (2005) Comparative Protein Structure Modeling. In: The Proteomics Protocols Handbook. Totowa, NJ: Humana Press, pp. 831-860.

Maier, T., B. Clantin, F. Gruss, F. Dewitte, A.-S. Delattre, F. Jacob-Dubuisson, S. Hiller & V. Villeret, (2015) Conserved Omp85 lid-lock structure and substrate recognition in FhaC. *Nature Communications* **6**: 7452.

Maier, T., M. Güell & L. Serrano, (2009) Correlation of mRNA and protein in complex biological samples. *FEBS Letters* **583**: 3966-3973.

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

Mallik, P., B.J. Paul, S.T. Rutherford, R.L. Gourse & R. Osuna, (2006) DksA is required for growth phase-dependent regulation, growth rate-dependent control, and stringent control of fis expression in Escherichia coli. *J. Bacteriol.* **188**: 5775-5782.

Matias, V.R.F., A. Al-amoudi, J. Dubochet & T.J. Beveridge, (2003) Cryo-Transmission Electron Microscopy of Frozen-Hydrated Sections of Escherichia coli and Pseudomonas aeruginosa Cryo-Transmission Electron Microscopy of Frozen-Hydrated Sections of Escherichia coli and Pseudomonas aeruginosa. *J. Bacteriol.* **185**: 6112-6118.

Matsushita, C., O. Matsushita, S. Katayama, J. Minami, K. Takai & A. Okabe, (1996) An upstream activating sequence containing curved DNA involved in activation of the Clostridium perfringens plc promoter. *Microbiology (Reading, England)* **142 (Pt 9**: 2561-2566.

Mazar, J. & P.A. Cotter, (2006) Topology and maturation of filamentous haemagglutinin suggest a new model for two-partner secretion. *Mol. Microbiol.* **62**: 641-654.

Mazurkiewicz, P., C.M. Tang, C. Boone & D.W. Holden, (2006) Signature-tagged mutagenesis: Barcoding mutants for genome-wide screens. *Nature Reviews Genetics* **7**: 929-939.

McCabe, A.L., D. Ricci, M. Adetunji & T.J. Silhavy, (2017) Conformational Changes That Coordinate the Activity of BamA and BamD Allowing β-Barrel Assembly. *J. Bacteriol.* **199**: 1-11.

McCarter, J.D., D. Stephens, K. Shoemaker, S. Rosenberg, J.F. Kirsch & G. Georgiou, (2004) Substrate specificity of the Escherichia coli outer membrane protease OmpT. *J. Bacteriol.* **186**: 5919-5925.

McGuffin, L.J., K. Bryson & D.T. Jones, (2000) The PSIPRED protein structure prediction server. *Bioinformatics* **16**: 404-405.

McMeechan, A., M. Roberts, T.A. Cogan, F. Jørgensen, A. Stevenson, C. Lewis, G. Rowley & T.J. Humphrey, (2007) Role of the alternative sigma factors sigmaE and sigmaS in survival of Salmonella enterica serovar Typhimurium during starvation, refrigeration and osmotic shock. *Microbiology (Reading, England)* **153**: 263-269.

Mecsas, J., P.E. Rouviere, J.W. Erickson, T.J. Donohue & C.A. Gross, (1993) The activity of sigma E, an Escherichia coli heat-inducible sigma-factor, is modulated by expression of outer membrane proteins. *Genes Dev* **7**: 2618-2628.

Merdanovic, M., T. Clausen, M. Kaiser, R. Huber & M. Ehrmann, (2011) Protein Quality Control in the Bacterial Periplasm. *Annual Review of Microbiology* **65**: 149-168.

Meyer, R.R. & P.S. Laine, (1990) The single-stranded DNA-binding protein of Escherichia coli. *Microbiological reviews* **54**: 342-380.

Meyer, S.E., S. Granett, J.U. Jung & M.R. Villarejo, (1990) Osmotic regulation of PhoE porin synthesis in Escherichia coli. *J. Bacteriol.* **172**: 5501-5502.

Miller, S.I. & N.R. Salama, (2018) The gram-negative bacterial periplasm: Size matters. *PLoS Biology* **16**: 1-7.

Miller, W.G., J.H. Leveau & S.E. Lindow, (2000) Improved gfp and inaZ broad-host-range promoter-probe vectors. *Molecular plant-microbe interactions : MPMI* **13**: 1243-1250.

Miot, M. & J.-M. Betton, (2004) Microbial Cell Factories Protein quality control in the bacterial periplasm. *Microbial Cell Factories* **3:4**: 1-13.

Mira, A. & R. Pushker, (2005) The silencing of pseudogenes. *Molecular Biology and Evolution* **22**: 2135-2138.

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

Misra, R., R. Stikeleather & R. Gabriele, (2015) In vivo roles of BamA, BamB and BamD in the biogenesis of BamA, a core protein of the β-barrel assembly machine of Escherichia coli. *J. Mol. Biol.* **427**: 1061-1074.

MOLL, I., (2003) Coincident Hfq binding and RNase E cleavage sites on mRNA and small regulatory RNAs. *RNA* **9**: 1308-1314.

Moslavac, S., O. Mirus, R. Bredemeier, J. Soll, A. Von Haeseler & E. Schleiff, (2005) Conserved pore-forming regions in polypeptide-transporting proteins. *FEBS Journal* **272**: 1367-1378.

Mřazek, J. & S. Xie, (2006) Pattern locator: A new tool for finding local sequence patterns in genomic DNA sequences. *Bioinformatics* **22**: 3099-3100.

Müller, C.M., U. Dobrindt, G. Nagy, L. Emödy, B.E. Uhlin & J. Hacker, (2006) Role of histone-like proteins H-NS and StpA in expression of virulence determinants of uropathogenic Escherichia coli. *J. Bacteriol.* **188**: 5428-5438.

Mutalik, V.K., G. Nonaka, S.E. Ades, V.A. Rhodius & C.A. Gross, (2009) Promoter strength properties of the complete sigma E regulon of Escherichia coli and Salmonella enterica. *J. Bacteriol.* **191**: 7279-7287.

Nakayama, H., K. Kurokawa & B.L. Lee, (2012) Lipoproteins in bacteria: Structures and biosynthetic pathways. *FEBS Journal* **279**: 4247-4268.

Navarre, W.W., M. McClelland, S.J. Libby & F.C. Fang, (2007) Silencing of xenogenic DNA by H-NS: facilitation of lateral gene transfer in bacteria by a defence system that recognizes foreign DNA. *Genes & Development* **21**: 1456-1471.

Newman, C.L. & C. Stathopoulos, (2004) Autotransporter and two-partner secretion: delivery of large-size virulence factors by gram-negative bacterial pathogens. *Critical reviews in microbiology* **30**: 275-286.

Ni, D., Y. Wang, X. Yang, H. Zhou, X. Hou, B. Cao, Z. Lu, X. Zhao, K. Yang & Y. Huang, (2014a) Structural and functional analysis of the β-barrel domain of BamA from Escherichia coli. *FASEB Journal* **28**: 2677-2685.

Ni, D., K. Yang & Y. Huang, (2014b) Refolding, crystallization and preliminary X-ray crystallographic studies of the β-barrel domain of BamA, a membrane protein essential for outer membrane protein biogenesis. *Acta Crystallographica Section F:Structural Biology Communications* **70**: 362-365.

Nikaido, H., (1989) Outer membrane barrier as a mechanism of resistance. *Antimicrob. Agents Chemother.* **33**: 1831-1836.

Nikaido, H., (1994) Porins and specific diffusion channels in bacterial outer membranes. *J. Biol. Chem.* **269**: 3905-3908.

Nikaido, H., (2003) Molecular Basis of Bacterial Outer Membrane Permeability Revisited. *Microbiol. Mol. Biol. Rev.* **67**: 593-656.

Nikaido, H. & M. Vaara, (1985) Molecular basis of bacterial outer membrane permeability. *Microbiol Rev* **49**: 1-32.

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

Noinaj, N., M. Guillier, Barnard, Travis J. & S.K. Buchanan, (2010) TonB-Dependent Transporters: Regulation, Structure, and Function. *Annual Review of Microbiology* **64**: 43-60.

Noinaj, N., J.C. Gumbart & S.K. Buchanan, (2017) The β-barrel assembly machinery in motion. *Nat. Rev. Microbiol.* **15**: 197-204.

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

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

Noinaj, N., S.E. Rollauer & S.K. Buchanan, (2015) The β-barrel membrane protein insertase machinery from Gram-negative bacteria. *Current Opinion in Structural Biology* **31**: 35-42.

Ogasawara, H., K. Yamada, A. Kori, K. Yamamoto & A. Ishihama, (2010) Regulation of the Escherichia coli csgD promoter: Interplay between five transcription factors. *Microbiology* **156**: 2470-2483.

Okamoto, T., R.B. Schwab, P.E. Scherer & M.P. Lisanti, (2001) Analysis of the association of proteins with membranes. *Current protocols in cell biology* **Chapter 5**: Unit 5.4.

Ono, S., Martin D. Goldberg, T. Olsson, D. Esposito, Jay C.D. Hinton & John E. Ladbury, (2005) H-NS is a part of a thermally controlled mechanism for bacterial gene regulation. *Biochemical Journal* **391**: 203-213.

Onufryk, C., M.-L. Crouch, F.C. Fang & C.A. Gross, (2005) Characterization of six lipoproteins in the sigmaE regulon. *J. Bacteriol.* **187**: 4552-4561.

Owen-Hughes, T.A., G.D. Pavitt, D.S. Santos, J.M. Sidebotham, C.S.J. Hulton, J.C.D. Hinton & C.F. Higgins, (1992) The chromatin-associated protein H-NS interacts with curved DNA to influence DNA topology and gene expression. *Cell* **71**: 255-265.

Padilla, E., E. Llobet, A. Doménech-Sánchez, L. Martínez-Martínez, J.A. Bengoechea & S. Albertí, (2010) Klebsiella pneumoniae AcrAB efflux pump contributes to antimicrobial resistance and virulence. *Antimicrob. Agents Chemother.* **54**: 177-183.

Paetzel, M., A. Karla, N.C.J. Strynadka & R.E. Dalbey, (2002) Signal peptidases. *Chemical Reviews* **102**: 4549-4579.

Palmer, K.L. & R.S. Munson, (1995) Cloning and characterization of the genes encoding the hemolysin of Haemophilus ducreyi. *Mol. Microbiol.* **18**: 821-830.

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

Palonen, E., M. Lindström, P. Somervuo & H. Korkeala, (2013) Alternative sigma factor σe has an important role in stress tolerance of yersinia pseudotuberculosis ip32953. *Applied and Environmental Microbiology* **79**: 5970-5977.

Papenfort, K., V. Pfeiffer, F. Mika, S. Lucchini, J.C.D. Hinton & J. Vogel, (2006) σE-dependent small RNAs of Salmonella respond to membrane stress by accelerating global omp mRNA decay. *Mol. Microbiol.* **62**: 1674-1688.

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

Parker, J.L. & S. Newstead, (2016) The Next Generation in Membrane Protein Structure Determination. **922**: 61-72.

Parshin, A., A.L. Shiver, J. Lee, M. Ozerova, D. Schneidman-Duhovny, C.A. Gross & S. Borukhov, (2015) DksA regulates RNA polymerase in Escherichia coli through a network of interactions in the secondary channel that includes Sequence Insertion 1. *Proc. Natl. Acad. Sci. U. S. A.* **112**: E6862-6871.

Parthasarathi, K., L.S. Ranganathan, V. Anandi & J. Zeyer, (2007) Diversity of microflora in the gut and casts of tropical composting earthworms reared on different substrates. *Journal of Environmental Biology* **28**: 87-97.

Paul, B.J., M.M. Barker, W. Ross, D.A. Schneider, C. Webb, J.W. Foster & R.L. Gourse, (2004) DksA: A critical component of the transcription initiation machinery that potentiates the regulation of rRNA promoters by ppGpp and the initiating NTP. *Cell* **118**: 311-322.

Paulus, J.K. & R.A.L. van der Hoorn, (2018) Tricked or trapped—Two decoy mechanisms in host–pathogen interactions. *PLoS Pathog.* **14**: 1-6.

Pavlova, O., J.H. Peterson, R. Ieva & H.D. Bernstein, (2013) Mechanistic link between β barrel assembly and the initiation of autotransporter secretion. *Proc. Natl. Acad. Sci. U. S. A.* **110**: E938-947.

Perica, T., Y. Kondo, S.P. Tiwari, S.H. McLaughlin, K.R. Kemplen, X. Zhang, A. Steward, N. Reuter, J. Clarke & S.A. Teichmann, (2014) Evolution of oligomeric state through allosteric pathways that mimic ligand binding. *Science* **346**.

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

Pfitzner, A.-K., N. Steblau, T. Ulrich, P. Oberhettinger, I.B. Autenrieth, M. Schütz & D. Rapaport, (2016) Mitochondrial-bacterial hybrids of BamA/Tob55 suggest variable requirements for the membrane integration of β-barrel proteins. *Scientific Reports* **6**: 39053.

Plummer, A.M. & K.G. Fleming, (2015) BamA Alone Accelerates Outer Membrane Protein Folding in Vitro through a Catalytic Mechanism. *Biochemistry* **54**.

Plummer, A.M. & K.G. Fleming, (2016) From Chaperones to the Membrane with a BAM! *Trends in Biochemical Sciences* **41**: 872-882.

Podschun, R. & U. Ullmann, (1998) Klebsiella spp. as nosocomial pathogens: Epidemiology, taxonomy, typing methods, and pathogenicity factors. *Clin. Microbiol. Rev.* **11**: 589-603.

Price, C.E. & A.J.M. Driessen, (2010) Conserved negative charges in the transmembrane segments of subunit K of the NADH:ubiquinone oxidoreductase determine its dependence on YidC for membrane insertion. *J. Biol. Chem.* **285**: 3575-3581.

Prieto, A.I., C. Kahramanoglou, R.M. Ali, G.M. Fraser, A.S.N. Seshasayee & N.M. Luscombe, (2012) Genomic analysis of DNA binding and gene regulation by homologous nucleoid-associated proteins IHF and HU in Escherichia coli K12. *Nucleic Acids Research* **40**: 3524-3537.

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

Raina, S., D. Missiakas & C. Georgopoulos, (1995) The rpoE gene encoding the sigma E (sigma 24) heat shock sigma factor of Escherichia coli. *The EMBO journal* **14**: 1043-1055.

Reznikoff, W.S., (2003) Tn5 as a model for understanding dna transposition. *Mol. Microbiol.* **47**: 1199-1206.

Rezuchova, B., H. Miticka, D. Homerova, M. Roberts & J. Kormanec, (2003) New members of the Escherichia coli σEregulon identified by a two-plasmid system. *FEMS Microbiology Letters* **225**: 1-7.

Ricci, D.P., C.L. Hagan, D. Kahne & T.J. Silhavy, (2012) Activation of the Escherichia coli β-barrel assembly machine (Bam) is required for essential components to interact properly with substrate. *Proc. Natl. Acad. Sci. U. S. A.* **109**: 3487-3491.

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

Ricci, D.P. & T.J. Silhavy, (2012) The Bam machine: A molecular cooper. *Biochim. Biophys. Acta Biomembr.* **1818**: 1067-1084.

Rigel, N.W., D.P. Ricci & T.J. Silhavy, (2013) Conformation-specific labeling of BamA and suppressor analysis suggest a cyclic mechanism for -barrel assembly in Escherichia coli. *Proceedings of the National Academy of Sciences* **110**: 5151-5156.

Rigel, N.W., J. Schwalm, D.P. Ricci & T.J. Silhavy, (2012) BamE modulates the Escherichia coli beta-barrel assembly machine component BamA. *J. Bacteriol.* **194**: 1002-1008.

Rizzitello, A.E.M.Y.E., J.R. Harper & T.J. Silhavy, (2001) Genetic Evidence for Parallel Pathways of Chaperone Activity in the Periplasm of Escherichia coli Genetic Evidence for Parallel Pathways of Chaperone Activity in the Periplasm of Escherichia coli. *J. Bacteriol.* **183**: 6794-6800.

Rocha, D.J.P., C.S. Santos & L.G.C. Pacheco, (2015) Bacterial reference genes for gene expression studies by RT-qPCR: survey and analysis. *Antonie van Leeuwenhoek* **108**: 685-693.

Rocque, W.J. & E.J. McGroarty, (1989) Isolation and preliminary characterization of wild-type OmpC porin dimers from Escherichia coli K-12. *Biochemistry* **28**: 3738-3743.

Rollauer, S.E., M.A. Sooreshjani, N. Noinaj & S.K. Buchanan, (2015) Outer membrane protein biogenesis in Gram-negative bacteria. *Philosophical Transactions of the Royal Society of London B: Biological Sciences* **370**.

Rosenblum, R., E. Khan, G. Gonzalez, R. Hasan & T. Schneiders, (2011) Genetic regulation of the ramA locus and its expression in clinical isolates of Klebsiella pneumoniae. *International Journal of Antimicrobial Agents* **38**: 39-45.

Rowley, G., H. Skovierova, A. Stevenson, B. Rezuchova, D. Homerova, C. Lewis, A. Sherry, J. Kormanec & M. Roberts, (2011) The periplasmic chaperone Skp is required for successful Salmonella Typhimurium infection in a murine typhoid model. *Microbiology* **157**: 848-858.

Roy, A., A. Kucukural & Y. Zhang, (2010) I-TASSER: A unified platform for automated protein structure and function prediction. *Nature Protocols* **5**: 725-738.

Ruijter, J.M., C. Ramakers, W.M. Hoogaars, Y. Karlen, O. Bakker, M.J. van den Hoff & A.F. Moorman, (2009) Amplification efficiency: linking baseline and bias in the analysis of quantitative PCR data. *Nucleic Acids Res* **37**: e45.

Ruiz, N., B. Falcone, D. Kahne & T.J. Silhavy, (2005) Chemical conditionality: A genetic strategy to probe organelle assembly. *Cell* **121**: 307-317.

Ruzin, A., F.W. Immermann & P.A. Bradford, (2008) Real-time PCR and statistical analyses of acrAB and ramA expression in clinical isolates of Klebsiella pneumoniae. *Antimicrob. Agents Chemother.* **52**: 3430-3432.

Safdar, N. & D.G. Maki, (2005) Risk of Catheter-Related Bloodstream Infection With Peripherally Inserted Central Venous Catheters Used in Hospitalized Patients. *Chest* **128**: 489-495.

Saio, T., X. Guan, P. Rossi, A. Economou & C.G. Kalodimos, (2014) Structural Basis for Protein Antiaggregation Activity of the Trigger Factor Chaperone. *Science* **344**: 1250494-1250494.

Salacha, R., F. Kovačić, C. Brochier-Armanet, S. Wilhelm, J. Tommassen, A. Filloux, R. Voulhoux & S. Bleves, (2010) The Pseudomonas aeruginosa patatin-like protein PlpD is the archetype of a novel Type V secretion system. *Environmental Microbiology* **12**: 1498-1512.

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

Sánchez-Pulido, L., D. Devos, S. Genevrois, M. Vicente & A. Valencia, (2003) POTRA: a conserved domain in the FtsQ family and a class of beta-barrel outer membrane proteins. *Trends in biochemical sciences* **28**: 523-526.

Sankaran, K. & H.C. Wu, (1994) Lipid modification of bacterial prolipoprotein. *The Journal of biological chemistry* **269**: 19701-19706.

Schmittgen, T.D. & K.J. Livak, (2008) Analyzing real-time PCR data by the comparative CT method. *Nature Protocols* **3**: 1101-1108.

Schulz, G.E., (2002) The structure of bacterial outer membrane proteins. *Biochimica et biophysica acta* **1565**: 308-317.

Selkrig, J., M.J. Belousoff, S.J. Headey, E. Heinz, T. Shiota, H.H. Shen, S.A. Beckham, R.S. Bamert, M.D. Phan, M.A. Schembri, M.C.J. Wilce, M.J. Scanlon, R.A. Strugnell & T. Lithgow, (2015) Conserved features in TamA enable interaction with TamB to drive the activity of the translocation and assembly module. *Scientific Reports* **5**.

Selkrig, J., D.L. Leyton, C.T. Webb & T. Lithgow, (2014) Assembly of β-barrel proteins into bacterial outer membranes. *Biochimica et biophysica acta* **1843**: 1542-1550.

Selkrig, J., K. Mosbahi, C.T. Webb, M.J. Belousoff, A.J. Perry, T.J. Wells, F. Morris, D.L. Leyton, M. Totsika, M.-D. Phan, N. Celik, M. Kelly, C. Oates, E.L. Hartland, R.M. Robins-Browne, S.H. Ramarathinam, A.W. Purcell, M.A. Schembri, R.A. Strugnell, I.R. Henderson, D.

Walker & T. Lithgow, (2012) Discovery of an archetypal protein transport system in bacterial outer membranes. *Nature Structural & Molecular Biology* **19**: 506-510.

Seo, J.-H., J. Hong, D. Kim, B.-K. Cho, T.-W. Huang, S.-F. Tsai, B.O. Palsson & P. Charusanti, (2012) Multiple-omic data analysis of Klebsiella pneumoniae MGH 78578 reveals its transcriptional architecture and regulatory features. *BMC Genomics* **13**: 679.

Serra, D.O., M.S. Conover, L. Arnal, G.P. Sloan, M.E. Rodriguez, O.M. Yantorno & R. Deora, (2011) FHA-mediated cell-substrate and cell-cell adhesions are critical for Bordetella pertussis biofilm formation on abiotic surfaces and in the mouse nose and the trachea. *PLoS ONE* **6**.

Shakhnovich, B. & E. Koonin, (2006) Origins and impact of constraints in evolution of gene families. *Genome research*: 1529-1536.

Shankar, C., B. Veeraraghavan, L.E.B. Nabarro, R. Ravi, N.K.D. Ragupathi & P. Rupali, (2018) Whole genome analysis of hypervirulent Klebsiella pneumoniae isolates from community and hospital acquired bloodstream infection. *BMC Microbiology* **18**: 1-9.

Shea, A., M. Wolcott, S. Daefler & D.A. Rozak, (2012) Biolog Phenotype Microarrays. In: Microbial Systems Biology: Methods and Protocols. A. Navid (ed). Totowa, NJ: Humana Press, pp. 331-373.

Shen, H.-H., D.L. Leyton, T. Shiota, M.J. Belousoff, N. Noinaj, J. Lu, S.A. Holt, K. Tan, J. Selkrig, C.T. Webb, S.K. Buchanan, L.L. Martin & T. Lithgow, (2014) Reconstitution of a nanomachine driving the assembly of proteins into bacterial outer membranes. *Nature Communications* **5**: 5078.

Shon, A.S., R.P.S. Bajwa & T.A. Russo, (2013) Hypervirulent (hypermucoviscous) Klebsiella Pneumoniae: A new and dangerous breed. *Virulence* **4**: 107-118.

Shruthi, H., P. Anand, V. Murugan & K. Sankaran, (2010) Twin arginine translocase pathway and fast-folding lipoprotein biosynthesis in E. coli: Interesting implications and applications. *Molecular BioSystems* **6**: 999-1007.

Shultis, D.D., M.D. Purdy, C.N. Banchs & M.C. Wiener, (2006) Outer Membrane Active Transport: Structure of the BtuB:TonB Complex. *Science* **312**: 1396-1399.

Silhavy, T.J., D. Kahne & S. Walker, (2010) The Bacterial Cell Envelope. *Cold Spring Harbor Perspectives in Biology* **2**: a000414-a000414.

Silhavy, T.J. & J.C. Malinverni, (2011) Assembly of Outer Membrane β-Barrel Proteins: the Bam Complex. *EcoSal Plus* **4**.

Simmerman, R.F., A.M. Dave & B.D. Bruce, (2014) Structure and Function of POTRA Domains of Omp85/TPS Superfamily. *International Review of Cell and Molecular Biology* **308**: 1-34.

Singh, R., N. Capalash & P. Sharma, (2017a) Immunoprotective potential of BamA, the outer membrane protein assembly factor, against MDR Acinetobacter baumannii. *Scientific Reports* **7**: 1-11.

Singh, R., C. Kraft, R. Jaiswal, K. Sejwal, V.B. Kasaragod, J. Kuper, J. Bürger, T. Mielke, J. Luirink & S. Bhushan, (2014) Cryo-electron microscopic structure of SecA protein bound to the 70S ribosome. *J. Biol. Chem.* **289**: 7190-7199.

Singh, S.K., R. Ekka, M. Mishra & H. Mohapatra, (2017b) Association study of multiple antibiotic resistance and virulence: a strategy to assess the extent of risk posed by bacterial population in aquatic environment. *Environmental Monitoring and Assessment* **189**.

Sinnige, T., M. Weingarth, M. Renault, L. Baker, J. Tommassen & M. Baldus, (2014) Solid-State NMR studies of full-length BamA in lipid bilayers suggest limited overall POTRA mobility. *J. Mol. Biol.* **426**: 2009-2021.

Siu, L.K., K.M. Yeh, J.C. Lin, C.P. Fung & F.Y. Chang, (2012) Klebsiella pneumoniae liver abscess: A new invasive syndrome. *The Lancet Infectious Diseases* **12**: 881-885.

Sklar, J.G., T. Wu, L.S. Gronenberg, J.C. Malinverni, D. Kahne & T.J. Silhavy, (2007a) Lipoprotein SmpA is a component of the YaeT complex that assembles outer membrane proteins in Escherichia coli. *Proc. Natl. Acad. Sci. U. S. A.* **104**: 6400-6405.

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

Smith, D.L., C.E. James, M.J. Sergeant, Y. Yaxian, J.R. Saunders, A.J. McCarthy & H.E. Allison, (2007) Short-tailed Stx phages exploit the conserved YaeT protein to disseminate Shiga toxin genes among enterobacteria. *J. Bacteriol.* **189**: 7223-7233.

Snijder, H.J. & B.W. Dijkstra, (2000) Bacterial phospholipase A: Structure and function of an integral membrane phospholipase. *Biochimica et Biophysica Acta - Molecular and Cell Biology of Lipids* **1488**: 91-101.

Snitkin, E.S., A.M. Zelazny, P.J. Thomas, F. Stock, D.K. Henderson, T.N. Palmore & J.A. Segre, (2012) Tracking a Hospital Outbreak of Carbapenem-Resistant Klebsiella pneumoniae with Whole-Genome Sequencing. *Science Translational Medicine* **4**: 148ra116-148ra116.

Sohlenkamp, C. & O. Geiger, (2015) Bacterial membrane lipids: Diversity in structures and pathways. *FEMS Microbiology Reviews* **40**: 133-159.

Soltes, G.R., N.R. Martin, E. Park, H.A. Sutterlin & T.J. Silhavy, (2017) Distinctive roles for periplasmic proteases in the maintenance of essential outer membrane protein assembly. *J. Bacteriol.* **199**.

Southward, C.M. & M.G. Surette, (2002) The dynamic microbe: Green fluorescent protein brings bacteria to light. *Mol. Microbiol.* **45**: 1191-1196.

Srinivasan, V.B., B.B. Singh, N. Priyadarshi, N.K. Chauhan & G. Rajamohan, (2014) Role of novel multidrug efflux pump involved in drug resistance in Klebsiella pneumoniae. *PLoS ONE* **9**.

St Geme, J.W., S. Falkow & S.J. Barenkamp, (1993) High-molecular-weight proteins of nontypable Haemophilus influenzae mediate attachment to human epithelial cells. *Proc. Natl. Acad. Sci. U. S. A.* **90**: 2875-2879.

Stahlhut, S.G., C. Struve, K.A. Krogfelt & A. Reisner, (2012) Biofilm formation of Klebsiella pneumoniae on urethral catheters requires either type 1 or type 3 fimbriae. *FEMS immunology and medical microbiology* **65**: 350-359.

Stenkat, J., M.E. Krautwald-Junghanns, A. Schmitz Ornés, A. Eilers & V. Schmidt, (2014) Aerobic cloacal and pharyngeal bacterial flora in six species of free-living birds. *Journal of Applied Microbiology* **117**: 1564-1571.

Stock, A.M., V.L. Robinson & P.N. Goudreau, (2000) Two-component signal transduction. *Annual review of biochemistry* **69**: 183-215.

Storek, K.M., M.R. Auerbach, H. Shi, N.K. Garcia, D. Sun, N.N. Nickerson, R. Vij, Z. Lin, N. Chiang, K. Schneider, A.T. Wecksler, E. Skippington, G. Nakamura, D. Seshasayee, J.T. Koerber, J. Payandeh, P.A. Smith & S.T. Rutherford, (2018) Monoclonal antibody targeting the β-barrel assembly machine of Escherichia coli is bactericidal. *Proceedings of the National Academy of Sciences* **115**: 3692-3697.

Stroud, D.A., T. Becker, J. Qiu, D. Stojanovski, S. Pfannschmidt, C. Wirth, C. Hunte, B. Guiard, C. Meisinger, N. Pfanner & N. Wiedemann, (2011) Biogenesis of mitochondrial -barrel proteins: the POTRA domain is involved in precursor release from the SAM complex. *Molecular Biology of the Cell* **22**: 2823-2833.

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

Studier, F.W., (2005) Protein production by auto-induction in high density shaking cultures. *Protein expression and purification* **41**: 207-234.

Stukenberg, P.T., P.S. Studwell-Vaughan, M. O'Donnell & M. O'Donnell, (1991) Mechanism of the sliding beta-clamp of DNA polymerase III holoenzyme. *J. Biol. Chem.* **266**: 11328-11334.

Sugawara, E., S. Kojima & H. Nikaido, (2016) Klebsiella pneumoniae major porins OmpK35 and OmpK36 Allow more efficient diffusion of β-lactams than their escherichia coli homologs OmpF and OmpC. *J. Bacteriol.* **198**: 3200-3208.

Surana, N.K., A.Z. Buscher, G.G. Hardy, S. Grass, T. Kehl-Fie & J.W. St. Geme, (2006) Translocator proteins in the two-partner secretion family have multiple domains. *J. Biol. Chem.* **281**: 18051-18058.

Tambyah, P.A., V. Knasinski & D.G. Maki, (2002) The direct costs of nosocomial catheter-associated urinary tract infection in the era of managed care. *Infection control and hospital epidemiology* **23**: 27-31.

Tamm, L.K., H. Hong & B. Liang, (2004) Folding and assembly of beta-barrel membrane proteins. *Biochimica et biophysica acta* **1666**: 250-263.

Thellin, O., B. ElMoualij, E. Heinen & W. Zorzi, (2009) A decade of improvements in quantification of gene expression and internal standard selection. *Biotechnology Advances* **27**: 323-333.

Thoma, J., Y. Sun, N. Ritzmann & D.J. Müller, (2018) POTRA Domains, Extracellular Lid, and Membrane Composition Modulate the Conformational Stability of the β Barrel Assembly Factor BamA. *Structure*: 987-996.

Tokuda, H. & S.-i. Matsuyama, (2004) Sorting of lipoproteins to the outer membrane in E. coli. *Biochimica et biophysica acta* **1694**: IN1-9.

Tokuda, H. & S.-I. Narita, (2010) Biogenesis and Membrane Targeting of Lipoproteins. *EcoSal Plus* **4**.

Tokunaga, M., H. Tokunaga & H.C. Wu, (1982) Post-translational modification and processing of Escherichia coli prolipoprotein in vitro. *Proc Natl Acad Sci U S A* **79**: 2255-2259.

Tomás, A., L. Lery, V. Regueiro, C. Pérez-Gutiérrez, V. Martínez, D. Moranta, E. Llobet, M. González-Nicolau, J.L. Insua, J.M. Tomas, P.J. Sansonetti, R. Tournebize & J.A. Bengoechea, (2015) Functional genomic screen identifies klebsiella pneumoniae factors implicated in blocking nuclear factor ??B (NF-??B) signaling. *J. Biol. Chem.* **290**: 16678-16697.

Torres, V.V.L., E. Heinz, S.C. J., W.J. J., C. Hanwei, Y. Ji, C. Abigail, D.R. A., A. Felicity, W.C. T., D. Gordon, S.R. A., H.I. D. & L. Trevor, (2018) An investigation into the Omp85 protein BamK in hypervirulent Klebsiella pneumoniae, and its role in outer membrane biogenesis. *Mol. Microbiol.* **0**: 1-35.

Trouplin, V., N. Boucherit, L. Gorvel, F. Conti, G. Mottola & E. Ghigo, (2013) Bone Marrow-derived Macrophage Production. *Journal of Visualized Experiments*: 1-6.

Tsai, Y.K., C.P. Fung, J.C. Lin, J.H. Chen, F.Y. Chang, T.L. Chen & L.K. Siu, (2011) Klebsiella pneumoniae Outer membrane porins OmpK35 and OmpK36 play roles in both antimicrobial resistance and virulence. *Antimicrob. Agents Chemother.* **55**: 1485-1493.

Tutar, Y., (2012) Pseudogenes. *Comparative and Functional Genomics* **2012**: 6-9.

Typas, A., G. Becker & R. Hengge, (2007) The molecular basis of selective promoter activation by the σs subunit of RNA polymerase. *Mol. Microbiol.* **63**: 1296-1306.

UniProt, C., (2015) UniProt: a hub for protein information. *Nucleic Acids Res* **43**: D204-212.

ur Rahman, S. & P. van Ulsen, (2013) System specificity of the TpsB transporters of coexpressed two-partner secretion systems of Neisseria meningitidis. *J. Bacteriol.* **195**: 788-797.

Urfer, M., J. Bogdanovic, F.L. Monte, K. Moehle, K. Zerbe, U. Omasits, C.H. Ahrens, G. Pessi, L. Eberl & J.A. Robinson, (2016) A peptidomimetic antibiotic targets outer membrane proteins and disrupts selectively the outer membrane in Escherichia coli. *J. Biol. Chem.* **291**: 1921-1932.

Vaara, M., (1992) Agents that increase the permeability of the outer membrane. *Microbiological reviews* **56**: 395-411.

Van Laar, T.A., T. Chen, T. You & K.P. Leung, (2015) Sublethal concentrations of carbapenems alter cell morphology and genomic expression of Klebsiella pneumoniae biofilms. *Antimicrob. Agents Chemother.* **59**: 1707-1717.

Van Ulsen, P., S.u. Rahman, W.S.P. Jong, M.H. Daleke-Schermerhorn & J. Luirink, (2014) Type V secretion: From biogenesis to biotechnology. In: Biochimica et Biophysica Acta - Molecular Cell Research. pp.

Van Ulsen, P., L. Rutten, M. Feller, J. Tommassen & A. Van Der Ende, (2008) Two-partner secretion systems of Neisseria meningitidis associated with invasive clonal complexes. *Infect. Immun.* **76**: 4649-4658.

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

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

Vij, R., Z. Lin, N. Chiang, J.M. Vernes, K.M. Storek, S. Park, J. Chan, Y.G. Meng, L. Comps-Agrar, P. Luan, S. Lee, K. Schneider, J. Bevers, I. Zilberleyb, C. Tam, C.M. Koth, M. Xu, A. Gill, M.R. Auerbach, P.A. Smith, S.T. Rutherford, G. Nakamura, D. Seshasayee, J. Payandeh & J.T. Koerber, (2018) A targeted boost-and-sort immunization strategy using Escherichia coli BamA identifies rare growth inhibitory antibodies. *Scientific Reports* **8**: 1-10.

Vilain, S., P. Cosette, G.-A. Junter & T. Jouenne, (2002) Phosphate deprivation is associated with high resistance to latamoxef of gel-entrapped, sessile-like Escherichia coli cells. *The Journal of antimicrobial chemotherapy* **49**: 315-320.

Vlahoviček, K., L. Kaján & S. Pongor, (2003) DNA analysis servers: Plot.it, bend.it, model.it and IS. *Nucleic Acids Research* **31**: 3686-3687.

Vogt, J. & G.E. Schulz, (1999) The structure of the outer membrane protein OmpX from Escherichia coli reveals possible mechanisms of virulence. *Structure (London, England : 1993)* **7**: 1301-1309.

Vollmer, W., D. Blanot & M.A. De Pedro, (2008) Peptidoglycan structure and architecture. *FEMS Microbiology Reviews* **32**: 149-167.

Volokhina, E.B., J. Grijpstra, F. Beckers, E. Lindh, V. Robert, J. Tommassen & M.P. Bos, (2013) Species-specificity of the bama component of the bacterial outer membrane protein-assembly machinery. *PLoS ONE* **8**: 1-12.

von Heijne, G., (1990) The signal peptide. *The Journal of membrane biology* **115**: 195-201.

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

Vuong, P., D. Bennion, J. Mantei, D. Frost & R. Misra, (2008) Analysis of YfgL and YaeT interactions through bioinformatics, mutagenesis, and biochemistry. *J. Bacteriol.* **190**: 1507-1517.

Waksman, G. & S.J. Hultgren, (2009) Structural biology of the chaperone-usher pathway of pilus biogenesis. *Nat. Rev. Microbiol.* **7**: 765-774.

Wallin, E. & G.V. Heijne, (2008) Genome-wide analysis of integral membrane proteins from eubacterial, archaean, and eukaryotic organisms. *Protein Science* **7**: 1029-1038.

Walther, D.M., D. Rapaport & J. Tommassen, (2009) Biogenesis of β-barrel membrane proteins in bacteria and eukaryotes: Evolutionary conservation and divergence. *Cell. Mol. Life Sci.* **66**: 2789-2804.

Walton, T.A., C.M. Sandoval, C.A. Fowler, A. Pardi & M.C. Sousa, (2009) The cavity-chaperone Skp protects its substrate from aggregation but allows independent folding of substrate domains. *Proceedings of the National Academy of Sciences* **106**: 1772-1777.

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

Ward, J.J., L.J. McGuffin, K. Bryson, B.F. Buxton & D.T. Jones, (2004) The DISOPRED server for the prediction of protein disorder. *Bioinformatics* **20**: 2138-2139.

Webb, C.T., E. Heinz & T. Lithgow, (2012a) Evolution of the beta-barrel assembly machinery. *Trends Microbiol* **20**: 612-620.

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

Wedege, E., K. Lie, K. Bolstad, V.E. Weynants, A. Halstensen, T.K. Herstad, J. Kreutzberger, L. Nome, L.M. Næss & A. Aase, (2013) Meningococcal Omp85 in detergent-extracted outer membrane vesicle vaccines induces high levels of non-functional antibodies in mice. *Scandinavian Journal of Immunology* **77**: 452-459.

Wilksch, J.J., J. Yang, A. Clements, J.L. Gabbe, K.R. Short, H. Cao, R. Cavaliere, C.E. James, C.B. Whitchurch, M.A. Schembri, M.L. Chuah, Z.X. Liang, O.L. Wijburg, A.W. Jenney, T. Lithgow & R.A. Strugnell, (2011) MrkH, a novel c-di-GMP-dependent transcriptional activator, controls Klebsiella pneumoniae biofilm formation by regulating type 3 fimbriae expression. *PLoS Pathog* **7**: e1002204.

Williams, A.B. & P.L. Foster, (2007) The Escherichia coli histone-like protein HU has a role in stationary phase adaptive mutation. *Genetics* **177**: 723-735.

Willis, L.M. & C. Whitfield, (2013) Capsule and lipopolysaccharide. *Escherichia coli: Pathotypes and Principles of Pathogenesis: Second Edition*: 533-556.

Wimley, W.C., (2002) Toward genomic identification of beta-barrel membrane proteins: composition and architecture of known structures. *Protein science : a publication of the Protein Society* **11**: 301-312.

Wimley, W.C., (2003) The versatile β-barrel membrane protein. *Current Opinion in Structural Biology* **13**: 404-411.

Wojtuszewski, K. & I. Mukerji, (2003) HU binding to bent DNA: A fluorescence resonance energy transfer and anisotropy study. *Biochemistry* **42**: 3096-3104.

Wu, T., J. Malinverni, N. Ruiz, S. Kim, T.J. Silhavy & D. Kahne, (2005) Identification of a multicomponent complex required for outer membrane biogenesis in Escherichia coli. *Cell* **121**: 235-245.

Wülfing, C. & A. Plückthun, (1994) Protein folding in the periplasm of Escherichia coli. *Mol. Microbiol.* **12**: 685-692.

Wzorek, J.S., J. Lee, D. Tomasek, C.L. Hagan & D.E. Kahne, (2017) Membrane integration of an essential β-barrel protein prerequires burial of an extracellular loop. *Proceedings of the National Academy of Sciences* **114**: 2598-2603.

Xu, X., S. Wang, Y.-X. Hu & D.B. McKay, (2007) The periplasmic bacterial molecular chaperone SurA adapts its structure to bind peptides in different conformations to assert a sequence preference for aromatic residues. *J. Mol. Biol.* **373**: 367-381.

Yamada, H., T. Yoshida, K.-i. Tanaka, C. Sasakawa & T. Mizuno, (1991) Molecular analysis of the Escherichia coli hns gene encoding a DNA-binding protein, which preferentially recognizes curved DNA sequences. *Molecular & general genetics : MGG* **230**: 332-336.

Yamagata, H., N. Taguchi, K. Daishima & S. Mizushima, (1983) Genetic characterization of a gene for prolipoprotein signal peptidase in Escherichia coli. *Mol Gen Genet* **192**: 10-14.

Yan, Z., M. Yin, D. Xu, Y. Zhu & X. Li, (2017) Structural insights into the secretin translocation channel in the type II secretion system. *Nature structural & molecular biology* **24**: 177-183.

Yang, F.L. & V. Braun, (2000) ShlB mutants of Serratia marcescens allow uncoupling of activation and secretion of the ShlA hemolysin. *International Journal of Medical Microbiology* **290**: 529-538.

Yen, M.R., C.R. Peabody, S.M. Partovi, Y. Zhai, Y.H. Tseng & M.H. Saier, (2002) Protein-translocating outer membrane porins of Gram-negative bacteria. *Biochim. Biophys. Acta Biomembr.* **1562**: 6-31.

Yoshida, K., T. Matsumoto, K. Tateda, K. Uchida, S. Tsujimoto & K. Yamaguchi, (2000) Role of bacterial capsule in local and systemic inflammatory responses of mice during pulmonary infection with Klebsiella pneumoniae. *Journal of medical microbiology* **49**: 1003-1010.

Yoshida, T., L. Qin, L.A. Egger & M. Inouye, (2006) Transcription regulation of ompF and ompC by a single transcription factor, OmpR. *J. Biol. Chem.* **281**: 17114-17123.

Zamze, S., L. Martinez-Pomares, H. Jones, P.R. Taylor, R.J. Stillion, S. Gordon & S.Y.C. Wong, (2002) Recognition of bacterial capsular polysaccharides and lipopolysaccharides by the macrophage mannose receptor. *J. Biol. Chem.* **277**: 41613-41623.

Zaslaver, A., A. Bren, M. Ronen, S. Itzkovitz, I. Kikoin, S. Shavit, W. Liebermeister, M.G. Surette & U. Alon, (2006) A comprehensive library of fluorescent transcriptional reporters for Escherichia coli. *Nature Methods* **3**: 623-628.

Zgurskaya, H.I., G. Krishnamoorthy, A. Ntreh & S. Lu, (2011) Mechanism and function of the outer membrane channel TolC in multidrug resistance and physiology of enterobacteria. *Front. Microbiol.* **2**: 1-13.

Zhang, a., S. Rimsky, M.E. Reaban, H. Buc & M. Belfort, (1996) Escherichia coli protein analogs StpA and H-NS: regulatory loops, similar and disparate effects on nucleic acid dynamics. *The EMBO journal* **15**: 1340-1349.

Zhang, G., T.C. Meredith & D. Kahne, (2013) On the essentiality of lipopolysaccharide to Gram-negative bacteria. *Current opinion in microbiology* **16**: 779-785.

Zhang, H., Z.Q. Gao, H.F. Hou, J.H. Xu, L.F. Li, X.D. Su & Y.H. Dong, (2011) High-resolution structure of a new crystal form of BamA POTRA4-5 from Escherichia coli. *Acta*

Crystallographica Section F: Structural Biology and Crystallization Communications **67**: 734-738.

Zhang, X.C., M. Liu & L. Han, (2017) Energy coupling mechanisms of AcrB-like RND transporters. *Biophysics Reports* **3**: 73-84.

Zhang, Y., (2008) I-TASSER server for protein 3D structure prediction. *BMC Bioinformatics* **9**: 1-8.

Ziegler, K., R. Benz & G.E. Schulz, (2008) A Putative α-Helical Porin from Corynebacterium glutamicum. *J. Mol. Biol.* **379**: 482-491.

Zilsel, J., P.H. Ma & J.T. Beatty, (1992) Derivation of a mathematical expression useful for the construction of complete genomic libraries. *Gene* **120**: 89-92.

Zuker, M., (2003) Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Research* **31**: 3406-3415.

Appendix 1.0
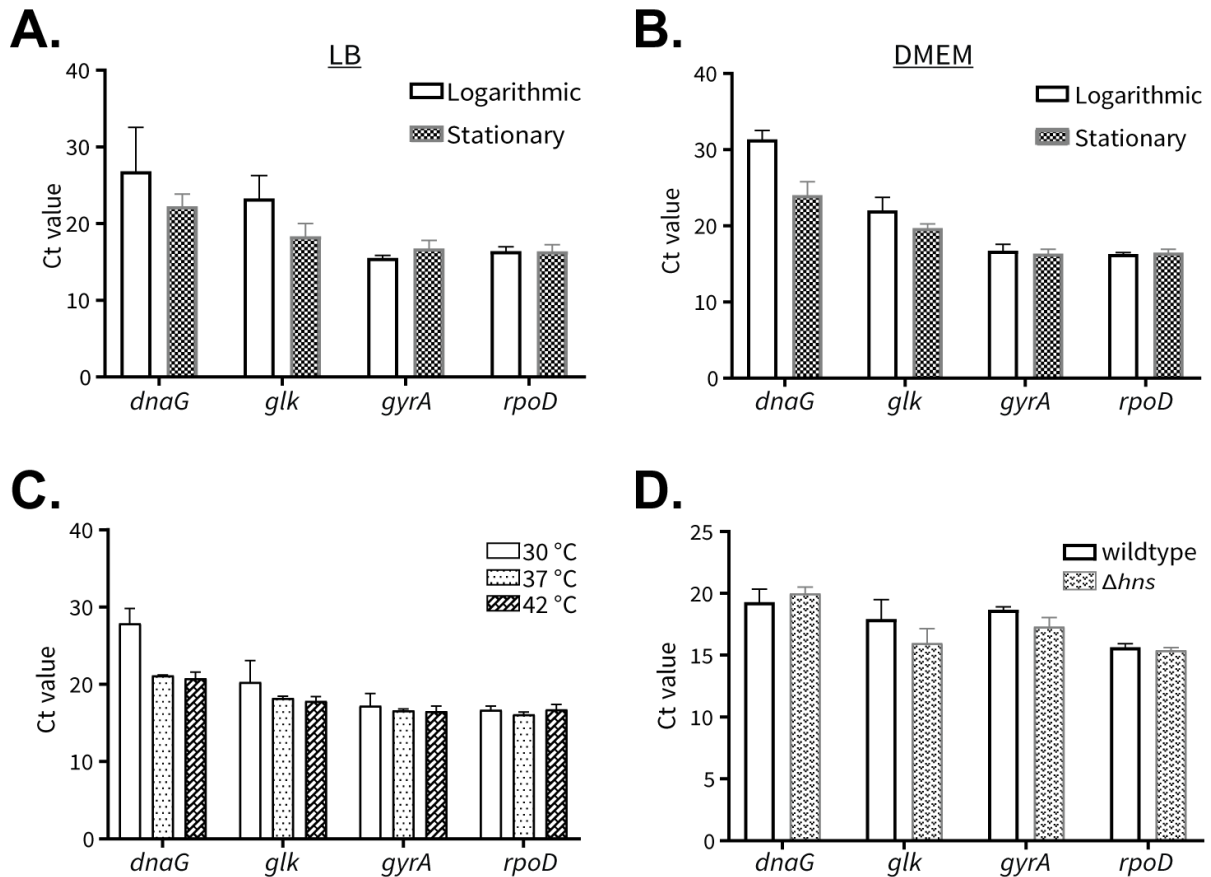
Comparison of candidate reference genes of K. pneumoniae B5055 cultured for thesis figures in Chapter 3.0. The expression levels of candidate reference (dnaG, glk, gyrA and rpoD) genes were determined using cycle threshold (Ct) values through real-time qPCR. A) Reference gene Ct values for Figure 3.1.2. B) Reference gene Ct values for Figure 3.1.3. C) Reference gene Ct values for Figure 3.1.4. D) Reference gene Ct values for Figure 3.2.5. A "Ct value" represents the amplification cycle number at which the fluorescence signal of the reporter dye reaches above an arbitrarily placed baseline threshold (Schmittgen & Livak, 2008). Thus, a lower Ct value corresponds to a higher abundance of a transcript, whereas a higher Ct value corresponds to a lower abundance of transcript. (n=3, error bars represent SD).

Appendix 2.0

Multiple sequence alignment of putative *bamK* promoters (500 bp upstream) from select strains representing
3 diverse phylogroups.

1) *K. pneumoniae*: B5055, AJ218 and MGH78578
2) *K. quasipnemuoniae*: UI9552 and AJ055
3) *K. variicola*: PUS13542 and AJ292

The putative -35 and -10 region of a σ^{70} promoter is labelled in red.

A palindromic sequence is underlined, probable ribosome binding site in blue. Translational start codon in green.

```
# Percent Identity Matrix - created by Clustal2.1
#
#
     1: UI9552    100.00    98.60    89.58    89.58    89.58    91.18    90.98
     2: AJ055      98.60   100.00    89.78    89.78    89.78    91.38    91.18
     3: B5055      89.58    89.78   100.00   100.00   100.00    93.20    93.00
     4: AJ218      89.58    89.78   100.00   100.00   100.00    93.20    93.00
     5: MGH78578   89.58    89.78   100.00   100.00   100.00    93.20    93.00
     6: Pus13542   91.18    91.38    93.20    93.20    93.20   100.00    99.80
     7: AJ292      90.98    91.18    93.00    93.00    93.00    99.80   100.00
```

```
CLUSTAL O (1.2.4) multiple sequence alignment
UI9552      aaggcgtccctactgtattctgctattttgaggaaaatatggagatgttcgccatcggcg    60
AJ055       aaggcgtccctaatgtattctgctattttgaggaaaatatgaagatgttcgccttcggcg    60
B5055       -aggcgaccctactgtattctgctattttgaggaaaatatggagatgtttgccaacggcg    59
AJ218       -aggcgaccctactgtattctgctattttgaggaaaatatggagatgtttgccaacggcg    59
MGH78578    -aggcgaccctactgtattctgctattttgaggaaaatatggagatgtttgccaacggcg    59
Pus13542    -aggcgaccctactgtattctgctattttgaggaaaatatggagatgttcgttaacggcc    59
AJ292       -aggcgaccctactgtattctgctattttgaggaaaatatggagatgttcgttaacggcc    59
              ***** ***** ************************** ******* *    ****

UI9552      tctggtggtgaatacaagataaaatgttgtgatgtcgattgcataatttattcaatttct   120
AJ055       tctggcggtgaatacaagatagaatgttgtgatgtcgattgcataatttattcaatttct   120
B5055       tctggcggtgaattcggggtagaatgttgtgatatcagttgcataatttattcaatttct   119
AJ218       tctggcggtgaattcggggtagaatgttgtgatatcagttgcataatttattcaatttct   119
MGH78578    tctggcggtgaattcggggtagaatgttgtgatatcagttgcataatttattcaatttct   119
Pus13542    tttggcagtgtatacggggtagaatgttgtgatatcagttgcataatttattcaatttct   119
AJ292       tttggcagtgtatacggggtagaatgttgtgatatcagttgcataatttattcaatttct   119
            * ***  ***  ** *  * ** ********** **  ****************** ****

UI9552      cttcattatatctgctggctttaagaatatgattgtgcctgttcctggtgattccagaga   180
AJ055       cttcattatatctgctggctttaagaatatgattgtgcctgttcctggtgattccagaga   180
B5055       cttcattaaatctgctgggttgcaggttattttctccctgttcctggtgatcccagaga   179
AJ218       cttcattaaatctgctggggttgcaggttattttttctccctgttcctggtgatcccagaga   179
MGH78578    cttcattaaatctgctgggttgcaggttattttttctccctgttcctggtgatcccagaga   179
Pus13542    cttcattatatctgctgggtttcaggatattttttctccctgttcctggtgatgccagaga   179
AJ292       cttcattatatctgctgggtttcaggatattttttctccctgttcctggtgatcccagaga   179
            ********  *********  **   **   ***   ** * *************** *******

UI9552      acaacattcctaacgagatccttcccccttttccattagcaagaaaggggggcttttttcaat   240
AJ055       acaacattcctaacgagatccttcccccttttccattagcaagaaaggggggcttttttcaat   240
B5055       gcatcattcctaacgagatccttcccccttttccattaataagaaaggggggcttttttatta   239
AJ218       gcatcattcctaacgagatccttcccccttttccattaataagaaaggggggcttttttatta   239
MGH78578    gcatcattcctaacgagatccttcccccttttccattaataagaaaggggggcttttttatta   239
Pus13542    gcatcattcctaacgagaattttcccccttttcattaacaagaaaggggggctttttttaat   239
AJ292       gcatcattcctaacgagaattttcccccttttcattaacaagaaaggggggcttttttttaat   239
             ** **************    **********  *****   ***************

UI9552      aaagaccgttcaggtaaatctaaatatagctctggtgttggccattagaatcacggcgtt   300
AJ055       aaagaccgttcaggtaaatctaaatatagctctggtgttggccattagaaccacggcgtt   300
B5055       ataaaacgttcaggcaaatctaaatatagtactggtgttggccattagaaccatgcgctt   299
AJ218       ataaaacgttcaggcaaatctaaatatagtactggtgttggccattagaaccatgcgctt   299
MGH78578    ataaaacgttcaggcaaatctaaatatagtactggtgttggccattagaaccatgcgctt   299
Pus13542    aaaaaacgttcaggtaaatctaaatatagccctggtgttggccattagaaccacggcgtt   299
AJ292       aaaaaacgttcaggtaaatctaaatatagccctggtgttggccattagaaccacggcgtt   299
            * * * ********  **************  ****************** ** *    **

UI9552      caggaaaagttccatcattaaatatattttttaaacatgatttaacatcctttttttctcc   360
AJ055       caggaaaagttccatcattaaatatatttttaaacatgatttaacatcctttttttctcc   360
B5055       cagaaaaagttccatcattaaatatatttttaaagatgctttaacatccatttttttctcc   359
AJ218       cagaaaaagttccatcattaaatatatttttaaagatgctttaacatccattttttctcc   359
MGH78578    cagaaaaagttccatcattaaatatatttttaaagatgctttaacatccattttttctcc   359
Pus13542    cagcaaaagttccatcattaaatatatttttaaagatgatttaacatcctttttttctcc   359
AJ292       cagcaaaagttccatcattaaatatatttttaaagatgatttaacatcctttttttctcc   359
            *** ************************************ *** ********** **********

UI9552      attttgtcgccataattgcatgtaattctgcgcctcgctcaggttat-ttctgtcgcata   419
AJ055       attttgtcgccataattgcatgtaattctgcgcctcgctcaggttat-ttctgtcgcata   419
B5055       attttgtcgccataattgcatgtaattctgcgcctcgctcaggttattttctgtcgcata   419
AJ218       attttgtcgccataattgcatgtaattctgcgcctcgctcaggttattttctgtcgcata   419
MGH78578    attttgtcgccataattgcatgtaattctgcgcctcgctcaggttattttctgtcgcata   419
Pus13542    attttatcgccataattgcatgtaattctgcgcctcgctcaggttattttctgtcgcaca   419
AJ292       attttatcgccataattgcatgtaattctgcgcctcgctcaggttattttctgtcgcaca   419
            ***** **************************************** ********** *

UI9552      tcttagatacgcatcctgaacaaataagctcaacgccagcagataactgctattgctcat   479
AJ055       tcttagatatgcatcctgaacaaataagctcaacgccagcagataactgctattgctcat   479
B5055       ttttagagatgcatcctgaacaaataagctcaacgccagcaggacactgctattgctcat   479
AJ218       ttttagatgatgcatcctgaacaaataagctcaacgccagcaggacactgctattgctcat   479
MGH78578    ttttagagatgcatcctgaacaaataagctcaacgccagcaggacactgctattgctcat   479
Pus13542    ttttagagatgcatcctgaacaaataagctcaacgccagcagaatactgctattgctcat   479
AJ292       ttttagagatgcatcctgaacaaataagctcaacgccagcagaatactgctattgctcat   479
            * ***** * **********************************     **************

UI9552      tatctcgtagacgacacattc    500 ATG
AJ055       tatctcgtagacgacacattc    500 ATG
B5055       tatctcgtagacgacacattc    500 ATG
AJ218       tatctcgtagacgacacattc    500 ATG
MGH78578    tatctcgtagacgacacattc    500 ATG
Pus13542    tatctcgtagacgacacattc    500 ATG
AJ292       tatctcgtagacgacacattc    500 ATG
            *********************
```

Appendix 3.0

Meta-analysis for proteins identified from suspected DNA::protein band of Figure 3.4.2.

Protein ID (UniProt)	Molecular weight (kDa)	Sequence Coverage (%)
W1G8T1_KLEPN	23.211	26.9
A6TB20_KLEPN	22.229	35.2
W1HVL4_KLEPN	17.984	21.8
A6TB36_KLEP7	9.9652	10
A6TEE8_KLEP7	28.611	15.2
A6TEK0_KLEPN	17.193	22.9
W1DQ00_KLEPN	35.48	21.5
W1DDG4_KLEPN	23.783	27.6
W1HWS1_KLEPN	54.906	8.1
A6TE29_KLEPN	23.447	16.6
W1DM81_KLEPN	28.647	25.5
W1HSH6_KLEPN	24.19	9.9
W1EA36_KLEPN	42.293	28.4
W1GBT1_KLEPN	28.926	22.6
A6T882_KLEP7	38.79	33.1
W1HS05_KLEPN	97.522	21.3
W1E9H0_KLEPN	36.363	26.1
A6T8M5_KLEP7	41.756	33
W1DMJ5_KLEPN	30.488	31.6
W1HTV3_KLEPN	16.32	22.6
W1H4H5_KLEPN	72.351	12
W1E4G7_KLEPN	35.69	27.6
W1H6R4_KLEPN	24.486	26.1
W1GNN4_KLEPN	30.824	61.3
W1HT41_KLEPN	75.094	28.6
J2DKG7_KLEPN	40.362	5.3
W1HLN9_KLEPN	12.576	29.9
W1E302_KLEPN	44.471	17
W1E2F9_KLEPN	40.644	7.6
W1H3J7_KLEPN	43.539	36.9
W1E1F6_KLEPN	20.913	54
W1DLV6_KLEPN	42.002	41.1
W1GMP7_KLEPN	27.333	11.7
W1E2G9_KLEPN	98.783	23.4
W1DQS8_KLEPN	31.572	18.6
W1E722_KLEPN	60.944	40
W1DCL9_KLEPN	39.641	47.8
W1I1L6_KLEPN	26.941	12.7
W1DU33_KLEPN	19.361	5.2
W1HY43_KLEPN	40.481	19.4
seq0008_CDS_-32	30.294	5.9
W1E3I1_KLEPN	47.113	21.1
W1E186_KLEPN	44.264	23.8

W1HS10_KLEPN	21.722	56.2
W1HYZ5_KLEPN	69.108	52
W1E8H4_KLEPN	22.594	31.4
W1DN72_KLEPN	28.627	20.9
W1GYA8_KLEPN	102.72	10.5
W1HYJ3_KLEPN	21.289	33.3
W1GYB2_KLEPN	32.674	13.2
W1E4R7_KLEPN	35.266	63.4
W1E3Q0_KLEPN	21.553	15.3
W1E673_KLEPN	39.877	50
W1H1J2_KLEPN	77.542	54.4
W1GXY7_KLEPN	48.848	39.8
W1DV12_KLEPN	59.547	56.5
W1DIA3_KLEPN	76.703	30.4
W1DNS4_KLEPN	20.743	22.6
W1GGT9_KLEPN	41.316	37.5
W1DVK8_KLEPN	38.946	25.8
W1DTJ5_KLEPN	23.244	37.9
W1DHT3_KLEPN	30.866	31.5
W1E358_KLEPN	17.572	18.6
W1DJJ9_KLEPN	28.647	19
W1HSQ0_KLEPN	32.77	5.9
W1HWJ7_KLEPN	12.456	28.3
W1DGI4_KLEPN	46.894	31.7
W1HS26_KLEPN	14.002	31.5
W1DN97_KLEPN	10.649	13.5
W1HU18_KLEPN	33.8	10.2
W1I0G4_KLEPN	84.344	30
W1HUS2_KLEPN	11.851	47.7
W1H297_KLEPN	24.504	52.6
W1DUL1_KLEPN	30.45	36.4
W1DCW5_KLEPN	8.986	36.2
W1GZN6_KLEPN	34.559	47.5
W1DFX4_KLEPN	18.272	17.6
UPI0003529D3B	34.269	5.9
A6T9G3_KLEP7	53.703	5.3
W1I1E4_KLEPN	6.4692	32.8
W1DNA0_KLEPN	16.061	21.9
W1DPZ8_KLEPN	34.49	14
W1HQ34_KLEPN	43.156	27.5
W1DZ02_KLEPN	34.953	41.9
W1DWY3_KLEPN	37.207	5.6
W1E946_KLEPN	26.175	38.2
W1DLE2_KLEPN	33.025	40.1
W1DWJ2_KLEPN	20.763	64.2
W1HWF9_KLEPN	36.99	50.9

W1DUW5_KLEPN	41.328	11
W1DKQ1_KLEPN	24.649	17.6
W1DWR4_KLEPN	27.121	10.6
W1GT99_KLEPN	63.598	42.9
W1H3T5_KLEPN	28.318	71.6
W1E5F2_KLEPN	97.055	37.6
W1I3B9_KLEPN	58.405	31.1
W1DE99_KLEPN	20.722	13.9
A6T6E3_KLEP7	27.013	4.5
W1E4E3_KLEPN	47.891	23.9
A6TD99_KLEP7	40.397	6.1
W1E8E8_KLEPN	58.302	11.9
W1DR59_KLEPN	45.549	46.1
W1DHP8_KLEPN	21.75	43.4
W1DIU6_KLEPN	58.1	25.7
W1DHD6_KLEPN	27.837	19.7
W1E669_KLEPN	43.579	19.9
W1HXL5_KLEPN	22.369	64.5
W1DTM5_KLEPN	49.398	40
W1HXB2_KLEPN	8.3884	16.5
W1GHD1_KLEPN	20.871	45.1
W1HNH7_KLEPN	16.934	27
W1E5Y8_KLEPN	71.106	47
J2LVZ9_KLEPN	33.154	7.1
W1HZ56_KLEPN	42.567	23.5
W1HVX2_KLEPN	48.156	27.8
W1DP63_KLEPN	32.497	28.9
W1DKD0_KLEPN	17.825	19.4
W1DMT2_KLEPN	33.287	27.1
W1HND3_KLEPN	34.984	26.9
W1HM18_KLEPN	25.21	13.4
W1HMR7_KLEPN	51.853	19.6
W1G642_KLEPN	14.728	9
W1HYH2_KLEPN	21.705	5.9
W1E781_KLEPN	30.547	25.5
W1GY35_KLEPN	54.493	38.4
W1H6M5_KLEPN	18.247	26
W1HQ84_KLEPN	34.589	77.1
W1HNM6_KLEPN	24.919	34.1
W1HV00_KLEPN	41.58	19.6
W1DY12_KLEPN	50.31	51.8
W1HSM9_KLEPN	12.429	23.2
W1HYV6_KLEPN	53.586	51.7
W1HPU5_KLEPN	45.439	51.1
W1E458_KLEPN	50.396	61.8
W1E731_KLEPN	58.792	59.6

J2E316_KLEPN	65.395	7.1
W1HTU4_KLEPN	45.056	50.5
W1DUP1_KLEPN	71.604	3.9
W1E5B3_KLEPN	15.526	62.2
W1HZU8_KLEPN	85.648	3.7
W1G8P3_KLEPN	43.507	43.4
W1DR90_KLEPN	63.267	21.7
W1DTR9_KLEPN	56.52	4.8
W1DUC4_KLEPN	43.519	43
W1DY56_KLEPN	37.576	14.4
W1HQH2_KLEPN	20.274	58.1
A6TBV0_KLEP7	43.441	15.1
W1DNU3_KLEPN	33.076	21.1
W1HPV1_KLEPN	21.946	9.6
W1H2R8_KLEPN	13.8	21.5
W1DRT6_KLEPN	69.095	12.2
W1GJ52_KLEPN	30.628	24.2
W1GWV4_KLEPN	15.154	8.1
W1I363_KLEPN	13.771	29.9
W1HRH2_KLEPN	65.489	15.1
W1DMM4_KLEPN	48.532	23.1
W1DFK8_KLEPN	19.721	55.7
W1HUA5_KLEPN	47.165	38.9
W1DVI8_KLEPN	63.383	22.3
W1HV18_KLEPN	108.25	32.2
W1EAN3_KLEPN	13.589	49.2
W1HUZ8_KLEPN	29.247	10.3
W1HR31_KLEPN	21.461	14.3
W1DZ11_KLEPN	32.564	17
W1HZ52_KLEPN	61.328	32.6
A6TIR6_KLEP7	17.867	72.9
A6T9B7_KLEP7	33.084	16.3
W1DEE5_KLEPN	28.541	52.3
W1H3C5_KLEPN	25.9	75.9
W1DSB4_KLEPN	77.159	24.2
W1DMT6_KLEPN	39.284	22.3
ADD_KLEP7	36.23	32.1
W1GL09_KLEPN	32.085	22.8
W1DSF7_KLEPN	27.636	11
W1HST0_KLEPN	34.625	9.1
W1HUP2_KLEPN	20.158	12.2
W1DUG1_KLEPN	10.483	19.8
W1HPP3_KLEPN	10.235	15.5
W1DI91_KLEPN	23.926	9.7
W1DNP4_KLEPN	9.4768	41.1
W1E4J4_KLEPN	12.341	25.6

W1GZP4_KLEPN	51.217	23.1
W1E2K4_KLEPN	39.142	36.2
W1H028_KLEPN	39.188	27.7
W1DRI7_KLEPN	57.546	37
W1DM95_KLEPN	49.35	5.5
W1GIF4_KLEPN	40.574	67.5
W1DM80_KLEPN7	32.833	7.5
UPI0001660147	31.843	11.9
W1H6G8_KLEPN	46.041	18.8
W1H2S6_KLEPN	30.292	11
W1E0Q5_KLEPN	59.283	60.7
W1DJ81_KLEPN	66.827	1.8
W1DIT5_KLEPN	32.09	19.1
W1DJ73_KLEPN	20.395	29.8
W1HU89_KLEPN	34.022	19
W1DYZ5_KLEPN	36.094	10.5
W1E189_KLEPN	56.416	34.3
W1GBN6_KLEPN	116.09	6
W1GYD5_KLEPN	46.038	17.3
W1GAJ8_KLEPN	64.242	49.4
W1H1E3_KLEPN	43.246	56.6
W1DY13_KLEPN	20.189	14.2
A6TC01_KLEP7	45.096	5.7
W1E2C3_KLEPN	37.101	39
A6TC88_KLEP7	71.862	21.6
A6TED1_KLEP	53.316	51.5
J2LPC2_KLEPN	28.444	53.2
W1HQ45_KLEPN	94.645	51.8
W1DIL7_KLEPN	10.359	18.7
W1H513_KLEPN	25.945	59.8
W1E8K6_KLEPN	111.98	2.9
W1E9W2_KLEPN	41.538	40.7
W1GZQ3_KLEPN	52.447	37.1
seq0003_CDS_B5055_0613	17.544	62.5
W1HLS5_KLEPN	9.3244	32.9
UPI0003527DF8	23.898	21
seq0038_CDS_B5055_4442	22.597	22.2

Appendix 4.0

Calculations theoretically estimating the number of transposon mutants to insert into every gene of *K. pneumoniae* B5055 (Genome size ~5.6 Mbp) and assumed average gene size of 1 kb.

Poisson based formula for calculating genome saturated with transposon insertions:

$$N = \ln(1 - P) \div \ln(1 - f)$$

Where *P* is the probability to disrupt a particular gene,
f is the fractional proportion of the genome represented by a single clone,
N is the number of clones required to get P

Formula adapted from:
J. Zilsel, P.H. Ma, J.T. Beatty, Derivation of a mathematical expression useful for the construction of complete genomic libraries, Gene, Volume 120, Issue 1,1992, pages 89-92,

Therefore	$f = 1/5600$			
	$\ln(1-f) = -0.000178587$			
P		$\ln(1-P)$	$N = \ln(1-P) \div \ln(1-f)$	Clones required (rounded up)
0.1		-0.10536	589.9662	590
0.5		-0.69315	3881.278	3881
0.9		-2.30259	12893.33	12893
0.99		-4.60517	25786.65	25787
0.999		-6.90776	38679.98	38680

Appendix 5.0
Strain IDs with nucleotide sequence identities details used for Figure 4.1.2.

Strain_ID	LupV sequence identity (%)	BamL sequence identity (%)
5197_8_8	99.529	98.854
9517_7_15	Stops	99.206
14936_3_39	99.435	99.03
Klebsiella_pneumoniae_subsp_pneumoniae_LZ	99.058	99.03
14936_2_30	99.058	98.942
5151_5_12	99.341	99.295
5299_7_4	99.341	99.295
9517_7_1	99.341	99.295
9517_7_16	Stops	99.206
5193_7_11	99.435	99.295
5193_5_7	99.529	99.471
9263_7_56	Stops	99.206
14936_3_24	99.341	99.295
14936_2_47	90.969	Stops
5235_7_8	99.529	99.383
14936_2_20	99.529	98.854
14893_8_89	99.341	99.118
5235_5_4	0	Contig
9517_7_5	99.341	99.295
10315_6_11	99.529	98.854
9263_7_77	Stops	99.206
5235_2_11	99.623	99.118
14936_3_66	99.529	98.942
14936_3_85	99.529	98.854
5197_2_8	99.341	99.206
14936_3_9	90.875	Stops
14936_3_22	99.341	99.295
5193_5_4	90.875	Stops
5299_7_7	99.435	93.04
14936_2_25	99.529	98.942
Klebsiella_pneumoniae_subsp_pneumoniae_MGH_78578	99.623	99.471
5235_2_4	90.781	Stops
14936_3_8	99.529	98.765
9878_1_8	99.812	99.206
9263_7_57	Stops	99.206
5151_2_7	91.345	89.075
14936_2_12	99.529	98.854
15277_3_61	99.341	99.206
5197_7_7	99.529	99.03
Klebsiella_pneumoniae_subsp_pneumoniae_KPNIH22	99.529	99.471
5151_3_1	92.192	89.427
5151_6_4	99.623	99.471
9878_1_12	99.812	99.206

Klebsiella_pneumoniae_subsp_pneumoniae_LCT-KP214	99.058	98.942
10315_6_62	99.529	98.854
5235_6_8	99.247	99.647
14936_2_62	90.969	Stops
14936_2_72	90.969	Stops
5150_2_10	99.529	98.854
5197_2_6	99.341	98.942
5235_7_10	99.529	99.383
14936_2_33	99.058	98.942
Klebsiella_pneumoniae_subsp_pneumoniae_KPNIH18	99.529	99.471
Klebsiella_pneumoniae_subsp_pneumoniae_HS11286	99.529	99.471
14936_3_51	99.529	99.471
5193_7_7	99.529	99.383
10356_5_82	99.812	99.206
14936_2_48	90.969	Stops
5235_6_11	99.529	99.295
5197_8_3	90.593	Stops
10315_6_1	99.623	99.03
14936_3_80	99.529	98.854
5193_2_1	99.529	99.471
5193_7_3	99.435	99.118
5235_3_8	99.812	99.647
9517_7_3	Stops	99.206
5299_1_7	99.623	99.03
5235_3_11	99.529	99.118
5151_3_9	Contig	99.295
5193_3_11	90.781	Stops
9263_7_61	Stops	99.206
14936_3_47	99.341	98.942
14936_3_13	99.529	98.854
5235_7_6	99.341	98.942
10315_6_95	99.812	99.206
5150_5_8	98.964	98.942
14936_2_42	99.058	98.942
5193_3_8	99.529	99.471
10315_6_21	99.529	98.854
Klebsiella_pneumoniae_subsp_pneumoniae_KPNIH7	99.529	99.471
14936_3_70	99.529	99.03
10315_6_22	99.529	99.471
14936_3_72	99.058	98.942
9263_7_55	Stops	99.206
5235_3_1	99.529	98.942
Klebsiella_pneumoniae_subsp_pneumoniae_KPNIH4	99.529	99.471
5235_7_9	99.529	99.383
5150_1_2	99.435	99.03
9263_7_48	Stops	99.206

14936_3_17	99.341	99.295
9263_7_70	Stops	99.206
14936_2_32	90.969	Stops
5193_8_9	90.781	Stops
14936_3_57	99.529	98.854
5193_8_3	99.529	99.471
14936_3_42	99.529	99.471
10315_6_41	99.341	99.118
Klebsiella_sp_MS_92-3	99.529	99.383
5150_3_2	100	Contig
9263_7_28	99.529	99.471
5193_7_10	90.499	Stops
Klebsiella_pneumoniae_subsp_pneumoniae_KPNIH12	99.529	99.471
5193_2_12	99.341	99.383
5197_8_2	99.341	99.118
5235_8_8	99.058	98.942
9878_1_2	99.529	98.854
14936_2_39	99.529	98.854
5235_3_4	99.435	98.942
9517_7_21	99.529	99.471
5193_6_12	99.153	99.383
14936_3_91	99.529	98.854
5193_6_4	99.529	99.383
5151_5_8	99.529	Stops
5193_1_6	99.247	99.647
Klebsiella_pneumoniae_1162281	99.247	99.383
5150_3_1	99.247	99.647
14936_2_63	90.969	Stops
14936_2_87	90.969	Stops
5299_7_1	99.153	99.383
5299_7_3	99.435	99.383
14936_3_86	99.529	98.854
10356_5_80	99.812	99.206
5197_8_6	99.058	98.942
9263_7_38	99.529	99.471
14936_2_83	90.969	Stops
10315_6_19	99.529	98.854
10315_6_78	99.529	98.854
5193_7_9	99.529	99.295
10315_6_42	99.435	99.471
5235_2_6	99.529	99.471
14936_3_84	99.341	99.206
5299_1_10	99.341	98.942
14936_3_28	99.341	99.118
5197_2_2	99.529	98.854
5193_8_7	91.345	89.251

5193_6_11	99.153	99.383
14936_3_83	99.529	99.118
9263_7_31	99.529	99.471
14936_3_10	90.969	Stops
14936_3_75	99.529	99.03
Klebsiella_sp_4_1_44FAA	99.623	99.206
5151_3_8	91.816	89.604
5193_8_2	99.435	99.471
14936_3_55	99.341	98.942
5197_7_1	99.623	99.383
14936_3_58	99.529	98.854
5151_6_12	91.533	89.075
5235_6_6	99.153	99.383
10315_6_87	99.812	99.206
14936_2_71	90.969	Stops
9517_7_26	Stops	99.206
5151_6_3	99.529	99.295
5193_1_12	99.341	99.295
10315_6_8	99.529	98.854
14936_2_38	99.529	98.854
14936_3_63	99.529	99.471
14936_2_73	99.058	98.942
5197_2_3	99.529	99.471
14936_2_57	90.969	Stops
5235_3_5	99.623	99.03
5151_5_6	91.627	89.163
14936_3_95	99.058	98.942
14936_2_8	99.058	98.942
5193_3_7	99.623	99.206
10356_5_77	99.529	98.854
14936_2_44	90.969	Stops
5299_1_2	99.153	98.942
9517_7_23	99.529	99.471
5151_6_10	99.529	99.118
5151_5_5	99.529	98.942
14936_3_62	99.058	98.942
5151_5_11	99.341	99.295
14936_2_52	90.969	Stops
14936_3_60	99.529	99.03
Klebsiella_pneumoniae_subsp_pneumoniae_KPNIH16	99.529	99.471
10315_6_27	99.341	99.118
14936_2_29	90.969	Stops
14936_2_53	90.969	Stops
14936_2_91	90.969	Stops
5193_6_8	99.153	99.118
14936_2_43	90.969	Stops

10315_6_74	99.341	99.295
10315_6_40	99.341	99.206
5235_2_8	99.906	99.912
5235_5_12	99.529	99.471
9263_7_69	Stops	99.206
14936_3_65	99.529	98.854
14936_3_40	99.529	99.471
5151_2_6	99.341	99.383
5150_1_4	99.529	99.383
5193_5_2	99.623	99.647
5150_2_6	99.435	0
14893_8_88	99.341	99.295
Klebsiella_pneumoniae_subsp_pneumoniae_KPNIH6	99.529	99.471
9878_1_11	99.812	99.206
5193_3_5	99.058	98.942
9263_7_24	99.529	99.471
9517_7_9	Stops	99.206
5151_5_2	99.247	99.383
5193_1_9	99.529	99.295
14936_3_49	99.341	98.942
10315_6_68	99.812	99.206
5150_5_7	99.247	98.854
5151_2_9	90.593	Stops
Klebs_pneumoniae_subsp_rhinoscleromatis_ATCC_13884	99.341	99.03
Klebsiella_pneumoniae_subsp_pneumoniae_WGLW3	99.058	98.942
9263_7_41	99.529	99.471
5193_8_5	99.529	99.471
14936_2_3	99.529	98.854
5151_2_1	99.247	99.647
14936_3_90	99.529	98.854
14936_3_44	99.341	99.295
14936_3_29	99.341	99.295
14936_2_59	99.058	98.942
Klebsiella_pneumoniae_subsp_pneumoniae_KPNIH11	99.529	99.471
5197_2_4	90.593	Stops
14936_3_14	99.529	98.854
14936_3_23	99.529	98.854
10315_6_82	99.529	98.854
10315_6_51	99.341	99.118
5193_6_6	99.623	99.383
5193_1_10	92.474	89.075
14936_2_34	90.969	Stops
14936_3_26	99.341	99.118
14936_2_2	99.529	98.854
14893_8_75	99.623	99.471
9263_7_68	Stops	99.206

9517_7_10	Stops	99.206
14936_2_4	99.435	99.295
5150_3_6	99.247	99.471
5151_6_1	99.529	99.471
5235_5_7	99.529	99.383
14936_3_30	99.341	99.295
9878_1_4	99.529	98.854
14936_2_21	99.058	98.942
9263_7_64	Stops	99.206
14936_2_54	90.969	Stops
10315_6_9	99.529	98.854
14936_2_16	99.058	98.942
14936_2_10	99.435	99.295
10315_6_34	99.435	99.471
5235_7_12	99.623	99.295
5193_5_10	99.529	98.765
14936_3_33	99.341	99.118
14936_2_67	90.969	Stops
14893_8_70	99.341	99.295
15277_3_60	90.969	Stops
5197_7_6	99.247	99.383
5235_5_8	99.623	99.647
5193_2_11	99.529	99.471
14936_3_48	99.341	98.942
10356_5_78	99.812	99.206
5197_2_12	90.593	Stops
14936_2_69	90.969	Stops
5151_6_6	99.529	99.206
5150_5_2	91.251	89.075
14936_3_50	99.529	99.206
Klebsiella_pneumoniae_subsp_pneumoniae_KPNIH5	99.529	99.471
5151_3_11	99.529	99.118
5235_5_10	0	Stops
5235_7_7	99.529	99.03
14936_3_34	99.529	98.854
5193_5_6	90.593	Stops
14936_3_93	99.529	99.471
5193_8_11	99.153	99.383
5151_5_7	99.812	99.295
5151_5_9	91.345	89.515
9263_7_44	99.529	99.471
5151_3_10	99.623	99.383
5299_7_6	99.529	99.383
9263_7_36	99.529	99.471
5235_1_11	99.435	99.03
9263_7_58	Stops	99.206

5151_2_3	99.529	99.471
14936_2_26	99.529	99.295
5193_5_11	99.529	98.765
14936_3_77	99.529	99.118
9263_7_39	99.529	99.471
5235_8_4	99.529	99.471
5299_7_2	99.529	98.854
14936_3_53	99.435	99.206
10315_6_75	99.812	99.206
14936_2_51	90.969	Stops
14893_8_69	90.875	Stops
5150_3_10	99.247	99.383
14936_2_6	99.529	98.854
5235_6_2	99.529	99.383
14936_3_76	99.529	99.471
14936_2_58	90.969	Stops
5193_1_7	99.623	99.383
5193_7_8	90.31	Stops
Klebsiella_pneumoniae_subsp_pneumoniae_KPNIH20	99.529	99.471
10315_6_5	99.529	98.854
5150_2_8	99.341	99.295
5150_2_2	90.687	89.515
9263_7_67	Stops	99.206
5197_8_12	99.718	99.295
14893_8_80	90.875	Stops
5150_2_5	99.623	99.03
5235_8_6	99.529	99.295
9263_7_43	99.529	99.471
14936_2_78	90.969	Stops
14936_2_27	99.529	99.295
5150_2_12	99.341	Contig
14893_8_72	99.341	99.206
5150_5_5	99.341	99.471
5150_2_7	99.247	99.295
5235_2_12	99.435	99.471
5151_6_5	99.623	99.118
14936_2_55	90.969	Stops
5193_3_1	90.969	Stops
5151_2_11	99.718	98.942
5193_3_10	99.529	99.295
14936_3_18	99.529	99.471
5151_6_7	99.529	98.942
9517_7_20	99.153	92.24
5193_3_6	99.058	98.942
14936_3_52	99.529	99.118
14936_3_16	99.247	99.295

5193_2_8	99.435	99.471
5193_1_8	Stops	99.559
5193_3_4	Stops	97.061
5197_7_3	99.247	99.559
14936_3_36	99.529	99.471
9517_7_4	99.341	99.295
5235_1_4	91.816	89.604
5193_8_6	99.529	99.471
5235_5_11	99.623	99.295
9878_1_3	99.529	98.854
14936_3_69	99.341	99.03
5193_5_1	99.247	99.295
14936_2_41	90.969	Stops
14936_2_64	90.969	Stops
14936_3_19	99.529	99.471
5193_2_4	90.593	Stops
10315_6_58	99.529	99.383
Klebsiella_pneumoniae_NTUH-K2044	99.247	99.647
14893_8_84	99.623	99.471
5151_2_8	99.341	99.471
5299_1_9	99.529	98.942
5151_5_3	99.529	99.118
10315_6_61	99.623	99.383
5235_1_1	99.435	99.471
5235_7_2	99.529	99.383
14936_3_7	90.969	Stops
14936_2_84	99.435	99.471
9263_7_50	Stops	99.206
5151_2_10	99.247	99.383
9263_7_29	99.529	99.471
10315_6_57	99.529	98.854
9517_7_28	99.529	98.765
10315_6_49	99.812	99.206
5150_5_4	99.058	98.942
14936_2_19	99.529	98.854
5193_2_9	99.529	99.471
10315_6_66	99.529	98.854
9263_7_49	Stops	99.206
14936_2_92	90.969	Stops
5151_5_10	92.192	89.515
5193_8_10	99.529	99.471
9263_7_60	Stops	99.206
9517_7_14	Stops	99.206
5235_1_12	99.247	98.942
5151_6_8	99.529	99.118
10356_5_86	99.435	99.295

5299_1_12	92.286	89.515
5193_8_4	99.341	99.118
5151_3_6	99.529	99.118
5150_3_7	99.623	98.942
5299_1_8	99.435	99.295
9263_7_76	Stops	99.206
Klebsiella_pneumoniae_JH1	99.435	99.206
5235_2_5	99.247	98.942
5193_1_1	Contig	99.295
10315_6_2	99.529	98.854
14936_3_71	99.153	99.559
10315_6_17	99.529	98.854
14936_2_74	90.969	Stops
9263_7_32	99.529	99.471
14893_8_90	99.341	99.118
Klebsiella_variicola_At-22	91.345	89.075
9263_7_63	Stops	99.206
5235_8_5	99.435	99.295
5151_5_4	99.529	98.942
5150_3_11	99.623	99.383
14936_2_65	90.969	Stops
5197_8_7	99.529	99.471
9517_7_25	Stops	99.206
5299_7_5	99.623	98.942
14936_3_68	99.058	98.942
10315_6_13	99.529	98.854
14893_8_86	99.341	99.206
9263_7_73	Stops	99.206
14893_8_87	99.623	99.471
5235_3_2	99.341	99.471
5235_7_11	99.529	99.383
5150_3_3	99.623	99.383
14936_2_81	90.969	Stops
14936_3_12	99.341	99.324
5151_3_5	99.247	99.03
9263_7_40	99.529	99.471
5235_8_2	99.247	99.559
14936_2_9	99.529	98.854
14936_2_61	90.969	Stops
14936_2_90	90.969	Stops
5235_3_9	99.718	99.206
10315_6_52	99.623	99.03
10315_6_96	99.341	99.295
5197_8_4	99.435	99.03
14936_2_11	99.529	99.383
5193_1_4	99.247	99.647

5150_2_4	99.341	99.295
10315_6_6	99.529	98.854
14936_3_11	90.875	Stops
5150_1_5	99.623	99.295
14936_2_60	90.969	Stops
14936_2_28	99.529	98.854
14893_8_76	90.875	Stops
10315_6_39	99.529	99.383
5151_5_1	99.623	99.383
9263_7_45	99.529	99.471
14936_2_82	90.969	Stops
5151_2_5	99.058	98.942
9263_7_33	99.529	99.471
Klebsiella_pneumoniae_1191100241	99.623	99.03
5151_6_2	99.529	99.03
5235_3_6	99.435	99.295
14936_2_75	90.969	Stops
5197_2_1	99.435	99.471
10315_6_70	99.341	99.295
5235_5_2	99.718	99.295
9517_7_17	99.529	99.471
5235_2_1	99.341	99.295
5197_2_9	99.623	99.03
Klebsiella_pneumoniae_KCTC_2242	99.341	99.383
5193_8_8	99.529	98.854
5151_6_11	99.529	99.118
14936_2_7	99.529	98.854
14936_3_2	99.341	99.118
14893_8_81	99.529	98.854
5150_5_11	99.529	98.765
14936_3_35	99.529	99.559
10315_6_28	99.529	98.854
5193_6_1	90.875	Stops
14936_2_86	90.969	Stops
10315_6_89	99.812	99.206
10315_6_80	99.529	98.854
9878_1_5	99.529	98.854
Klebsiella_pneumoniae_subsp_pneumoniae_1084	99.247	99.647
14936_2_35	99.529	98.854
5299_7_10	99.435	99.471
Klebsiella_pneumoniae_subsp_pneumoniae_KPNIH21	99.529	99.471
10315_6_10	99.529	98.854
9263_7_71	Stops	99.206
5299_1_4	99.529	98.942
10356_5_76	99.812	99.206
14936_2_76	90.969	Stops

14936_3_54	99.341	98.942
5235_7_4	99.529	99.118
14936_2_31	99.058	98.942
5193_6_3	99.153	99.118
10315_6_71	99.812	99.206
5235_5_5	99.529	99.471
5193_7_5	99.341	99.295
10356_5_85	99.812	99.206
5235_5_6	99.529	98.765
14936_3_37	99.529	99.471
9263_7_27	99.529	99.471
5193_3_12	99.623	99.383
14936_3_94	99.058	98.942
5150_5_9	99.341	98.942
5193_8_1	99.529	99.471
5150_3_8	91.345	89.075
10315_6_77	99.529	98.854
5235_6_12	99.623	93.122
10315_6_92	99.812	99.206
14936_2_50	90.969	Stops
10356_5_79	99.812	99.206
10315_6_65	99.341	98.942
9517_7_13	Stops	99.206
9263_7_72	Stops	99.206
5193_3_9	99.435	99.383
14936_2_93	90.969	Stops
9263_7_74	Stops	99.206
14936_3_27	99.341	99.118
14936_3_45	99.341	98.942
9263_7_54	Stops	99.206
14936_3_21	99.529	98.854
14936_2_36	99.529	98.854
5193_6_10	99.812	99.912
5193_7_2	99.529	99.471
14893_8_92	99.341	99.118
9263_7_35	99.529	99.471
5150_5_10	99.812	99.206
14936_3_64	99.529	98.854
9878_1_6	99.529	98.854
5299_1_3	99.529	99.206
14893_8_85	99.529	99.471
5150_1_11	99.435	99.03
9263_7_46	99.529	99.471
Klebsiella_sp_1_1_55	91.345	89.251
14893_8_74	99.623	99.471
14936_2_66	99.529	98.942

14936_2_95	99.435	99.471
9517_7_12	Stops	99.206
14936_3_31	99.435	99.03
10315_6_88	99.529	99.471
10356_5_87	99.529	Stops
9263_7_52	Stops	99.206
10315_6_45	99.529	98.854
14893_8_71	99.341	99.295
Klebsiella_pneumoniae_subsp_pneumoniae_KPNIH17	99.529	99.471
5150_5_12	Contig	Contig
5299_1_1	99.529	99.559
10315_6_60	99.058	98.942
5235_3_7	99.529	99.383
5193_2_5	99.529	99.471
5197_8_1	99.623	99.03
14936_3_79	99.529	98.854
5150_1_9	99.529	Contig
10315_6_59	99.529	99.471
14936_3_56	99.529	98.854
10315_6_31	99.529	99.471
9263_7_51	Stops	99.206
14936_3_6	99.341	99.118
10315_6_7	99.529	98.854
14936_2_68	90.969	Stops
10315_6_38	99.058	98.942
14936_2_88	99.058	98.942
14893_8_91	90.969	Stops
10315_6_20	99.529	98.854
5193_2_7	99.435	99.03
10315_6_53	99.341	99.295
10315_6_79	99.529	98.854
14936_3_61	99.529	99.206
5235_6_7	99.435	99.206
14936_2_79	90.969	Stops
5235_1_9	99.435	99.295
14893_8_82	99.341	99.295
14893_8_77	99.341	99.206
5150_5_1	99.435	99.03
14936_2_14	99.529	98.854
9263_7_53	Stops	99.206
10315_6_44	99.529	99.03
14936_2_46	90.969	Stops
14936_3_43	99.341	99.295
5150_2_1	90.687	Stops
14936_3_3	99.529	98.765
5235_1_7	99.529	99.383

14936_2_89	99.058	98.942
5197_8_11	99.906	99.912
5235_1_5	99.529	98.942
9263_7_30	99.529	99.471
10315_6_76	99.529	99.471
Klebsiella_pneumoniae_342	91.157	88.899
5235_8_7	99.153	99.383
5193_2_3	90.687	Stops
14936_2_49	90.969	Stops
14936_2_24	99.529	98.942
5299_7_8	99.529	98.854
5299_1_11	99.247	99.383
Klebsiella_pneumoniae_subsp_pneumoniae_KPNIH14	99.529	99.471
9263_7_75	Stops	99.206
5151_2_2	99.247	99.647
9517_7_18	99.529	98.854
5193_2_2	99.529	99.471
10315_6_67	99.812	99.206
14936_3_92	99.529	98.854
14936_3_89	99.529	98.854
9517_7_27	Stops	99.206
Klebsiella_pneumoniae_subsp_pneumoniae_KPNIH19	99.529	99.471
9517_7_24	Stops	99.206
9517_7_7	99.529	99.471
5151_6_9	99.623	99.206
5235_1_8	99.529	98.854
9263_7_25	99.529	99.471
10315_6_93	99.812	99.206
5150_2_11	99.529	98.854
5193_1_5	99.529	99.295
14936_2_17	99.529	98.854
14936_2_5	99.529	98.854
5150_5_3	99.247	99.118
5197_7_8	99.247	99.647
Klebsiella_pneumoniae_subsp_pneumoniae_WGLW1	99.058	98.942
5197_8_5	90.687	Stops
9263_7_42	99.529	99.471
10315_6_23	99.435	99.206
5235_7_3	99.529	99.118
14936_3_74	99.529	98.854
9517_7_22	Stops	99.206
Klebsiella_pneumoniae_subsp_pneumoniae_KPNIH8	99.529	99.471
5151_2_4	99.247	99.647
5235_5_3	Contig	99.118
14936_3_1	90.969	Stops
Klebsiella_pneumoniae_subsp_pneumoniae_KPNIH2	99.529	99.471

9517_7_8	99.529	99.471
14893_8_83	99.529	98.854
5235_6_5	99.153	99.383
5235_7_1	99.529	99.383
14936_2_94	99.435	99.471
5197_2_10	99.529	99.471
5235_1_6	99.623	99.295
5197_2_7	99.529	99.471
5235_1_3	99.247	99.206
5299_7_9	99.341	99.824
14936_2_22	99.058	98.942
5235_1_2	99.529	99.383
14936_2_18	99.058	98.942
14893_8_73	99.341	99.295
14936_2_56	90.969	Stops
Klebsiella_pneumoniae_subsp_pneumoniae_KPNIH10	99.529	99.471
5235_2_7	99.529	99.03
9263_7_37	99.529	99.471
5235_8_10	91.533	89.251
5151_3_7	99.623	99.03
9263_7_47	99.529	99.471
14936_3_5	90.969	Stops
5193_8_12	0	99.912
5299_1_5	91.902	Stops
10315_6_24	99.529	98.854
14936_3_41	99.529	98.854
14936_3_59	99.341	99.118
14936_3_81	99.529	98.854
5150_1_7	91.345	89.075
14936_2_1	99.153	99.559
14936_3_25	99.529	98.854
9263_7_65	Stops	99.206
14936_2_13	99.435	99.295
5193_2_10	90.593	Stops
10315_6_54	99.341	99.295
5235_5_1	0	Stops
10315_6_29	99.529	98.854
5193_6_2	99.906	99.559
5235_6_1	99.529	98.942
5193_6_7	Stops	99.118
14936_3_15	99.529	98.854
14936_3_4	99.341	99.118
10315_6_18	99.529	98.854
14936_3_46	99.341	98.942
9517_7_19	99.529	98.854
5197_8_9	99.435	99.295

9263_7_66	Stops	99.206
14936_3_88	99.623	99.03
5150_1_10	99.529	Contig
5197_7_2	99.058	99.295
5197_7_4	99.529	98.854
5197_7_5	92.286	89.075
10315_6_69	99.341	99.295
5193_7_1	99.529	98.854
14936_2_40	99.058	98.942
5150_3_9	99.529	99.295
10315_6_73	99.529	98.854
10315_6_43	99.529	98.854
5151_2_12	91.063	88.987
14936_2_70	90.969	Stops
14936_3_38	99.529	99.471
5235_6_10	99.529	99.471
5193_5_9	99.529	98.942
14936_2_45	90.969	Stops
5150_1_3	99.529	99.206
9517_7_11	Stops	99.206
5150_1_8	91.722	89.075
5150_2_3	99.529	99.559
14936_2_85	99.435	99.471
9878_1_9	99.812	99.206
Klebsiella_pneumoniae_subsp_pneumoniae_KPNIH1	99.529	99.471
9263_7_62	Stops	99.206
14936_3_67	99.435	99.118
14936_3_73	99.341	99.295
10356_5_81	99.341	99.295
5235_3_12	99.623	99.295
5151_3_2	99.529	99.647
5235_2_2	99.435	97.316
10315_6_3	99.529	98.854
5197_2_11	99.247	99.295
14936_3_32	99.435	99.03
5193_6_5	99.529	98.942
5193_5_8	98.964	99.03
14936_3_82	99.529	99.118
5193_1_11	99.529	99.295
Klebsiella_pneumoniae_subsp_pneumoniae_JC2877	99.247	Stops
Klebsiella_pneumoniae_subsp_pneumoniae_KPNIH9	99.529	99.471
5193_6_9	99.529	98.942
5193_7_6	99.623	99.383
5235_8_9	99.529	99.03
14936_3_20	99.341	99.295
14936_2_37	90.969	Stops

10315_6_72		99.341	99.206
5193_2_6		99.341	99.383
5193_5_12		99.529	99.383
9517_7_2	Stops		99.206
5235_1_10		99.341	98.942
5235_6_9		99.341	99.471
5235_2_3		99.718	98.942
5235_6_3		99.435	99.295
14936_3_87		99.529	98.854
14936_2_77		90.969	Stops
5197_2_5		90.593	Stops
14936_2_80		99.058	98.942
14936_2_23		99.529	98.942
Klebsiella_pneumoniae_subsp_pneumoniae_KPNIH23		99.529	99.471
5235_8_3		99.623	99.03
5150_3_5	Contig		99.295
9263_7_59	Stops		99.206
5235_7_5		99.341	98.942