A ROLE FOR PROTEOBACTERIAL MAMMALIAN CELL ENTRY DOMAINS IN PHOSPHOLIPID TRAFFICKING AND INFECTION

by

Georgia Louise Isom

A thesis submitted to the University of Birmingham for the degree of DOCTOR OF PHILOSOPHY

Institute of Microbiology and Infection
College of Medical and Dental Sciences
University of Birmingham
November 2016

ABSTRACT

Mammalian cell entry (MCE) domains are so called due to the reported ability of an *Escherichia coli* strain harbouring the *mce1* gene from *Mycobacterium tuberculosis* to invade mammalian cells. Bioinformatic analyses presented here demonstrated that proteins containing a single MCE domain are widespread in bacteria and that proteins containing multiple MCE domains are specific to and have evolved within Proteobacteria. Gene neighbourhood analyses revealed that MCE domain containing proteins are components of transporters and that multi MCE domain containing proteins constitute a novel type of transporter. *E. coli* was shown to harbour three MCE proteins: the single MCE domain protein MlaD and two multi-domain proteins PqiB and YebT. All three proteins were shown to locate to the inner membrane and bind phospholipids. Phenotypic studies revealed that their functions overlap but are distinct. Infection studies with *Salmonella* showed that the proteins are important for systemic infection but are not required for mammalian cell entry. Phospholipid growth experiments with *Salmonella* demonstrated that they are important for phospholipid uptake. These findings suggest that MCE domain containing proteins in Proteobacteria are not directly involved in mammalian cell entry and instead play a role in other aspects of mammalian infection related to phospholipid trafficking.

ACKNOWLEDGMENTS

First I would like to thank my supervisor, Professor Ian Henderson, for his supervision, guidance and advice during the last four years. He has encouraged me to think strategically and has helped me make contacts that will help my future career in science. I would also like to say a very big thank you to Professor Jeff Cole. He has helped me enormously with my thesis and upcoming papers, the process of which has helped me improve my scientific writing. He has also given me advice for my future and has given me confidence in my ability as a scientist, for which I will always be grateful.

For help with all the bioinformatics aspects of the project, I owe a huge thanks to Nathaniel Davies. I would also like to thank him for his continuous support over the last three years in both my work life and our personal lives. I also owe thanks to many members of the T101 lab. I would like to thank Dr Jack Bryant, who in the last year of my PhD has been an excellent work partner and friend and has been a great help with sounding out ideas and overcoming any problems. Thanks must also go to Dr Mohammed Jamshad who taught me everything I know to do with proteins and has given a lot of his time to helping me, for which I am very grateful. I also owe thanks to Dr Tim Wells, Dr Amanda Rossiter and Dr Yanina Sevastsyanovich for teaching me new techniques and helping me with experiments. Thanks are also owed to Tim for his advice with reports and my thesis and always being keen to help. I am grateful to Emily Goodall, Chris Icke and Emma Sheehan for being great friends both in and outside work, and helping with experiments over the last year, and to Karl Dunne for sounding out ideas. I am also grateful for members of the "magic rainbow five" (Jess Rooke, Irene Beriotto, Ash Robinson and Tim Wells) for their friendship and support, and for making me very welcome when I first joined the lab, and also to Jess for putting up with living with me for a year. Outside T101, I owe thanks to members of the Lovering lab, particular Richard Meek, for all their help with crystallography. I'd also like to thank Dr James Haycocks for being an excellent housemate during the writing-up period and cooking for me.

Finally I would like to thank my family, in particular my Mum, my sister Ellie and brother Joe for being the core three people supporting me throughout my life, and Andy, who together with my mum has done a lot to help me. I'd also like to thank my Dad, grandparents, friends and the rest of my family for their continuing support.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

GENERAL INTRODUCTION

1.1. Introduction

Mammalian cell entry (MCE) domains are a type of protein domain that were first identified in *Mycobacterium tuberculosis*, where insertion of a gene from *M. tuberculosis* into non-pathogenic *Escherichia coli* (*E. coli*) was shown to allow the bacterium to enter and survive in mammalian cells (Arruda et al., 1993). It is because of this acquired ability that this gene was named *mce1A*, for mammalian cell entry.

MCE domains are widespread in bacteria but their function has primarily been studied in Actinobacteria. Most literature suggests that MCE domain containing proteins are important for infection and are often implicated in lipid transportation and maintenance of the bacterial cell surface. The few studies that examine MCE domains in Gram-negative bacteria generally report a link to lipid trafficking, particularly of phospholipids, which are essential components of cell membranes.

In this thesis, the evolution and distribution of MCE domains is investigated through a thorough bioinformatic analysis. The function of MCE domains is then studied in *E. coli* in order to understand more about their function in Proteobacteria, with particular focus on their role in phospholipid trafficking. Furthermore, *Salmonella* infection is used as a model to investigate whether MCE domain containing proteins are important for infection in Proteobacteria, given the known importance of MCE domains in Actinobacterial infection.

This introduction chapter summarises what is currently known about the field, and lays the foundations for an understanding of why the work presented in this thesis is important. The cell envelope of Gram-negative bacteria and the pathways involved in assembly and maintenance of the outer membrane are summarised, with particular focus on fatty acid and phospholipid metabolism. The *Salmonella* murine infection model is explained, and the existing literature on MCE domains is reviewed.

1.2. The cell envelope of Gram negative bacteria

The Gram-negative bacterial cell envelope is a complex entity that is essential for the protection of bacterial cells against the external environment. The envelope consists of three major structures: the outer membrane, peptidoglycan and the inner membrane (Figure 1.1). The outer membrane is an asymmetrical lipid bilayer that consists of a lipopolysaccharide (LPS) outer leaflet and phospholipid inner leaflet. The outer membrane contains many integral outer membrane proteins (OMPs) and lipoproteins that are anchored to phospholipids and either face into the periplasm or are on the cell surface. In contrast to the outer membrane, the inner (cytoplasmic) membrane is a symmetrical phospholipid bilayer. The inner membrane contains integral inner membrane proteins (IMPs) and lipoproteins that anchor to the outer leaflet of the inner membrane. Between the inner and outer membranes lies the periplasm, which contains a polymer essential for maintaining cell structure and rigidity: the peptidoglycan cell wall.

The major difference between Gram-negative and Gram-positive bacteria is that Gram-positive bacteria lack the outer membrane and have a much thicker layer of peptidoglycan, which is surface exposed. The additional membrane in Gram-negative bacteria acts as a selectively permeable barrier that allows influx of nutrients while preventing entry of harmful compounds such as antibiotics, detergents and antimicrobial peptides (Nikaido, 2003).

1.2.1. Integral inner membrane proteins and the Sec translocon

The inner membrane of Gram-negative bacteria contains many important inner membrane proteins involved in transport, energy production, lipid biosynthesis and protein secretion (Silhavy et al., 2010). One of the most important complexes of the inner membrane is SecYEG, which is responsible for the insertion or translocation of nearly all cell envelope and secreted proteins across the inner membrane. Most IMPs are targeted to the inner membrane by the first transmembrane (TM) region. Other cell envelope proteins (lipoproteins, OMPs etc) are targeted by specific N-terminal signal sequences. The methods

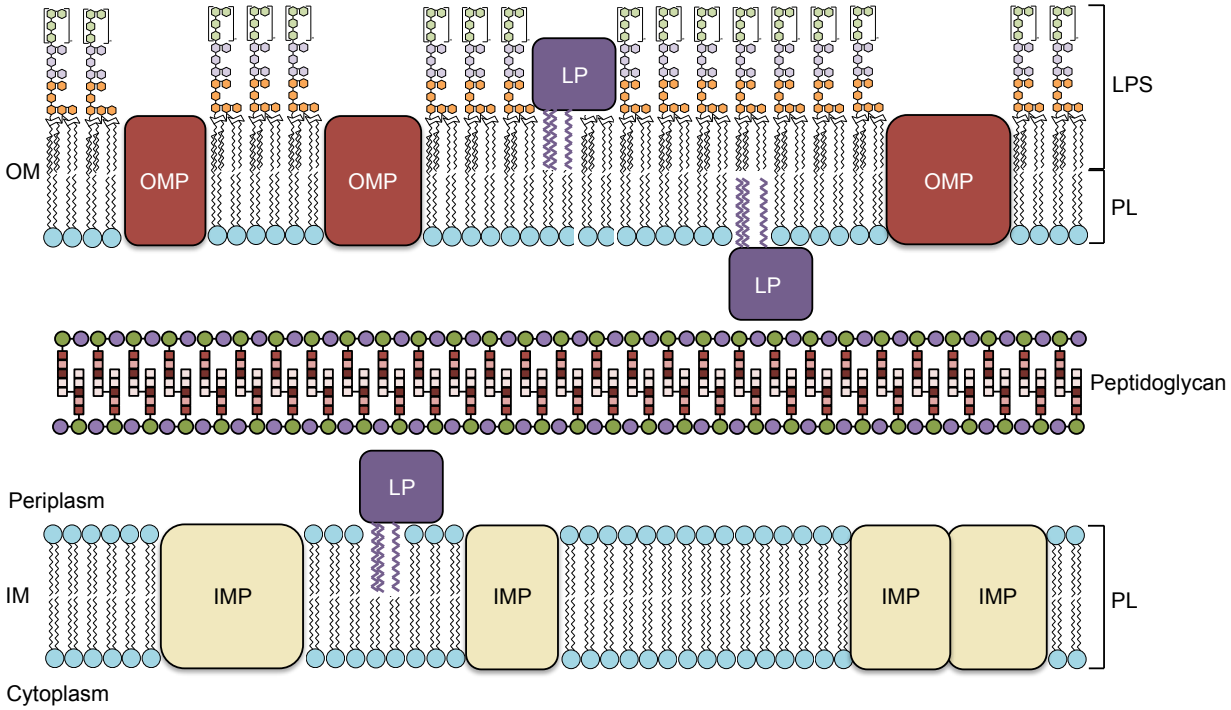


Figure 1.1. The structure of cell envelope of Gram-negative bacteria. The inner membrane (IM) consists of a phospholipid bilayer and contains inner membrane proteins (IMPs). The outer membrane consists of an inner leaflet of phospholipids and an outer layer of lipopolysaccharide (LPS) and contains outer membrane proteins (OMPs). The membranes enclose the periplasm that contains the peptidoglycan cell wall. Lipoproteins (LP) interact with both the outer leaflet of the inner membrane and the inner leaflet of the outer membrane and face into the periplasm. Alternatively they can locate to the cell surface.

of protein translocation and incorporation into the inner membrane are reviewed by Luirink et al. (2012). Protein translation occurs in the cytoplasm, at the ribosome. Beyond this, there are two possible routes to the Sec translocon: the co-translational signal recognition particle (SRP) – dependent pathway or the post-translational (SecB - dependent) pathway (Figure 1.2). The SRP pathway usually targets the more hydrophobic TM regions of IMPs (Luirink and Sinning, 2004). In this co-translational pathway, the SRP binds to its target as it emerges from the ribosome exit tunnel (Figure 1.2A). Binding of SRP to the nascent peptide chain results in elongation arrest and the ribosome-SRP-peptide complex passes towards the inner membrane to bind the FtsY receptor. Upon binding to FtsY, SRP is released and elongation is continued (Luirink et al., 2012; Miller et al., 1994). At the inner membrane, the ribosome interacts with the SecYEG/YidC complex (Nissen et al., 2000). YidC is important for insertion, folding and assembly of IMPs, and can act both with and independently from the Sec translocon (Scotti et al., 2000). The Sec complex consists of the core SecYEG, through which the TM regions of IMPs are inserted into the inner membrane. The SecDF-YajC complex that transiently associates with SecYEG may also have a role in IMP insertion (Duong and Wickner, 1997).

In the post-translational pathway, the less hydrophobic signal sequences of non-IMP proteins are targeted via the SecB chaperone pathway (Luirink et al., 2005). Although this pathway is known as “post-translational”, more recent evidence suggests the pathway targets proteins during translation (Huber et al., 2011; Huber et al., 2016). This pathway is the usual method of translocation of lipoproteins, periplasmic proteins, OMPs and secreted proteins across the inner membrane. SecB chaperones peptides to the SecYEG complex independently of the ribosome (Figure 1.2B) (Ullers et al., 2004). Translocation across the membrane requires the ATPase SecA that inserts itself into the SecYEG complex (Brundage et al., 1990). In particular, it is thought that SecA is required for the translocation of large periplasmic loops, which can also occasionally be found in IMPs (Andersson and von Heijne, 1993; Luirink et al., 2012). The accessory complex SecDF-YajC also aids in the post-translational pathway by stabilising the inserted state of SecA (Mori and Ito, 2001). Once

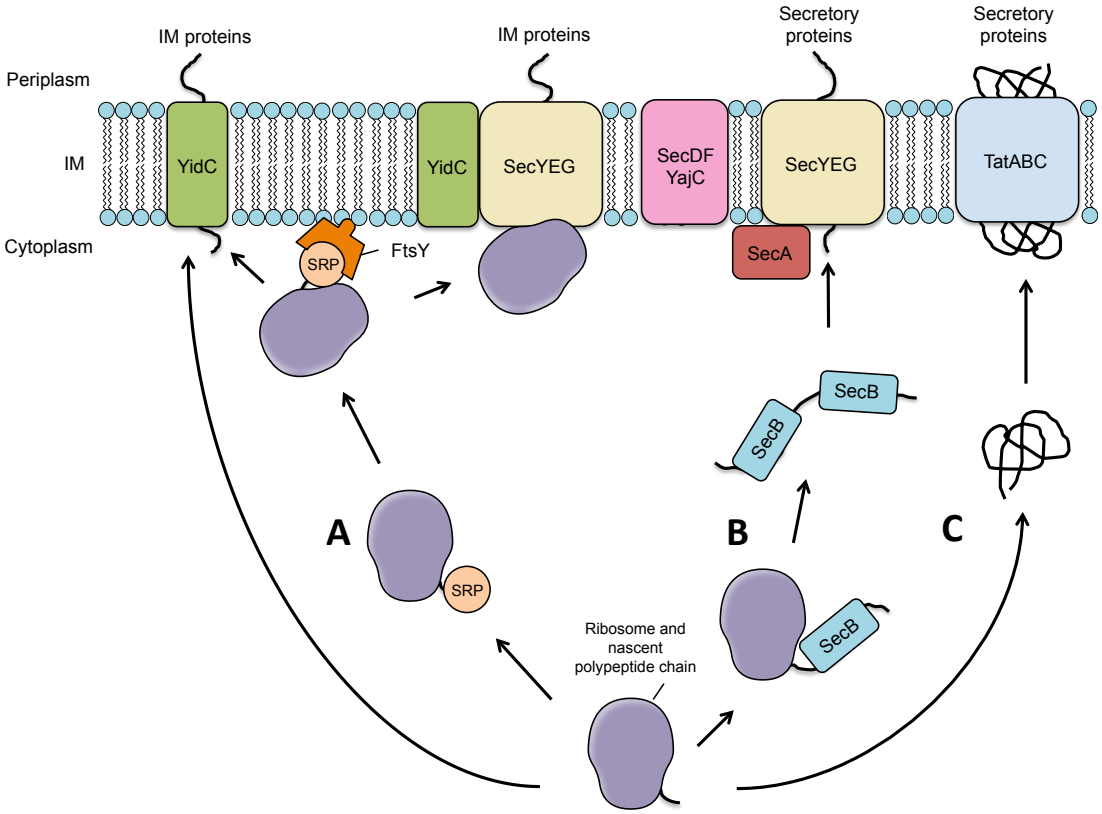


Figure 1.2. Pathways for the incorporation or translocation of proteins across the inner membrane. **A.** Pathways for the incorporation of inner membrane proteins (IMPs) into the inner membrane. Proteins targeted to the inner membrane via the SRP co-translational pathway are bound by SRP during translation. The ribosome-nascent chain-SRP complex interacts with the FtsY, releasing the ribosome and nascent chain from SRP and allowing further translation through SecYEG/YidC at the inner membrane. Some IMPs are translocated via YidC alone independently of the Sec translocon and sometimes completely independently of SRP. **B.** The SecB-dependent post-translational pathway for translocation of lipoproteins, periplasmic, outer membrane and secreted proteins. SecB chaperones polypeptides with a signal sequence to the SecYEG, where the protein is translocated with the aid of SecA. In the periplasm the signal sequence is removed. The SecDFYajC complex aids the Sec translocon in both the co- and post-translational pathways. **C.** The TAT pathway for the translocation of proteins that are folded in the cytoplasm, with major components TatA, B and C.

the protein has been translocated through SecYEG, signal peptidase I (or II for lipoproteins) cleaves it from its signal sequence and other cell envelope proteins transport it to its final destination (Dev and Ray, 1984; Paetzel et al., 2002).

SecYEG/YidC are not the only protein complexes involved in translocation across the inner membrane. The twin arginine translocation (TAT) system is important for translocation of proteins that are folded in the cytoplasm (Figure 1.2C) (Stanley et al., 2001). In particular, proteins that fold with cytoplasmic cofactors are commonly translocated via this pathway (Berks et al., 2003). Proteins translocated via the TAT system have a specific N-terminal signal sequence that is defined by two consecutive arginine residues, giving rise to the pathway name (Sargent et al., 1998). Generally, it is the lower hydrophobicity of the N-terminal signal sequence that may prevent TAT proteins going via the Sec pathway (Cristobal et al., 1999). The core proteins in the TAT pathway are TatA, B and C. In their inactive state, TatB and C form a complex in the inner membrane, while TatA proteins are dispersed in the membrane (Palmer and Berks, 2012). Upon binding of a TAT signal sequence to the TatBC complex, TatA is recruited to form an active TatABC complex (Palmer and Berks, 2012). The targeted protein passes through the membrane via TatA whilst the N-terminal signal sequence remains bound to TatBC and is cleaved by a signal peptidase (Palmer and Berks, 2012). Other proteins involved in the TAT system include TatD and TatE. TatD is a quality control protein that is involved in the detection of misfolded proteins (Matos et al., 2009). TatE is homolgous to TatA and might be recruited to either the TatBC or TatABC complex for the translocation of particularly large proteins (Baglieri et al., 2012).

1.2.2. Integral outer membrane proteins

Integral outer membrane proteins are a subset of proteins in Gram-negative bacteria that form β-barrels in the outer membrane. One of the most abundant types of OMP are porins, which allow passive diffusion of small hydrophilic molecules across the outer membrane (Silhavy et al., 2010). Two of the most abundant porins, OmpC and OmpF, are used for

general diffusion of molecules across the outer membrane (including amino acids, sugars etc) and together they have around 250,000 copies per cell (Cowan et al., 1992). Other OMPs, such as TolC, are required for specific functions such as efflux or secretion. In terms of efflux, TolC forms the outer membrane component of the multiple efflux pumps, including the major AcrAB-TolC system (Fralick, 1996). TolC is also involved in type I secretion systems (T1SS), in which it forms the outer membrane factor of multiple secretion systems (Durand et al., 2009).

As with most proteins destined for the periplasm or outer membrane, OMPs are synthesised in the cytoplasm with an N-terminal signal peptide and translocated across the inner membrane via the SecYEG complex (Figure 1.3A) (de Keyzer et al., 2003). Periplasmic chaperones Skp and SurA are responsible for transporting proteins across the periplasm to the outer membrane (Bos and Tommassen, 2004). Although proteins remain largely unfolded at this stage, the periplasmic enzyme DsbA can form disulfide bonds between cysteine residues before the transported protein reaches the outer membrane (Eppens et al., 1997). For insertion, OMPs are transported to the β-barrel assembly machinery (BAM) complex in the outer membrane. In *E. coli,* the BAM complex consists of the OMP BamA, which forms a complex with four lipoproteins (BamB-E), where both BamA and BamD are essential (Malinverni et al., 2006; Wu et al., 2005). In addition to a large β-barrel, there are five polypeptide transport-associated (POTRA) domains in BamA that might directly interact with OMPs in the periplasm (Sanchez-Pulido et al., 2003).

The entire process of OMP synthesis and transport is tightly regulated to prevent toxic accumulation of unfolded peptides in the cell. DegP is the major protease responsible for degrading misfolded proteins in the periplasm, but can also act as a chaperone under certain conditions (Spiess et al., 1999). An accumulation of OMPs in the cytoplasm or periplasm can be toxic for the cell. OMP accumulation in the periplasm induces a σ^E stress response that in turn increases the production of periplasmic chaperones, proteases and the BAM complex (Ruiz et al., 2008).

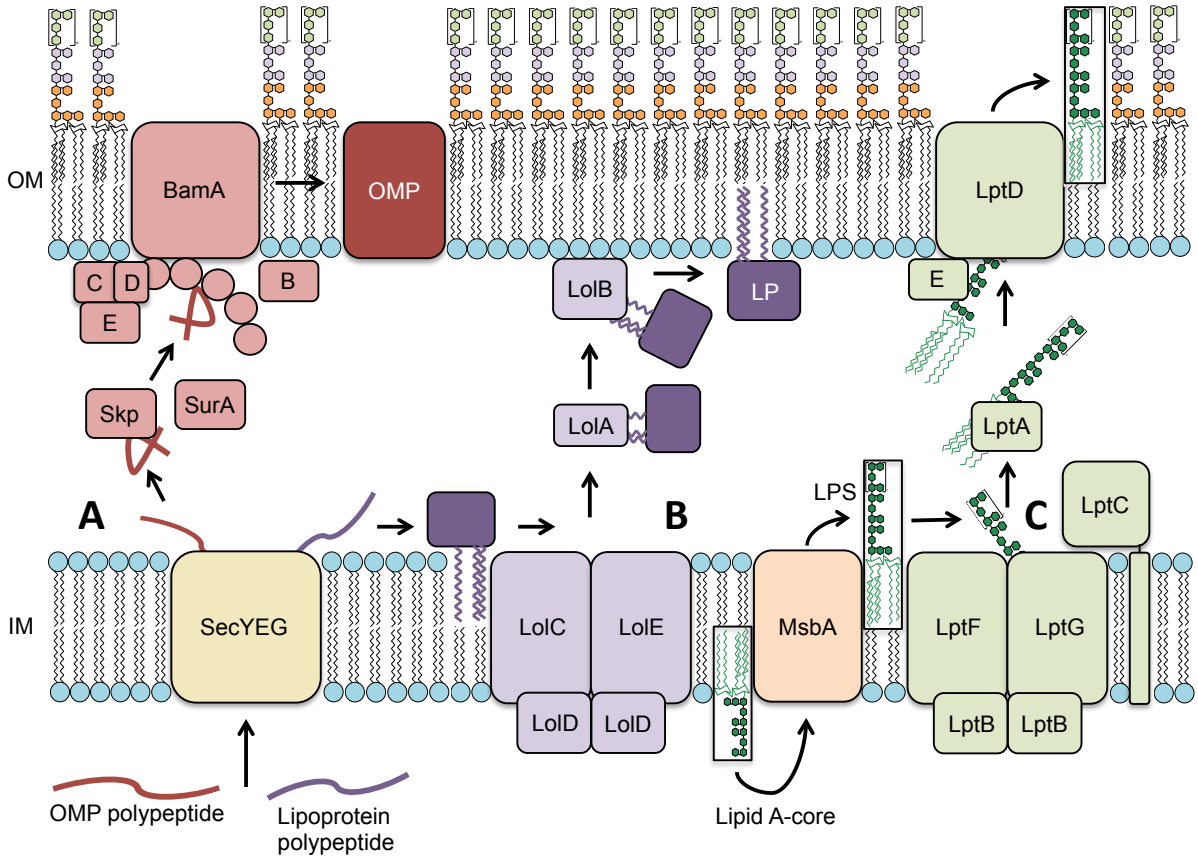


Figure 1.3. The pathways for delivery of outer membrane proteins (OMP), lipoproteins (LP) and lipopolysaccharide (LPS) to the outer membrane. **A.** OMPs are translocated across the inner membrane and transported across the periplasm by periplasmic chaperones (Skp/SurA) to the Bam complex. OMPs are inserted through BamA to form a β-barrel in the outer membrane. **B.** Lipoproteins (LP) destined for the outer membrane are transported from the outer leaflet of the inner membrane to the inner leaflet of the outer membrane via the Lol pathway. **C.** The Lipid A-core is synthesised in the periplasm and flipped across the inner membrane by MsbA. The O-antigen component is made independently and ligated to the Lipid A core at the periplasmic leaflet of the inner membrane. The intact LPS molecule is transported to the outer leaflet of the outer membrane via that Lpt pathway.

1.2.3. Lipoproteins

In Gram-negative bacteria, lipoproteins are found on the periplasmic side of the inner or outer membrane and on the cell surface. There are around 750,000 copies of the most abundant lipoprotein, Lpp, in *E. coli*. It creates a bridge between the OM and peptidoglycan, improving cell stability (Braun and Rehn, 1969; Braun, 1973; Inouye et al., 1972). There are around 100 different OM lipoproteins in *E. coli*, for most of which the function is unknown (Miyadai et al., 2004).

As with other cell envelope proteins, lipoproteins are translated in the cytoplasm and targeted for translocation across the inner membrane by the Sec machinery (Figure 1.3B). Whilst still attached to the Sec machinery, Lgt transfers a diacylglycerol group from phosphatidylglycerol to the N terminal cysteine residue of the lipoprotein (usually at the +1 position after the signal sequence) (Sankaran and Wu, 1994). The lipoprotein is cleaved by signal peptidase II allowing further acylation of the cysteine by the enzyme Lnt (Dev and Ray, 1984; Gupta et al., 1993). At this stage the protein is attached to the membrane via the N-terminal acylated cysteine and is considered a mature lipoprotein (Bos and Tommassen, 2004). The final location of lipoproteins is dependent on whether it harbours an IM retention (Lol avoidance) signal. A Lol avoidance signal is recognised as an aspartate residue at the +2 position of the mature protein, without which the lipoprotein is transported to the OM via the Lol pathway (Bos and Tommassen, 2004; Matsuyama et al., 1995).

The Lol pathway consists of an inner membrane ATP-binding cassette (ABC) transporter (LolCDE), a periplasmic chaperone (LolA) and an outer membrane lipoprotein receptor (LolB) (Figure 1.3C) (Bos and Tommassen, 2004; Matsuyama et al., 1997). Lipoproteins fated for the outer membrane are bound independently of ATP by the permeases LolC/LolE. This binding creates a conformational change and increases affinity of the LolD ATPase for ATP (Bos and Tommassen, 2004). The ATP binding causes a further conformational change in LolC/LolE that decreases its affinity for lipoproteins. ATP hydrolysis releases the lipoprotein from the LolC/LolE complex onto LolA (Bos and

Tommassen, 2004; Ito et al., 2006). The LolA-lipoprotein complex crosses the periplasm to interact with the OM lipoprotein, LolB (Matsuyama et al., 1997; Yokota et al., 1999). The method by which the lipoprotein is incorporated into the outer membrane remains unknown (Okuda and Tokuda, 2011). All five components of the Lol pathway (LolABCDE) are essential for cell viability in *E. coli* (Okuda and Tokuda, 2011).

1.2.4. Lipopolysaccharide biogenesis and transport to the outer membrane

LPS is the major component of the outer leaflet of the outer membrane in Gram-negative bacteria. The molecule forms a tight layer that is low in fluidity and has low permeability to hydrophobic solutes, particularly in the presence of divalent cations that form bridges between the molecules (Sperandeo et al., 2009). Loss of proteins involved in the biosynthesis or trafficking of LPS are lethal in many Gram-negative bacteria (Bos and Tommassen, 2004). LPS is made up of three main components: lipid A, core (inner and outer) and O-antigen (Figure 1.4) (Bos and Tommassen, 2004). Lipid A is the most conserved of these components and consists of two glucosamine units with 6 fatty acids and 2 phosphate groups (Bos and Tommassen, 2004). It is also known as endotoxin and is detected by the TLR4 receptor of mammalian innate immunity (Sperandeo et al., 2009). The inner core is also well conserved and is anchored to lipid A, consisting of the unusual sugars KDO (3-deoxy-D-manno-oct-2-2-ulosonic acid) and heptose (Sperandeo et al., 2009). Linked to the inner core is the outer core, which is slightly more variable and contains more common sugars such as hexoses and hexosamines (Erridge et al., 2002). Finally, the outer core is linked to O-antigen, a polysaccharide chain that is highly variable in bacteria. O-antigen is important in protecting bacteria during infection, particularly against phagocytosis and complement-mediated killing (Trent et al., 2006). As a surface-exposed polysaccharide it is a prime antibody target and the different variations are used to serotype bacterial strains (Erridge et al., 2002). In the laboratory, O-antigen is non-essential and is missing in the laboratory strain *Escherichia coli* K-12.

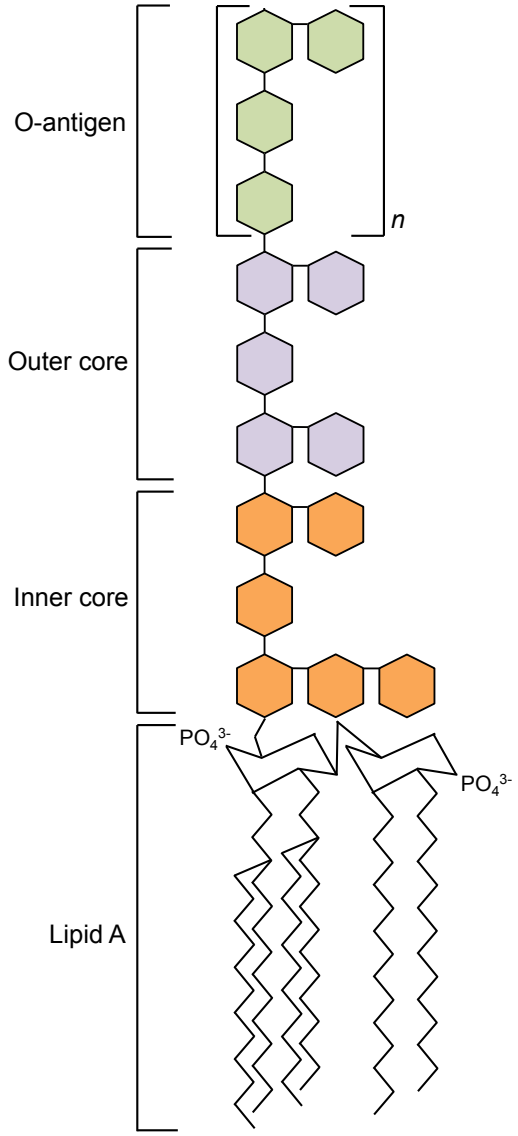


Figure 1.4. The structure of lipopolysaccharide (LPS). LPS is made up of three main components: lipid A, core (inner and outer) and O-antigen. Lipid A consists of two glucosamine units with 6 fatty acids and 2 phosphate groups. The inner core is anchored to lipid A and consists of KDO (3-deoxy-D-manno-oct-2-2-ulosonic acid) and heptose. Linked to the inner core is the outer core, which contains sugars such as hexoses and hexosamines. In the organisms in which it is found, an O-antigen polysaccharide chain is anchored to the outer core.

The lipid A and inner core components of LPS are synthesised separately and brought together at the inner leaflet of the inner membrane. The first few steps of lipid A biogenesis generate lipid IVA, the tetra-acylated precursor of lipid A. Lipid IVA consists of a disaccharide backbone with four fatty acid chains (Sperandeo et al., 2009). Lipid IVA is made from two uridine diphosphate N-acetylglucosamine (UDP-NAG) molecules by a series of Lpx proteins, C-D, H and K (Raetz et al., 2009). WaaA adds two cytidine monophosphate-KDO molecules to lipid IVA, which are synthesised by a separate pathway, to form $(KDO)_2$-lipid IVA. LpxL and M catalyse two more acylation reactions, forming the hexa-acylated $(KDO)_2$-Lipid A (Sperandeo et al., 2009). The core oligosaccharide (outer core) is synthesised from KDO_2-Lipid A by a series of ten Waa enzymes to form the lipid A-core structure (Karp et al., 2002). The lipid A-core is flipped across the inner membrane by the flippase MsbA (Figure 1.3) (Karow and Georgopoulos, 1993). When present, the O-antigen component is made independently in the cytoplasm, flipped across the inner membrane and ligated to the lipid A-core at the periplasmic leaflet of the inner membrane by the ligase WaaL (Abeyrathne et al., 2005).

Once an intact LPS molecule has been assembled, it is transported to the outer membrane by the Lpt pathway (Figure 1.3). The inner membrane component of this pathway is a predicted ABC transporter, consisting of LptB, LptC, LptF and LptG (Narita and Tokuda, 2009). LptB is a cytoplasmic facing ATPase that interacts with inner membrane permeases LptG and LptF to form the core component of the ABC transporter (Ruiz et al., 2008). This complex associates with periplasmic facing LptC (Sperandeo 2006). The LptBCFG complex is thought to extract LPS from the inner membrane and pass it to the periplasmic chaperone LptA (Ma et al., 2008; Tran et al., 2008). LptA transfers LPS to the LptD–LptE outer membrane complex, which incorporates it into the outer leaflet (Wu et al., 2006). It is thought that LptC, A and D may form a continuous transmembrane bridge across the periplasm for LPS transportation (Freinkman et al., 2012).

1.2.5. Phospholipids

Phospholipids are the main component of the inner membrane and the inner leaflet of the outer membrane. The major phospholipid species in *E. coli,* phosphatidylethanolamine (PE), phosphatidylglycerol (PG) and cardiolipin (CL), are synthesised from phosphatidic acid (PA) at the cytoplasmic side of the IM (see section 1.3.2 for details). The first evidence that phospholipids move between membranes was seen in the 1970s, when radiolabelled phospholipids were shown to rapidly move between the outer and inner membranes (Jones and Osborn, 1977). While some phospholipids remain at the inner leaflet of the inner membrane, others are flipped across the membrane, potentially by the flippase MsbA (Zhou et al., 1998). Some of these phospholipids are destined for the inner leaflet of the outer membrane, but the method of transport is currently unknown. Instead a pathway for retrograde phospholipid trafficking, the Mla pathway, has been identified.

The Mla pathway consists of an inner membrane ABC transporter (MlaBDEF), a periplasmic chaperone MlaC, and an outer membrane lipoprotein (MlaA) (Figure 1.5B) (Malinverni and Silhavy, 2009). This pathway has been shown to be important for the prevention of phospholipid accumulation in the outer leaflet of the outer membrane (Malinverni and Silhavy, 2009). The outer membrane lipoprotein MlaA (VacJ) interacts with OmpC in the outer membrane and together they function to extract phospholipids from the outer leaflet (Chong et al., 2015; Malinverni and Silhavy, 2009). The phospholipid is passed from MlaA via the MlaC chaperone onto the inner membrane ABC transporter complex. This complex consists of MlaD, a predicted substrate binding protein that contains a sinlge MCE domain, MlaE, a transmembrane permease, MlaF, a cytoplasmic facing ATPase and MlaB, a predicted NTP binding protein (Malinverni and Silhavy, 2009). The fate of the phospholipids beyond this stage is unknown but there are two possible outcomes: 1) they are incorporated into the inner membrane or 2) they move into the cytoplasm for degradation (Malinverni and Silhavy, 2009).

While the Mla pathway might have a role in phospholipid degradation, certain enzymes in the outer membrane are known to be involved in this process. PldA is an

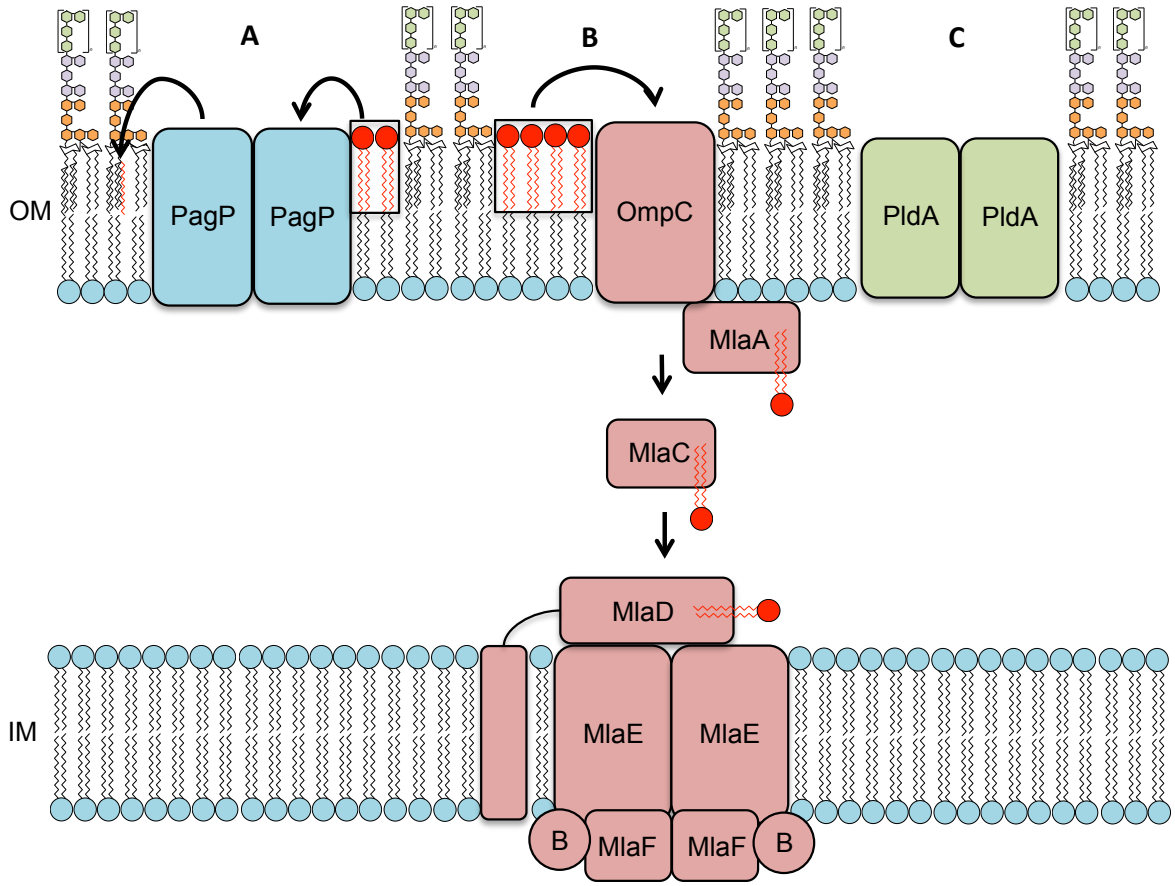


Figure 1.5. Processes involved in phospholipid degradation in the outer membrane. **A.** PagP removes an acyl chain from phospholipids and donates it to lipid A in the outer leaflet. **B.** The Mla pathway and OmpC are involved in trafficking phospholipids from the outer leaflet of the outer membrane back into the cell. **C.** PldA is a phospholipase that degrades phospholipids in the outer membrane.

enzyme that has lipolytic activity and degrades phospholipids at the outer membrane (Figure 1.5A) (Dekker, 2000). Also involved in the process is the enzyme PagP, which takes palmitate from phospholipids and donates it to lipid A (Bishop et al., 2000). This increases the acylation state of lipid A from hexa to hepta-acylated, which might prevent entry of cationic antimicrobial peptides (CAMPs) into the cell (Guo et al., 1998). Overexpression of PldA can compensate for the loss of the Mla pathway suggesting that they both play a part in removal of unwanted phospholipids from the outer membrane (Malinverni and Silhavy, 2009).

1.2.6. The periplasm and peptidoglycan

The periplasmic space is a protein-dense environment that is considerably more viscous than the cytoplasm (Mullineaux et al., 2006). Compartmentalisation allows Gram-negative bacteria to harbour enzymes in the periplasm that would otherwise be harmful (MacAlister et al., 1972; Silhavy et al., 2010). Other proteins that occupy this space include chaperones and periplasmic binding proteins for sugar and amino acid transport (Erhmann, 2007; Silhavy et al., 2010). As ATP is absent from the periplasm, energy required for transport of several outer membrane components across the periplasm is obtained from ATP hydrolysis in the cytoplasm (Okuda et al., 2012). The important peptidoglycan cell wall is also located in the periplasm. It is a rigid entity that determines cell shape and prevents cell lysis (Silhavy et al., 2010). Peptidoglycan is a polysaccharide consisting of alternating N-acetyl glucosamine (NAG) and N-acetyl muramic acid (NAM) monosaccharides (Vollmer et al., 2008), and its synthesis is reviewed by Lovering et al. (2012). Attached to NAM is a pentapeptide chain consisting of L-alanine, D-glutamic acid, mesodiaminopimelic (meso-A_2pm) acid and two D-alanine residues.

The bulk of peptidoglycan biosynthesis occurs in the cytoplasm, and starts with the conversion of uridine diphosphate (UDP)-NAG to UDP-NAM in a two-step process catalysed by MurA and B (Figure 1.6) (Benson et al., 1996; Benson et al., 1997; Brown et al., 1995). The five amino acids are added to UDP-NAM in a stepwise fashion by four different

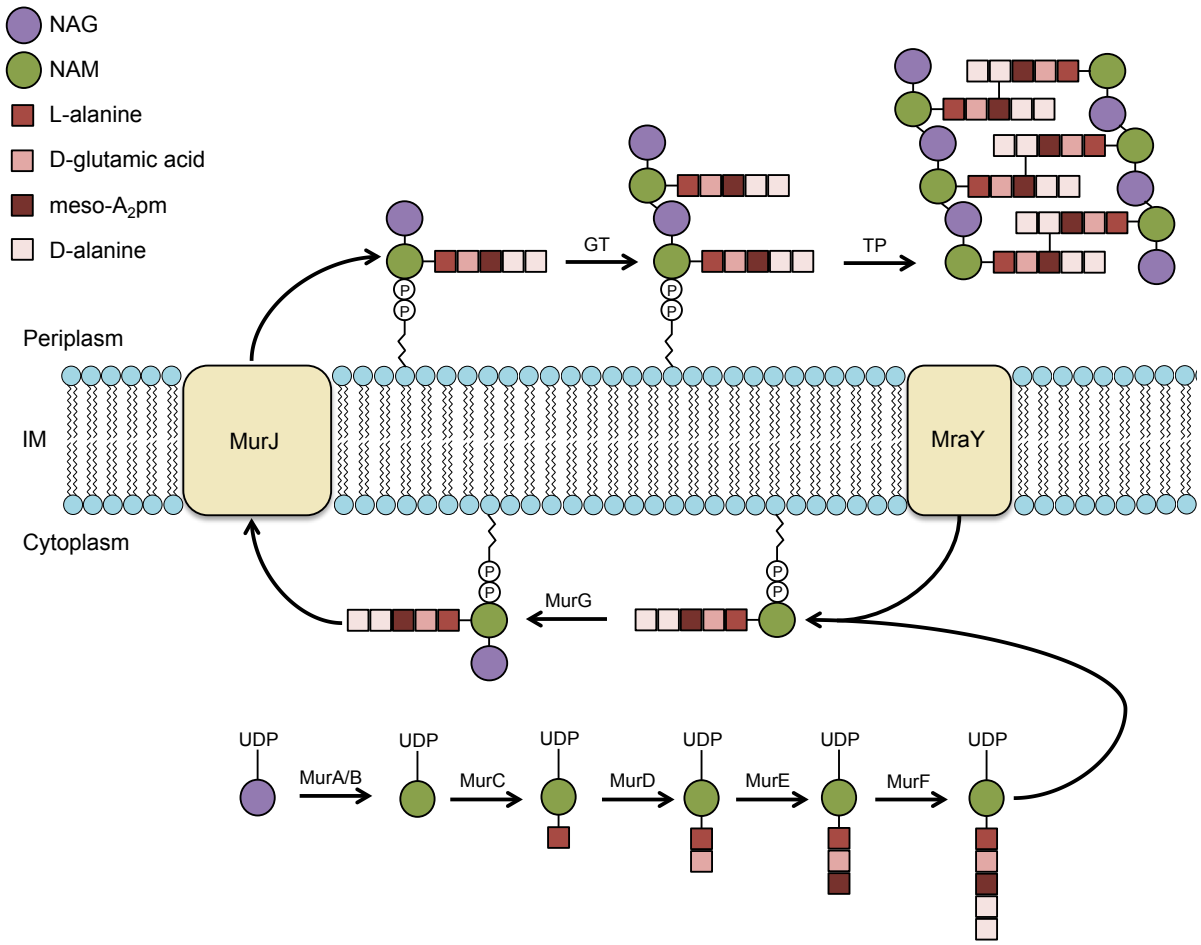


Figure 1.6. The biosynthesis of the peptidoglycan cell wall. Synthesis starts in the cytoplasm with the conversion of uridine diphosphate (UDP)-NAG to UDP-NAM, catalysed by MurA and B. MurC-F add the five amino acids to form NAM-pentapeptide. At the membrane, the NAM-pentapeptide is converted to lipid I by its transfer onto the undecaprenyl phosphate (UP) carrier by MraY. This is coupled to the transfer of UDP-activated N-acetyl D-glucosamine to lipid I (to form lipid II), by MurG. Lipid II is flipped across the membrane by MurJ. In the periplasm lipid II is elongated with the repeating unit of the NAG-NAM disaccharide by a glycosyltransferase (GT). A transpeptidase (TP) creates crosslinks between the pentapeptides.

ligases: MurC-F. At the membrane, the NAM-pentapeptide is converted to lipid I by its transfer onto the undecaprenyl phosphate (UP) carrier, which is catalysed by the inner membrane protein MraY (MurX) (Ikeda et al., 1991). This step is coupled to the MurG catalysed transfer of UDP-activated N-acetyl D-glucosamine to lipid I, to form lipid II (Lovering et al., 2012; Mengin-Lecreulx et al., 1991). Lipid II is flipped to the periplasmic leaflet by an inner membrane flippase, which was recently identified as MurJ (Sham et al., 2014). In the periplasm lipid II is elongated with the repeating unit of the NAG-NAM disaccharide by glycosyltransferase, forming a nascent peptidoglycan strand. A transpeptidase creates crosslinks between the pentapeptides of different nascent chains to form the complex mesh that is the bacterial cell wall (Lovering et al., 2012). Both glycosyltransferase and transpeptidase are types of penicillin binding protein, which are one of the most common targets for treatment of Gram-positive infections. In particular, vancomycin affects transpeptidase activity by binding the double D-alanine amino acids of the peptide chain, preventing cross-linkage of nascent peptidoglycan strands. Gram-positive bacteria can become resistant to vancomycin by alteration of the double D-alanine moiety to D-alanine D-lactate (Xie et al., 2011).

1.3 Fatty acid metabolism

Fatty acids are an important source of energy and an essential component of cell membranes. *E. coli* primarily synthesises unbranched fatty acid chains with an even number of carbons, whereas several Gram-positive bacteria favour the production of branched fatty acid chains with an odd number of carbons (Choi et al., 2000; Khandekar et al., 2001). The biosynthesis and degradation of fatty acids are two separate pathways that are switched on and off depending on fatty acid availability in the cell (Fujita et al., 2007). The regulator FadR oversees fatty acid metabolism by repressing degradation and activating some genes involved in biosynthesis (DiRusso et al., 1992). FadR is repressed by the binding of long acyl-CoA thioesters, a substrate for fatty acid degradation (DiRusso et al., 1992).

1.3.1. Fatty acid biosynthesis

There are two types of fatty acid biosynthesis pathways: Fatty acid synthase (FAS) I and II. Animals and fungi primarily use a FAS I system, which involves multifunctional enzymes catalysing multiple steps in biosynthesis (Jenke-Kodama et al., 2005). FAS II is found in bacteria and plants and includes several monofunctional enzymes (Jenke-Kodama et al., 2005).

The stages involved in bacterial fatty acids biosynthesis are reviewed by Parsons and Rock (2013). The FAS II pathway is divided into two main stages: the initiation module and the elongation module. The initiation module begins with the carboxylation of acetyl-CoA to malonyl-CoA by the acetyl-CoA carboxylase AccABCD (Figure 1.7) (Cronan and Waldrop, 2002). Next, FabD transfers the malonyl group to the acyl carrier protein (ACP), which is necessary for subsequent recognition by FAS II enzymes (Zhang and Rock, 2008). Malonyl-ACP is used for one of two things: initiation of the elongation cycle (via β-ketoacyl-ACP) or extension of the fatty acid chain during the elongation cycle (Tsay et al., 1992). For initiation of the elongation module, malonyl-ACP is first condensed with a short chain acyl-CoA to form β-ketoacyl-ACP, a reaction catalysed by FabH (Tsay et al., 1992). FabH substrate specificity differs between different bacteria and is in part responsible for the major fatty acid type found within a species (Choi et al., 2000; Qiu et al., 2005).

To begin the elongation cycle, β-ketoacyl-ACP is reduced by FabG to form β-hydroxyl-ACP (Figure 1.8). Either FabA or FabZ dehydrates β-hydroxyl-ACP to form trans-2-enoyl-ACP (Heath and Rock, 1996). The specificities of FabA and FabZ largely overlap, but FabA can convert trans-2-enoyl-ACP to cis-2-enoyl-ACP, a precursor required for the generation of unsaturated fatty acids (Heath and Rock, 1996). Following dehydration, trans-2-enoyl-ACP is reduced by FabI to form acyl-ACP (Heath and Rock, 1995). Acyl-ACP can either be used in the acyl transfer module (see 1.3.2) or it can be condensed by FabB/FabF to initiate a new round of elongation (Jackowski et al., 2002).

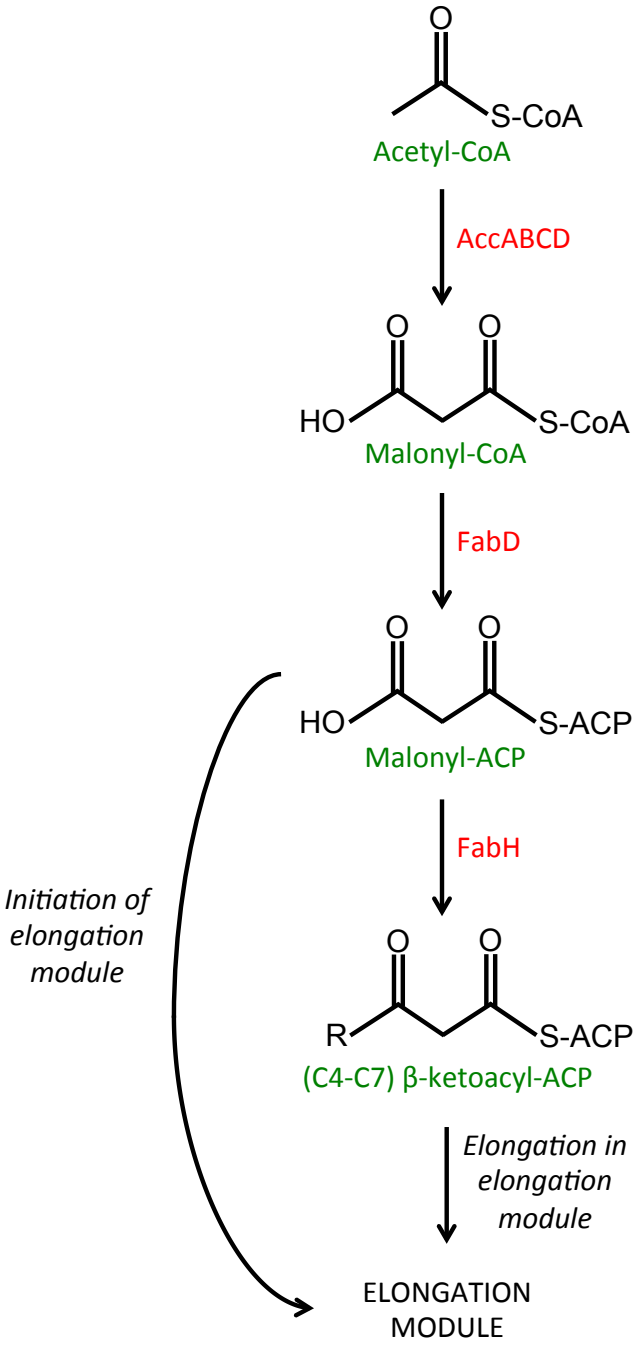


Figure 1.7. The initiation module of fatty acid biosynthesis. Acetyl-CoA is carboxylated to malonyl-CoA by AccABCD. FabD transfers the malonyl group to acyl carrier protein (ACP). Malonyl-ACP is either used for the initiation of the elongation cycle (via conversion to β-ketoacyl-ACP by FabH) or extension of the fatty acid chain during the elongation cycle.

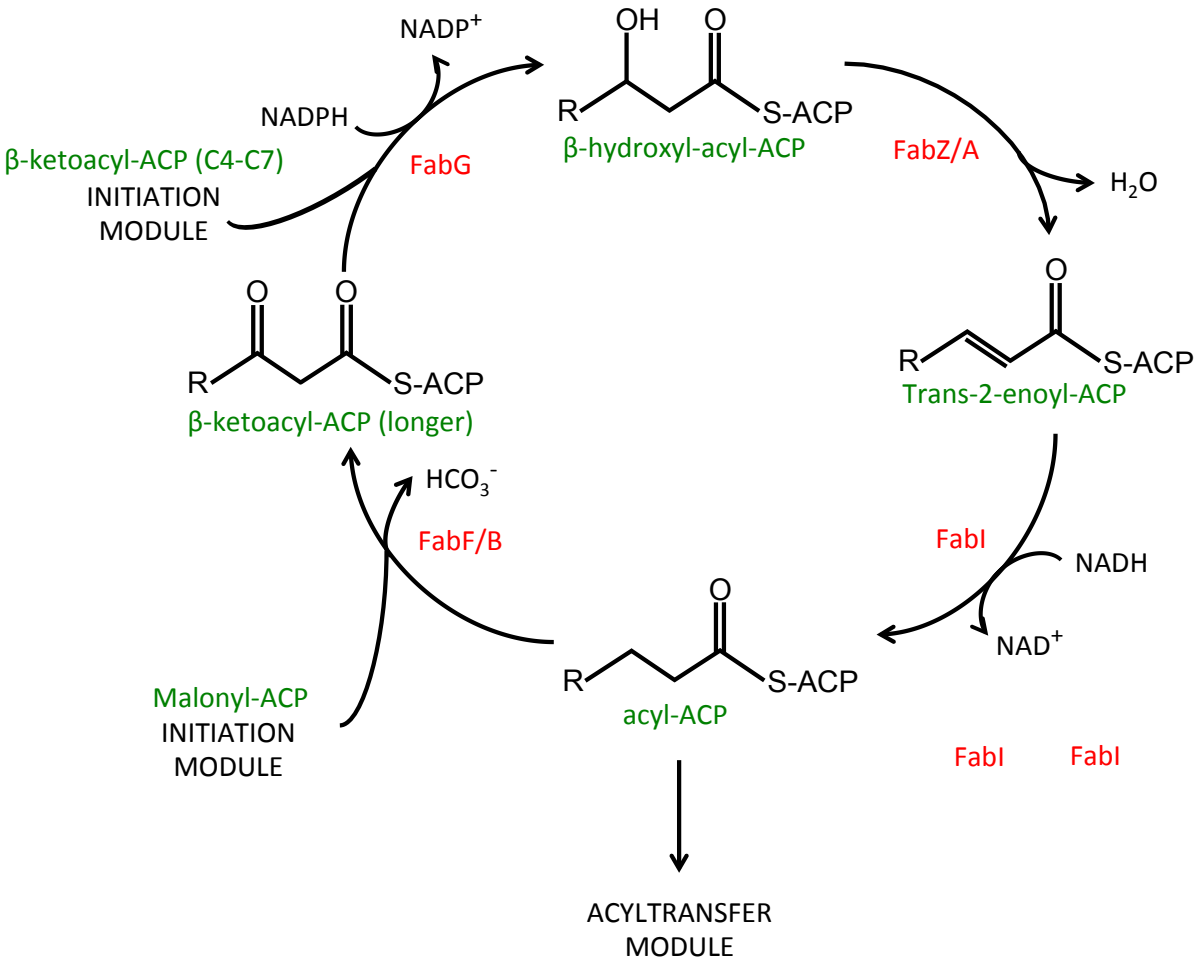


Figure 1.8. The elongation module of fatty acid biosynthesis. β-ketoacyl-ACP is reduced to β-hydroxyl-ACP by FabG. FabA/FabZ dehydrates β-hydroxyl-ACP to trans-2-enoyl-ACP. Trans-2-enoyl-ACP is reduced to acyl-ACP by FabI. Acyl-ACP can either be used in the acyl transfer module or it can be condensed by FabB/FabF to initiate a new round of elongation.

The elongation module is primarily used for the formation of saturated fatty acids, but unsaturated fatty acids can also utilise parts of the cycle. The isomerisation catalysed by FabA specifically occurs at the 10-carbon stage of the elongation module, forming cis-2-decenoyl-ACP (Heath and Rock, 1996). The molecule is then elongated by FabB/FabF to form an unsaturated fatty acid (Heath and Rock, 1996). Fatty acid biosynthesis is partly regulated by internal feedback loops. For example, AccABCD that catalyses the first step of the initiation module is inhibited by acyl-ACP, the final product of the elongation module (Davis and Cronan, 2001). Several intermediates of fatty acid biosynthesis are used for other important cellular processes. Enzymes of lipid A biosynthesis are specific for ACP thioesters and therefore the elongation cycle provides the only source of hydroxyl fatty acids for lipid A biosynthesis (Raetz et al., 2007).

1.3.2. The acyltransferase module and phospholipid biosynthesis

The acyltransferase module involves the transfer of acyl chains from acyl-ACP onto glycerol-3-phosphate to begin the synthesis of phopsholipids (Figure 1.9). The enzymes involved in this process in *E. coli* are the phospholipid acyltransferases PlsB and PlsC (Bell, 1975). PlsB ligates a fatty acid at the sn-1 carbon position of glycerol-2-phosphate to form 1-acyl-glycerol-3-phosphate or lysophosphatidic acid. PlsC subsequently ligates a fatty acid at the sn-2 position producing 1,2-diacyl-sn-glycerol-3-phosphate or phosphatidic acid (PA), the phospholipid precursor for all other phospholipids (Bell, 1975). Like all naturally occurring phospholipids, a saturated fatty acid is found at position 1 of PA and an unsaturated at position 2 (Cronan and Vagelos, 1972).

In *E. coli* the membrane phospholipid content is primarily made up of phosphatidylethanolamine (PE), followed by phosphatidylglycerol (PG) and cardiolipin (CL). To start phospholipid synthesis, CdsA converts PA to the intermediate molecule, cytidine diphosphate diacylglycerol (CDP-DAG) (Parsons and Rock, 2013). At this stage, phospholipid head group synthesis can be split into two pathways: zwitterionic phospholipid synthesis (PE); and anionic phospholipid synthesis (PG and CL).

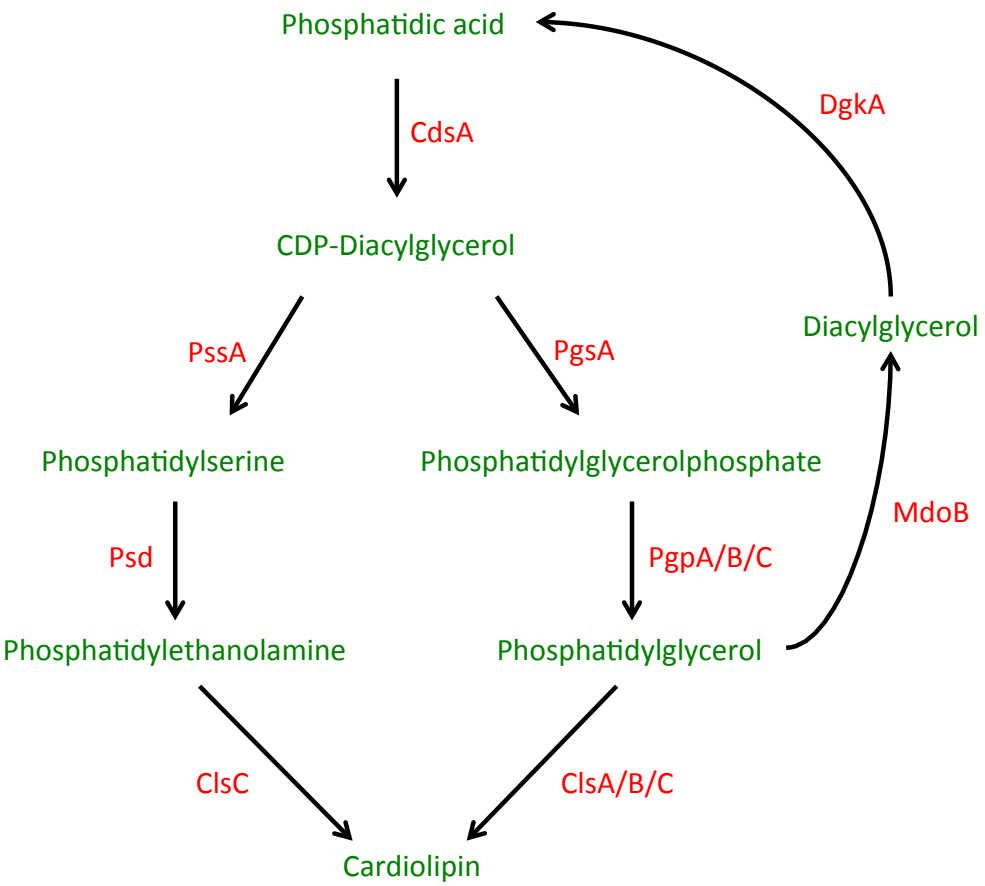


Figure 1.9. Phospholipid biosynthesis pathway. Phosphatidic acid (PA) is the precursor for all other phospholipids. PA is converted to CDP-diacylglycerol by CdsA. PssA starts the zwitterionic phospholipids synthesis pathway by conversion of CDP-diacylglycerol to phoshatidylserine (PS), which is subsequently converted to phosphatidylethanolamine (PE) by Psd. Alternatively, CDP-DAG can be converted to phosphadtidylglycerol-phosphate by PgsA to begin anionic phospholipids synthesis. Phosphadtidylglycerol-phosphate is converted to phosphatidylglycerol by the three enzymes PgpA, B and C. Cardiolipin is synthesised by ClsA, B and C, either from two PG molecules or one PE and one PG molecule. PG can be recycled back to PA by MdoB and DgkA, via the intermediate diacylglycerol.

The first step of the PE pathway is the condensation of CDP-DAG with serine to produce phosphatidylserine (PS), a reaction that is catalysed by PssA. PS is then decarboxylated by Psd to form PE (Parsons and Rock, 2013). As the most abundant phospholipid in the cell, mutants defective in PE synthesis can only grow under certain conditions. The temperature-sensitive phenotype of a *psd* mutant is a consequence of PS accumulation in the membrane (Hawrot and Kennedy, 1978). A *pssA* deletion strain cannot produce PS or PE and compensates with increased amounts of PG and CL. A *pssA* mutant is only viable when grown in media supplemented with Ca^{2+}, Mg^{2+} or Sr^{2+} (DeChavigny et al., 1991).

The first step of the PG pathway involves the replacement of cytidine monophosphate on CDP-DAG by glycerol-phosphate, a reaction catalyzed by the enzyme PgsA (Miyazaki et al., 1985). The product phosphadtidylglycerol-phosphate is desphosphorylated by the three enzymes PgpA, B and C to form PG (Lu et al., 2011). A *pgsA* deletion is only viable in the presence of a secondary mutation in *lpp*. This is because PG donates an acyl group to the N-terminal cysteine of lipoproteins, allowing correct localization (Suzuki 2002). Lpp is a highly expressed lipoprotein that locates to the OM and forms interactions with peptidoglycan (Braun and Rehn, 1969). In the case of a *pgsA* knockout, Lpp cannot localise correctly and builds up in the inner membrane, interacting with peptidoglycan and resulting in cell lysis (Suzuki et al., 2002). The viability of a *pgsA lpp* double mutant suggests that the *E. coli* does not need anionic phospholipids under laboratory conditions.

CL is synthesised by three different enzymes: ClsA, B and C. ClsA is the main CL synthase during exponential growth, while all three are active during stationary phase (Tan et al., 2012). ClsA and B both form CL by the condensation of two PG molecules. ClsC uses PE as a phosphatidyl donor to PG to form CL, providing some crossover between the zwitterionic and anionic phospholipid pathways (Tan et al., 2012). A triple *clsABC* mutant completely lacks CL synthesis but is viable, suggesting that CL is not essential in *E. coli* (Tan et al., 2012).

A phospholipid breakdown pathway has been defined, which mainly targets PG as a substrate. The sn-glycerol-1-phosphate headgroup is removed from PG by the enzyme MdoB, producing DAG (Jackson et al., 1984). This headgroup is donated to the formation of membrane-derived oligosaccharides (MDOs). MDOs are known to be important in the regulation of osmolality in the periplasm, as an *mdoB* mutant is sensitive in environments that cause osmotic stress (Jackson et al., 1984; Rajagopal et al., 2003). Following removal of the headgroup, DAG is transformed back into PA by the enzyme DgkA (Raetz and Newman, 1978). Mutants in *dgkA* show accumulation of DAG, which is lethal in osmotically challenging environments (Jackson et al., 1984).

As with fatty acid biosynthesis, a feedback loop for phospholipid biosynthesis helps regulate the inhibition and activation of different pathways. For example, an increase in the number of anionic phospholipids in the membrane activates Pss to increase PE synthesis (Louie et al., 1986). The phospholipid content of the membrane also changes in response to external stimuli, such as temperature. As temperature decreases, the membrane becomes more rigid, which is compensated by the increased incorporation of unsaturated fatty acids into phospholipids (Jackowski et al., 2002). When the temperature increases, the membrane fluidity increases and the proportion of unsaturated fatty acids in the membrane is decreased (Jackowski et al., 2002).

1.3.3. Fatty acid degradation

E. coli can import fatty acids from outside the cell for use as either a carbon source, or as precursors for phospholipid biosynthesis (Fujita et al., 2007). The pathway for fatty acid degradation is known as the β-oxidation pathway. To enter the cell, short-chain fatty acids can diffuse across the membrane and do not require a specific uptake mechanism. Longer chain fatty acids are too large for this diffusion and are transported across the membrane via the FadL/FadD system (Figure 1.10). FadL is responsible for transport of long-chain (C12-18) fatty acids across the OM (Black, 1991). At the inner membrane, fatty acids are transported and activated by FadD to form acyl-CoA in the cytoplasm (Black et al., 1992).

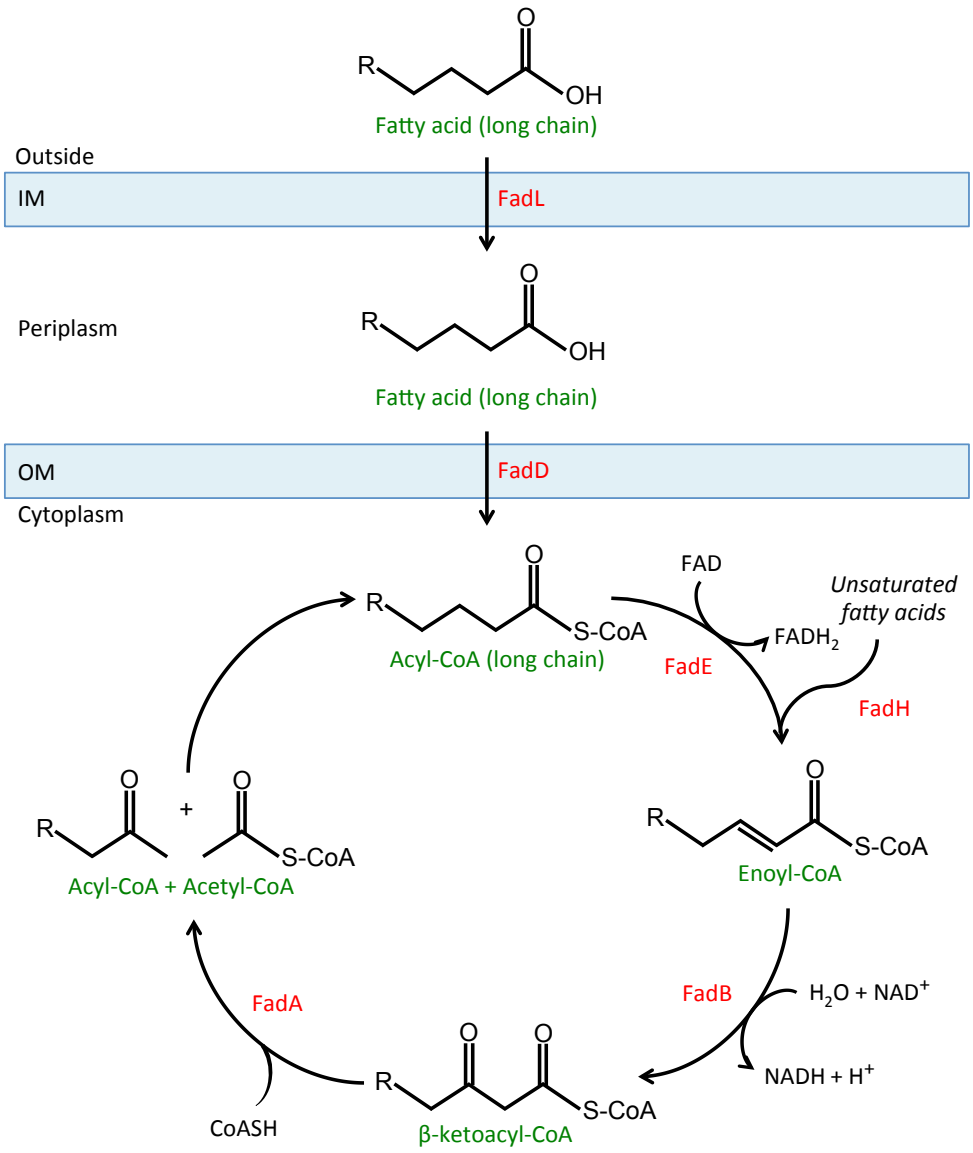


Figure 1.10. Fatty acid degradation in *Escherichia coli*. Short chain fatty acids can diffuse across the outer and inner membranes without aid. FadL is required for transport of long chain fatty acids across the outer membrane. FadD helps transport of long chain fatty acids across the inner membrane and activates all fatty acids for degradation, forming acyl-CoA. The β-oxidation cycle begins with the conversion of acyl-CoA to enoyl-CoA by FadE. FadB converts enoyl-CoA to β-ketoacyl-CoA, which is then shortened by FadA shortens, producing a shortened acyl-CoA plus acetyl-CoA. The acyl-CoA then enters another round of the β-oxidation cycle. FadH is responsible for the introduction of unsaturated fatty acids into the cycle for degradation.

The method by which fatty acids cross the periplasm is currently unknown (Parsons and Rock, 2013). In the cytoplasm, the β-oxidation cycle begins with the conversion of acyl-CoA to enoyl-CoA by FadE, the only acyl-CoA dehydrogenase found in *E. coli* (Campbell and Cronan, 2002; Fujita et al., 2007). FadB hydrates and oxidises enoyl-CoA to β-ketoacyl-CoA via the intermediate β-hydroxyacyl-CoA (Pramanik et al., 1979). FadA shortens β-ketoacyl-CoA by two carbons per cycle, producing a shortened acyl-CoA plus acetyl-CoA (Pramanik et al., 1979). The β-oxidation cycle can also degrade unsaturated fatty acids via their derivative, 2,4-dienolyl-CoA. FadH transforms 2,4-dienolyl-CoA into enoyl-CoA, allowing subsequent hydration and oxidation by FadB and cleavage by FadA (Fujita et al., 2007). During anaerobic β-oxidation, FadI, FadJ and FadK replace FadA, FadB and FadD respectively, but carry out equivalent functions (Campbell et al., 2003).

1.4. The Salmonella infection model

Salmonella is a clinically important Gram-negative pathogen belonging to the family Enterobacteriaceae. Together, *Salmonella* and *Campylobacter* are the most commonly diagnosed foodborne pathogens (Silva et al., 2011). The species *S. enterica* contributes to the majority of *Salmonella* infection in mammals with over one billion cases reported annually (Coburn et al., 2007). There are multiple subspecies of *S. enterica*, including *S. enteria* subsp. *enterica*. The *enterica* subspecies is further derived into serovars based on their flagella, carbohydrate and LPS properties (Fierer and Guiney, 2001). In humans most *enterica* serovars cause enteritis, with the exception of Typhi, Paratyphi and Sendai, which can cause enteric (typhoid) fever (Coburn et al., 2007). *S. enteria* subsp. *enterica* serovar Typhimurium (*S.* Typhimurium) is a serovar that causes enteritis in humans but can cause typhoid fever in susceptible mouse strains.

S. Typhimurium infection in mice is used as a model of *S.* Typhi infection in humans and is the source of much of the understanding of systemic *Salmonella* infection (Coburn et al., 2007). In short, *Salmonella* are ingested and infect epithelial cells in Peyer's patches in the small intestine. The bacteria enter phagocytes and disseminate through the

reticuloendothelial system, gaining access to reticuloendothelial system organs, where the infection can persist long term (Figure 1.11) (Monack et al., 2004).

The innate immune response is the first line of defence against invading pathogens. It includes physical barriers such as skin, mucus and stomach acid and induced response such as inflammation and phagocytosis-mediated killing (Medzhitov, 2010). Pattern-recognition receptors are a group of receptors of the host immune system that detect pathogen-associated molecular patterns (PAMPs) (Vance et al., 2009). They include the well studies Toll-like receptors (TLRs) that play an important role in the innate inflammatory response (Broz et al., 2012). Factors of the innate immune response work to clear *Salmonella* infection and prevent systemic spread and persistence. If this fails, the adaptive immune response is required for clearance of more persistent infections.

During infection, *Salmonella* has to overcome these barriers of the immune system, which change as infection proceeds and the bacteria move to different locations. The stages of infection can be split as follows: 1) passage through the stomach, 2) invasion of the small intestine, 3) establishment of systemic infection and 4) persistence and the adaptive immune response.

1.4.1. Passage through the stomach

Salmonella infections begin with ingestion. The first barrier to overcome is the low pH of stomach acid, an environment to which *Salmonella* is exposed only for a short time (Ohl and Miller, 2001). *Salmonella* is usually ingested with food, when the pH of the stomach is increased along with the chances of survival (Waterman and Small, 1998). *Salmonella* have evolved an acid tolerance response (Garcia-del Portillo et al., 1993). To maintain intracellular pH, proton pumps continuously export H^+ ions out of the cytoplasm (Foster, 2000). Similarly, low pH induces amino acid decarboxylase systems that incorporate protons into certain amino acids and export them out of the cell (Alvarez-Ordonez et al., 2011). Other changes include a decrease in membrane fluidity (by decrease in membrane

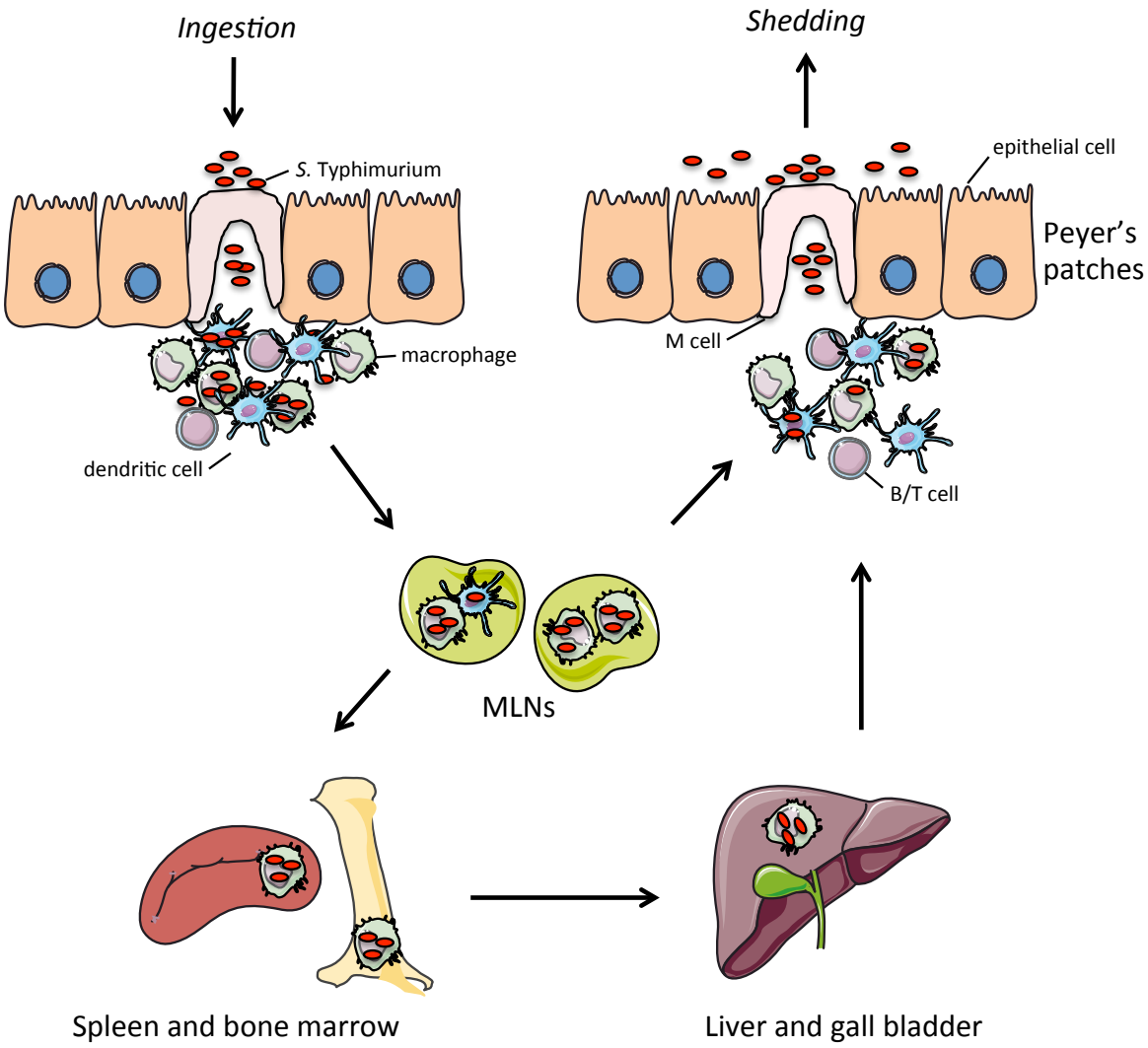


Figure 1.11. The *Salmonella* Typhimurium murine model for systemic infection. After ingestion, the bacteria infect epithelial cells in the Peyer's patches of the small intestine. The bacteria cross the epithelial barrier to reach immune cells including dendritic cells and submucosal macrophages. The bacteria enter phagocytes and disseminate through the mesenteric lymph nodes (MLNs) and onto the liver, gall bladder, spleen and bone marrow. During persistent infection, which involves long-term colonisation of the gall bladder, the bacteria continually shed through the faeces.

unsaturated fatty acids) and induction of acid shock proteins that repair damage caused by low pH (Alvarez-Ordonez et al., 2011; Audia et al., 2001).

1.4.2. Invasion of the small intestine

Having survived passage through the stomach, bacteria are exposed to new challenges in the small intestine. The intestinal lumen is high in salt, and *Salmonella* respond to this by increasing intracellular levels of solutes (including K^+, betaine, carnitine and proline) either by uptake or synthesis (Alvarez-Ordonez et al., 2011). *Salmonella* also have to respond to the reduced oxygen levels in the gut by inducing expression of genes involved in anaerobic metabolism (Alvarez-Ordonez et al., 2011). In the small intestine, *Salmonella* adhere to epithelial cells, with a particular preference for microfold (M) cells (Ohl and Miller, 2001). M cells are found in Peyer's patches and monitor intestinal antigens via pinocytosis, a specialised type of endocytosis for fluids (Brandtzaeg, 1989). *Salmonella* adhere to epithelial cells via their fimbriae and other surface structures. At this stage, genes located on the chromosomal *Salmonella* pathogenicity island 1 (SPI-1) are activated. SPI-1 includes components of a type 3 secretion system (T3SS), which transports effectors via an injectisome from the cytoplasm of the bacterium into the cytoplasm of the host cell (Ramos-Morales, 2012). The effector protein SopE is translocated via this SPI-1 T3SS and interacts with members of the Rho-family proteins in the host that regulate the actin cytoskeleton (Hall, 1998). SipA, SipC, SopB, SopD and SopE2 are also involved in actin rearrangement of the host cell (McGhie et al., 2001; McGhie et al., 2009; Raffatellu et al., 2005). This actin rearrangement results in *Salmonella* engulfment into the host cell to form *Salmonella* containing vesicles (Francis et al., 1992). Once the *Salmonella* is internalised, another SPI-1 effector, SptP, returns the host cell back to its resting state (Patel and Galan, 2005).

At this stage, serovars associated with enteritis trigger a large secretory response in the epithelia via the host pattern-recognition receptors such as TLRs (Broz et al., 2012). Following detection by TLRs, host cells express proinflammatory cytokines that induce the production of mucins, antimicrobial peptides and chemokines (Galyov et al., 1997). Certain

chemokines, such as IL-8, recruit neutrophils that subsequently accumulate in the intestinal lumen (Galyov et al., 1997). In *Salmonella,* SopB (a SPI-1 effector) is required for neutrophil accumulation and increased secretion of chloride ions in the lumen, both of which promote diarrhoea (Norris et al., 1998).

1.4.3. Systemic infection

After uptake into epithelial cells, *Salmonella* migrates through the epithelia to reach sub-mucosal macrophages at the sub-mucosal membrane (Ohl and Miller, 2001). Within the macrophage phagosome, *Salmonella* has to withstand factors designed to kill pathogens, including reactive oxygen species, cationic antimicrobial peptides (CAMPs) and hydrolytic enzymes (Ohl and Miller, 2001). At this stage, bacteria that cause enteritis proceed no further, while enteric fever serovars can use the macrophage for replication and hide from the immune system. One of the major regulators of survival in these conditions is the PhoP-PhoQ two-component system, which is activated inside the macrophage phagosome by antimicrobial peptides and low Mg^{2+} levels (Hancock and McPhee, 2005; Miller et al., 1989; Ohl and Miller, 2001). Amongst other things, this system activates modifications in LPS, such as acylation by the enzyme PagP, which prevent membrane disruption and cell death by CAMPs (Ganz and Weiss, 1997; Guo et al., 1998). Furthermore, these modifications reduce the immunogenicity of LPS by up to 100-fold (Kawasaki et al., 2004). In total, the PhoPQ system regulates over 200 different genes including those involved in drug resistance, motility and defence against oxidative stress (Minagawa et al., 2003).

In the phagosome, genes found on another pathogenicity island, SPI-2, are induced via the OmpR-EnvZ system. OmpR-EnvZ is activated in conditions of low osmolality and nutrients, and by acidification of the *Salmonella* containing vesicle in the macrophage (Kaur and Jain, 2012). These genes encode another T3SS that injects effector proteins into the host cell. The SPI-2 T3SS injects an array of proteins (including SifA, SifB, SseF, SseG, SseJ, PipB and SopD2) involved in formation of *Salmonella*-induced filaments (Waterman and Holden, 2003). These filaments extend from the vesicle and might have a role in the

intracellular replication of *Salmonella* (Knodler and Steele-Mortimer, 2003). Other SPI-2 effector proteins are involved in avoidance of NADPH-oxidase dependent killing, delay of host cell apoptosis-like death, and maintenance of the *Salmonella* containing vesicle membrane (Waterman and Holden, 2003).

In the sub-mucosal environment, it is possible that *Salmonella* specifically targets host cells that disseminate through the lymphatic system. As well as the sub-mucosal macrophages, *Salmonella* can also enter dendritic cells where they cannot replicate but remain viable (Kaur and Jain, 2012). Dendritic cells migrate widely within the body and are important for antigen presentation. *Salmonella* harbouring dendritic cells and macrophages disseminate through the lymphatic system to the mesenteric lymph nodes (MLNs) and onto the liver, gall bladder, spleen and bone marrow (Kaur and Jain, 2012).

1.4.4. Persistence and the adaptive immune response

Once systemic, *Salmonella* infection can persist for months with periodical shedding of bacteria in the stool (Monack et al., 2004). In mice, early prevention of persistence is partly due to *nramp1* found in monocytes/macrophages. It prevents the intracellular replication of *Salmonella* by depriving the bacteria of divalent cations, controlling bacterial growth in RES organs during the first week of infection (Gruenheid et al., 1997; Vidal et al., 1995). Later clearance of *Salmonella* (3-4 weeks of infection) involves input from the adaptive immune response. In particular, CD4+ and CD8+ T lymphocytes and the humoral (antibody) response are all important for late stage infection (Monack et al., 2004). Antigen presentation via the major histocompatibility complex (MHC) II on the surface of antigen presenting cells, such as macrophages and dendritic cells, activates CD4+ T cells that provide positive feedback to the innate immune system (Monack et al., 2004). CD4+ T helper cells are important for the humoral response, by aiding in the stimulation of B-cells and subsequent antibody production. CD8+ T cells function by inducing apoptosis of infected cells following the detection of foreign antigens on the cell surface (via MHC I molecules) (Monack et al., 2004).

Most systemic *Salmonella* infections will be successfully cleared by the immune system. Persistent *Salmonella* infection is in part due to effectors of the SPI-2 T3SS that can suppress antigen presentation, and limit the chance of an adaptive immune response (Waterman and Holden, 2003). Furthermore, *Salmonella* can modulate the immune response for example by inducing the production of cytokines that play a role in immunosuppression (Pie et al., 1996). However, its most successful survival tactic is simply to hide from the immune system long term. Aside from intracellular protection, *S.* Typhi is adapted to long-term colonisation of the gall bladder by forming a biofilm that protects it from the antimicrobial effects of bile and the immune system (Gunn et al., 2014). During persistent infection, bacteria continually shed via faeces. In some cases *S.* Typhi can cause chronic life-long infection that is often asymptomatic. These carriers shed *Salmonella* in their stool and act as a reservoir for the pathogen that is easily spread to new, susceptible hosts (Monack et al., 2004).

1.5. Mammalian cell entry domains

The PFAM database defines an MCE domain through a hidden Markov model (HMM) describing an 81 amino acid sequence. The hidden Markov model contains particularly well-conserved hydrophobic residues including valines, leucines, isoleucines and glycines (Finn et al., 2016). The glycine residue at position 29 is by far the most conserved amino acid in the model (Figure 1.12).

MCE domain-containing proteins have been reported to resemble components of ABC transporters (Casali and Riley, 2007). A typical ABC transporter consists of four core protein domains. Two of these domains are permeases, which span the membrane 6 times forming a pore through which solutes pass. The other two domains are ATP binding domains, which face the cytoplasm and are responsible for the hydrolysis of ATP (Higgins, 2001). It is these conserved ATPase domains that are used to classify proteins as part of the ABC transporter family (Higgins et al., 1986).

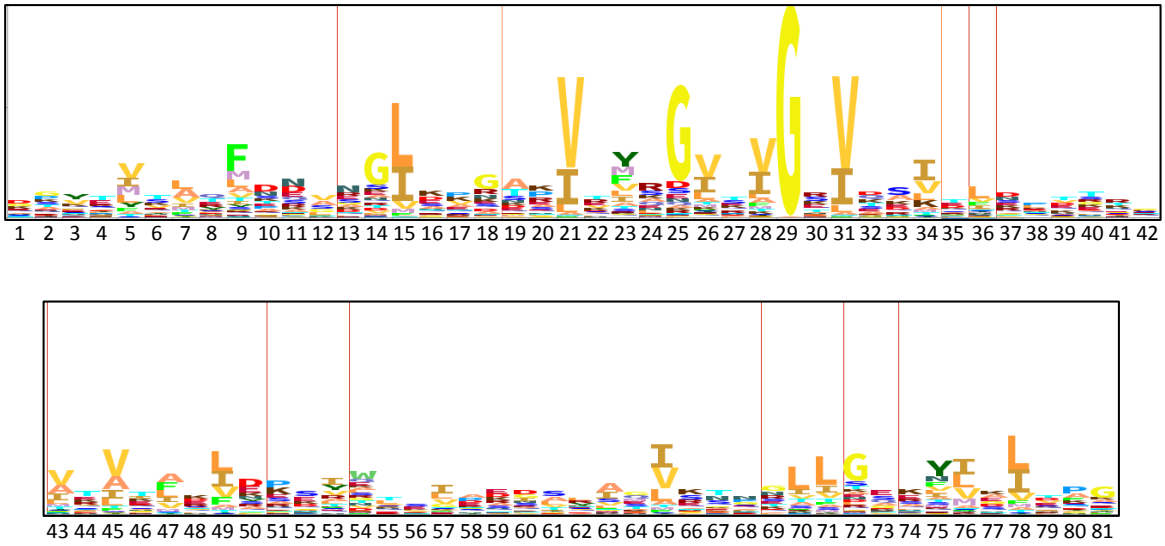


Figure 1.12. The PFAM HMM for an MCE domain. The numbers represent amino acid position and the height of the letters represent the relative importance of amino acids in the model.

MCE domain containing proteins resemble substrate-binding proteins (SBPs) (Kumar et al., 2005), which are usually specific to ABC importer systems (Casali and Riley, 2007; Higgins, 2001). SBPs are thought to bind substrates in the periplasm (or externally in Gram-positive bacteria) and deliver them to the membrane component of the transporter. The presence of SBPs tends to increase the specificity of an ABC transporter (Higgins, 2001). MCE domains are widespread across bacteria and are found in both Gram- positive and Gram-negative phyla. In particular, they are prominent in Proteobacteria, Bacteroidetes, and Actinobacteria (Finn et al., 2016). While some studies have looked at the role of MCE domains in Proteobacteria, most information on MCE domains has been acquired from studies in Actinobacteria.

1.5.1. Mammalian cell entry domains in Actinobacteria

Actinobacteria is a phylum that includes clinically important *Mycobacterium* and *Streptomyces* species. As a whole the phylum is classed as Gram-positive, but *Mycobacterium* species do not Gram stain due to their unusual mycolic acid-rich outer membrane.

The first MCE domain was identified in the gene *mce1A* from *M. tuberculosis* in 1993 (Arruda et al., 1993). This was later identified as being part of an operon containing six *mce* genes and two *yrbE* genes, the latter of which encode permeases and are orthologues of *mlaE* (Figure 1.13) (Casali and Riley, 2007). The Mce proteins each contain one MCE domain and are sometimes associated with a domain of unknown function, DUF3407, within the same protein. The sequencing of the *M. tuberculosis* genome in 1999 revealed that there are four variants of this operon in the genome, presumably following duplication events (Casali and Riley, 2007; Cole et al., 1998). These operons all contain the two *yrbE* genes and six *mce* genes, and MCE-associated genes are found downstream in three of the four operons (Figure 1.13) (Casali and Riley, 2007). Operons one to three contain genes that regulate the operon: *mce1R, mce2R* and *mce3R*. The *mce1* operon also contains *fadD5*, suggesting that it could be linked to fatty acid degradation.

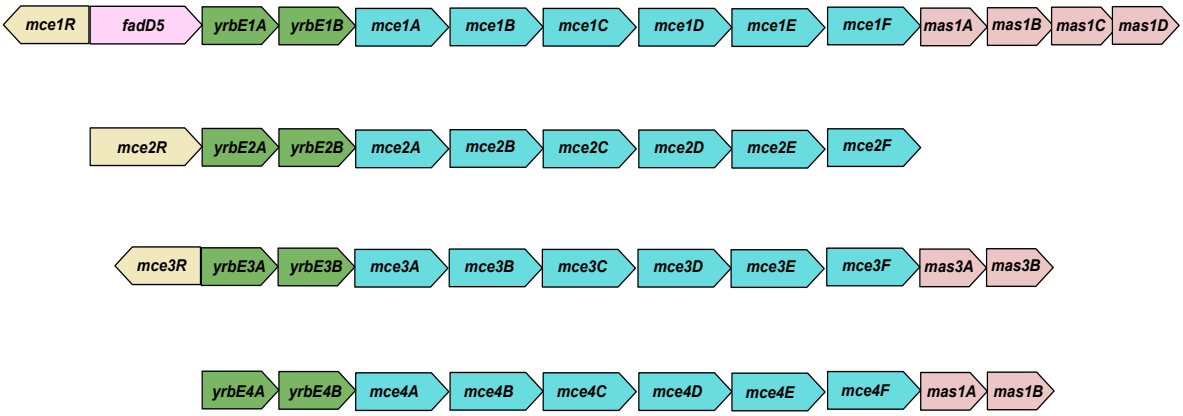


Figure 1.13. The four *mce* operons found in *Mycobacterium tuberculosis*. Each operons contains two *yrbE* genes followed by six *mce* genes. Operons 1, 3 and 4 contain mce-associated (*mas)* genes at the end of the operon. Operons 1, 2 and 3 contain *mce* regulator genes *(*e.g. *mce3R)* at the start of the operon. The *mce1* operon also contains *fadD5*, a gene involved in fatty acids degradation.

Several studies have shown that the different *mce* operons in *M. tuberculosis* are expressed at different times during growth and infection, with the *mce4* operon having particular importance in late persistent infection (Haile et al., 2002; Kumar et al., 2003). In general, the Mce proteins are prominent during growth of *M. tuberculosis* in stationary phase (Singh et al., 2016). This *M. tuberculosis* operon structure is fairly conserved in Actinobacteria, but the number of operons per species can vary between 1 and 12 (Casali and Riley, 2007).

As the name suggests, several studies indicate that Mce proteins have a direct function in mammalian cell entry. Beads coated with *M. tuberculosis* proteins Mce1A or Mce3A have the ability to enter HeLa cells, and both Mce1A and Mce4A proteins allow the entry of non-pathogenic *E. coli* (Arruda et al., 1993; Chitale et al., 2001; El-Shazly et al., 2007; Saini et al., 2008). The Mce2A protein cannot aid entry as it lacks the "cell entry epitope" found in Mce1A, Mce3A and Mce4A, suggesting its function might differ from the other proteins (Chitale et al., 2001; Mitra et al., 2005). Despite their apparent role in mammalian cell entry, mutants disrupted in the *mce1* genes in *M. tuberculosis* show no decreased ability to enter and survive in mammalian cells (Marjanovic et al., 2010; Shimono et al., 2003). This suggests that the links to mammalian cell entry could be an artifact of inserting *M. tuberculosis* proteins into *E. coli* and the artificial coupling of the proteins to beads.

Several studies have suggested that Mce proteins in Actinobacteria are important during infection. The presence of antibodies against Mce proteins in patients with tuberculosis indicates that they are expressed during infection and that they are immunogenic (Ahmad et al., 1999; Ahmad et al., 2004). Some Mce proteins have been shown to directly induce an inflammatory response and *mce* mutants in *M. tuberculosis* induce less of an inflammatory response than the WT in mice (Marjanovic et al., 2010; Shimono et al., 2003; Xu et al., 2007; Xu et al., 2008; Xue et al., 2007). In *M. tuberculosis*, mutants in either the *mce2, 3* or *4* operons are attenuated in mice (Marjanovic et al., 2010; Senaratne et al., 2008). In contrast, an *mce1* mutant in *M. tuberculosis* is hypervirulent

(Lima et al., 2007; Shimono et al., 2003; Uchida et al., 2007). Similarly, a *Streptomyces mce* operon mutant is hypervirulent in the amoebae vacuole (Clark et al., 2013). However, the spores of the *Streptomyces* strain lacking the *mce* operon are more wrinkled and weak, and the mutant cannot colonise plants as well as the parent stain (Clark et al., 2013).

Whilst all of these results indicate that Mce proteins are important for infection and/or colonisation, there is no indication of the specific function of these proteins in Actinobacteria. Some studies of *M. tuberculosis* propose a more specific role of Mce proteins as lipid transporters. An *mce1* mutant in *M. tuberculosis* accumulates mycolic acid in its cell wall and is less efficient at the uptake of palmitic acid than the parent (Cantrell et al., 2013; Forrellad et al., 2014). Strains disrupted in the *mce2* operon accumulate sulpholipids and are therefore likely to be involved in sulpholipid degradation (Marjanovic et al., 2011). The 3rd MCE operon is regulated by Mce3R, which also regulates expression of genes involved in lipid metabolism, including fatty acid degradation (de la Paz Santangelo et al., 2009). The *mce4* operons of *M. tuberculosis*, *M. smegmatis* and *Rhodococcus* have all been linked to cholesterol uptake (Klepp et al., 2012; Mohn et al., 2008; Pandey and Sassetti, 2008). Cholesterol utilisation is an important step in the survival of *M. tuberculosis* in macrophages, explaining why the *mce4* operon in particular is important for persistent infection in mice (Pandey and Sassetti, 2008).

1.5.2. Mammalian cell entry domains in Proteobacteria

The most recognised MCE domain containing protein in Proteobacteria is MlaD, which contains a single MCE domain and is the substrate-binding component of the Mla retrograde phospholipid trafficking pathway in *E. coli* (Malinverni and Silhavy, 2009). Predicted MlaD-like ABC transporters are found across Proteobacteria. In *Sphingobium japonicu*m the transporter is required for the uptake of gamma-hexachlorohexane, which it uses as its sole carbon source (Endo et al., 2007). In *Shigella flexneri*, it is required for multiplication and spreading within epithelial cells (Hong et al., 1998). In *Pseudomonas putida* the operon is hypothesised to function by altering phospholipid composition of the

outer membrane, which makes it essential for resistance to toluene (Kim et al., 1998). In *Neisseria meningitidis*, the transporter is required for uptake of L-glutamate, which is essential for survival in low sodium conditions (Monaco et al., 2006).

Proteobacteria is the only phylum in which proteins containing multiple MCE domains have been studied. The best-studied multi MCE domain containing protein is multivalent adhesion molecule- 7 (MAM-7) in *Vibrio parahaemolyticus*, in which there are seven MCE domains (Krachler et al., 2011). MAM-7 is reported to be an adhesin involved in early infection that binds to mammalian cells via phosphatidic acid and fibronectin (Krachler et al., 2011; Krachler and Orth, 2011). Mutants in MAM-7 are less able to adhere to mammalian cells and are less cytotoxic than the parent strain (Krachler et al., 2011). More recent research has identified that once MAM-7 establishes attachment with the host cell it can induce signalling processes including the activation of RhoA, which is involved in cytoskeleton rearrangement (Lim et al., 2014). This rearrangement increases the permeability of the intestinal epithelial barrier allowing the bacteria to establish deeper infection. Studies have also focused on the use of MAM-7 as an anti-adhesion therapy. The attachment and cytotoxic effects of clinically isolated pathogens on mammalian cells can be decreased by the addition of bacteria overproducing MAM-7 or by beads coupled to the protein (Hawley et al., 2013; Krachler et al., 2012a; Krachler et al., 2012b).

Despite their similarity to MAM-7, research into other multi-MCE domain containing proteins is sparse and the functions of those found in *E. coli* are currently unknown. The two multi-domain proteins in *E. coli* are PqiB and YebT, the latter of which is a MAM-7 orthologue. Given the literature on MCE domain containing proteins, PqiB and YebT could be involved in a range of cellular functions from inducing mammalian cell entry and playing a key role in infection to lipid transportation and maintenance of the cell membrane.

1.6. Aims

In *E. coli*, there are three genes predicted to encode MCE domain-containing proteins: *mlaD; pqiB;* and *yebT* (UniProt Consortium, 2015). There are predicted to be one, three and seven MCE domains in MlaD, PqiB and YebT, respectively. While the function of MlaD has been studied to some extent, the functions of PqiB and YebT in *E. coli* are currently unknown. The main focus of this study is to investigate MCE domain containing proteins in Proteobacteria, with particular focus on the *E. coli* multi-domain proteins PqiB and YebT. The aims were as follows:

1) Conduct a large-scale bioinformatic study on all MCE domain containing proteins to determine their prevalence, co-occurring protein domains, neighbouring genes, and evolution with particular focus on Proteobacteria
2) Identify a phenotype for *pqiB* and *yebT* mutants in *E. coli* and determine the cellular locations of PqiB and YebT to provide an insight into their functions
3) Investigate the role of MlaD, PqiB and YebT as components of lipid transporters by testing their affinity to lipids and whether they have a role in maintaining the lipid composition of the cell membranes
4) Solve the structure of PqiB, as no structure for an MCE domain containing protein currently exists
5) Use *Salmonella* Typhimurium, in which MlaD, PqiB and YebT are also found, to investigate the importance of MCE domain containing proteins in murine infection.

CHAPTER 2

MATERIALS AND METHODS

2.1. Culture media, growth conditions and strains

Strains were inoculated into lysogeny broth (LB) consisting of 10 g/l tryptone, 10 g/l NaCl and 5 g/l yeast extract. Overnight cultures were always 5 ml in volume grown in a 20 ml universal bottle. All other liquid cultures were grown in Erlenmeyer flasks with 1:10 culture to flask volume ratio. Liquid cultures were incubated in a shaking incubator at 180 rpm. For LB agar (LA), 1.5% (w/v) nutrient agar was added to LB. For growth on a sole carbon source, M9 minimal medium was used (5 x M9 salts from Sigma Aldrich) and supplemented with 200 mM $MgSO_4$, 1 mM $CaCl_2$ and the desired carbon source (e.g. 10 mg/ml glucose). Unless otherwise stated, cultures were incubated at 37°C, or 30°C for cultures retaining a temperature sensitive plasmid. If required, medium was supplemented with 50 μg/ml kanamycin sulfate, 100 μg/ml carbenicillin disodium salt or 100 μg/ml chloramphenicol.

2.1.1. Dilution plates

Square petri dishes (120 mm x 120 mm, Greiner Bio One) were filled with 50 ml of LA and the desired chemical. Overnight cultures were adjusted to an OD_{600} of 1, and serially diluted (1 in 10) to 10^{-7} in a 96-well plate (Greiner Bio One). Per strain, 2 μl of each dilution was pipetted on the square plate using a multichannel pipette.

2.1.2. Bacterial strains and plasmids

Table 2.1 shows a list of strain used in this study, and table 2.2 shows a list of plasmids.

2.2. Molecular genetic techniques

Unless otherwise specified, all kits were used following the manufacturers' instructions.

Table 2.1. Strains used in this study.

Strain	Genotype	Source
E. coli K-12 BW25113	*F^- λ^- rrnB3 ΔlacZ4787 hsdR514 Δ(araBAD)567 Δ(rhaBAD)568 rph-1*	Baba et al., 2006
BW25113 *pqiAB::aph*	*E. coli* K-12 BW25113 with *pqiAB* disrupted by a kanamycin cassette	This study
BW25113 *ΔpqiAB*	*E. coli* K-12 BW25113 with a *pqiAB* deletion	This study
BW25113 *yebST::aph*	*E. coli* K-12 BW25113 with *yebST* disrupted by a kanamycin cassette	This study
BW25113 *ΔyebST*	*E. coli* K-12 BW25113 with a *yebST* deletion	This study
BW25113 *ΔpqiAB, yebST::aph*	*E. coli* K-12 BW25113 with a *pqiAB* deletion and *yebST* disrupted by a kanamycin cassette	This study
BW25113 *ΔpqiAB ΔyebST*	*E. coli* K-12 BW25113 with a *pqiAB* and *yebST* deletions	This study
BW25113 *mlaD::aph*	*E. coli* K-12 BW25113 with *mlaD* disrupted by a kanamycin cassette	This study
BW25113 *ΔpqiAB mlaD::aph*	*E. coli* K-12 BW25113 with a *pqiAB* deletion and *mlaD* disrupted by a kanamycin cassette	This study
BW25113 *ΔyebST mlaD::aph*	*E. coli* K-12 BW25113 with a *yebST* deletion and *mlaD* disrupted by a kanamycin cassette	This study
BW25113 *ΔpqiAB ΔyebST mlaD::aph*	*E. coli* K-12 BW25113 with a *pqiAB* and *yebST* deletion and *mlaD* disrupted by a kanamycin cassette	This study
BW25113 *ΔmlaD*	*E. coli* K-12 BW25113 with an *mlaD* deletion	This study
BW25113 *ΔmlaD ΔpqiAB*	*E. coli* K-12 BW25113 with an *mlaD* and *pqiAB* deletion	This study
BW25113 *ΔmlaD ΔyebST*	*E. coli* K-12 BW25113 with an *mlaD* and *yebST* deletion	This study
BW25113 *ΔmlaD ΔpqiAB ΔyebST*	*E. coli* K-12 BW25113 with an *mlaD, pqiAB* and *yebST* deletions	This study
BW25113 *pqiA::aph*	*E. coli* K-12 BW25113 with *pqiA* disrupted by a kanamycin cassette	This study
BW25113 *pqiB::aph*	*E. coli* K-12 BW25113 with *pqiB* disrupted by a kanamycin cassette	This study
BW25113 *ymbA::aph*	*E. coli* K-12 BW25113 with *ymbA* disrupted by a kanamycin cassette	This study
BW25113 *mlaC::aph*	*E. coli* K-12 BW25113 with *mlaC* disrupted by a kanamycin cassette	This study

Strain	Genotype	Source
BW25113 *ΔpqiAB*, *mlaC::aph*	*E. coli* K-12 BW25113 with a *pqiAB* deletion and *mlaC* disrupted by a kanamycin cassette	This study
BW25113 *ΔpqiAB ΔyebST, mlaC::aph*	*E. coli* K-12 BW25113 with a *pqiAB* and *yebST* deletion and *mlaC* disrupted by a kanamycin cassette	This study
E. coli BL21 DE3	DE3 phage T7 protein expression strain	Invitrogen
Salmonella Typhimurium SL1344	*Salmonella enterica* subsp. *enterica* serovar Typhimurium	Wray et al.,1978
SL1344 *pqiAB::aph*	*Salmonella* Typhimurium SL1344 with *pqiAB* disrupted by a kanamycin cassette	This study
SL1344 *yebST::aph*	*Salmonella* Typhimurium SL1344 with *yebST* disrupted by a kanamycin cassette	This study
SL1344 *ΔpqiAB*	*Salmonella* Typhimurium SL1344 with a deletion in *pqiAB*	This study
SL1344 *ΔyebST*	*Salmonella* Typhimurium SL1344 with a deletion in *yebST*	This study
SL1344 *ΔpqiAB, yebST::aph*	*Salmonella* Typhimurium SL1344 with a deletion in *pqiAB* and *yebST* disrupted by a kanamycin cassette	This study
SL1344 *ΔpqiAB ΔyebST*	*Salmonella* Typhimurium SL1344 with a deletion in *pqiAB* and *yebST*	This study
SL1344 *mlaD::aph*	*Salmonella* Typhimurium SL1344 with *mlaD* disrupted by a kanamycin cassette	This study
SL1344 *ΔpqiAB ΔyebST, mlaD::aph*	*Salmonella* Typhimurium SL1344 with a deletion in *pqiAB* and *yebST* and *mlaD* disrupted by a kanamycin cassette	This study

Table 2.2. Plasmids used in this study.

Plasmid	Properties	Source
pKD4	Contains kanamycin resistance cassette flanked by FLP sites. Kan^R	Datsenko and Wanner, 2000
pKD46	Temperature sensitive plasmid with arabinose inducible promotor. Contains λ Red recombination machinery. Amp^R	Datsenko and Wanner, 2000
pCP20	Temperature sensitive plasmid with FLP recombinase. Amp^R	Datsenko and Wanner, 2000
pET17b	Constitutively active T7 expression vector, IPTG inducible, Amp^R	Novagen
pET22b (+)	T7 protein expression vector, N terminal his tag, IPTG inducible, Amp^R	Novagen

2.2.1. Preparation of DNA

Plasmid DNA was isolated from an overnight culture using the Qiagen QIAprep Spin Miniprep kit. For isolation of genomic DNA, the STRATEC RTP Bacteria DNA Mini kit was used following the protocol for extraction of DNA from a Gram-negative bacterial cell pellet.

2.2.2. Polymerase chain reaction

Primers used in this study are found in table 2.3. Fragments required for cloning and mutagenesis were amplified using Phusion High-Fidelity DNA Polymerase (New England Biolabs) and included the optional use of 5x GC buffer and 3% (v/v) DMSO. Unless otherwise specified the thermocycling conditions for Phusion PCR were as in table 2.4. PCR products amplified by Phusion were purified using the Qiagen QIAquick PCR purification kit.

For colony PCR, MyTaq Red Mix (Bioline) was used. A small amount of bacterial colony was added straight into the reaction mix. The thermocycling conditions included a 10-minute initial denaturation step (at 95^{o}C) to lyse the cells, and an annealing temperature of 55^{o}C (unless otherwise specified).

2.2.3. Agarose Gel Electrophoresis

DNA fragments were separated using 1% (w/v) agarose (Bioline) gels in 1 x TAE buffer (50x TAE buffer = 2 M Tris, 1 M acetic acid, 0.05 M EDTA in water) with Midori Green (Nippon Genetics) as the DNA dye. DNA samples were loaded using 5 x DNA loading buffer Blue (Bioline) at a 1:4 buffer:sample ratio, with the exception of PCR products from MyTaq Red Mix, which includes a dye. In all cases Hyperladder 1 kb (Bioline) was used as a marker. Gels were run in 1 x TAE buffer at 100 – 120 V until the loading dye had run sufficiently. Analytical gels were viewed under UV light (300 nm UV light) in a Gel Doc (Bio-Rad), and preparative gels were viewed in a blue light LED illuminator (Geneflow) to prevent DNA damage.

Table 2.3. Primers used in this study.

Primer	Sequence (5′ to 3′)	Description
pqiA_DW_F	AGTCATCTGTAAAATAGCGCATCATTAAGGAGTACCAA TGGTGTAGGCTGGAGCTGCTTC	Forward primer for amplification of kanamycin resistance cassette *pqiAB* gene deletion
pqiB_DW_R	ATCGTCACTAGCCACTTTTTCATTGTTTCGCCCTCTTC GGCATATGAATATCCTCCTTAG	Reverse primer for amplification of kanamycin resistance cassette for *pqiAB* gene deletion
pqiA_check_F	AACAGAGACAGCTGACAACG	Forward flanking primer to check *pqiAB* gene deletion in *E. coli* BW25113
pqiB_check_R	TGTTGCGCAACTGTTGATCC	Reverse flanking primer to check *pqiAB* gene deletion in *E. coli* BW25113
yebS_DW_F	TAAGATTAATGTTAATTCTTATTACATTTGGCACGTCAT GGTGTAGGCTGGAGCTGCTTC	Forward primer for amplification of kanamycin resistance cassette for *yebST* gene deletion in *E. coli* BW25113
yebT_DW_R	CACGCCGGAGCAGTGGGCATTATTTGGGAAGCGCAG TTCCCATATGAATATCCTCCTTAG	Reverse primer for amplification of kanamycin resistance cassette for *yebST* gene deletion in *E. coli* BW25113
yebS_check_F	TAACAACGCACTGGTGTTCG	Forward flanking primer to check *yebST* gene deletion in *E. coli* BW25113
yebT_check_R	CAGCGTATTAACGCGAATGC	Forward flanking primer to check *yebST* gene deletion *E. coli* BW25113
mlaD_check_F	TGGAAAGGCATTGATAGCGG	Forward flanking primer to check *mlaD* gene deletion in *E. coli* BW25113
mlaD_check_R	TCTTGTAATACTGGCCCAGC	Reverse flanking primer to check *mlaD* gene deletion in *E. coli* BW25113
pqiA_NdeI_F	GGCGGCCATATGTGCGAACATCATCATGCCGC	Forward primer for generation of a *pqiA/pqiAB* fragment for cloning into pET17b/pET22b(+)
pqiB_XhoI_R	CGAAGCCTCGAGTTGTTTCGCCCTCTTCGGCT	Reverse primer for generation of a *pqiB* fragment for cloning into pET22b(+)
pqiB_comp_R	TCACGAAGCCTCGAGTTGTTTCGCCCTCTTCG	Reverse primer for generation of a *pqiAB* fragment for cloning into pET17b
yebS_NdeI_F	GGCGGCCATATGGCTCTTAACACACCACAAAT	Forward primer for generation of a yebS/ *yebST* fragment for cloning into pET17b/ pET22b(+)
yebT_XhoI_R	CGAAGCCTCGAGTTTGGGAAGCGCAGTTCCCC	Reverse primer for generation of a *yebT* fragment for cloning into pET22b(+)
yebT_comp_R	TCACGAAGCCTCGAGTTTGGGAAGCGCAGTTC	Reverse primer for generation of a *yebST* fragment for cloning into pET17b
mlaD_NdeI_F	CATATGATGCAAACGAAAAAAAATGAA	Forward primer for generation of a *mlaD* fragment for cloning into pET17b/pET22b(+)
mlaD_XhoI_R	CTCGAGTTATTTCGTTGTACCCACAGG	Reverse primer for generation of a *mlaD* fragment for cloning into pET17b/pET22b(+)
T7_F	TAATACGACTCACTATAGGG	Forward primer used for colony PCR/ sequencing of pET17b/pET22b(+) constructs
T7_R	GCTAGTTATTGCTCAGCGG	Reverse primer used for colony PCR/ sequencing of pET17b/pET22b(+) constructs
ymbA_check_F	AGCCTACAACAAGATGGTGG	Forward flanking primer to check *ymbA* gene deletion in *E. coli* BW25113
ymbA_check_R	CACAGTTAAGCAATACCGCG	Reverse flanking primer to check *ymbA* gene deletion in *E. coli* BW25113

Primer	Sequence (5′ to 3′)	Description
mlaC_check_F	ACTGCTATCCTGAAGGATGG	Forward flanking primer to check *mlaC* gene deletion in *E. coli* BW25113
mlaC_check_R	CCAGGGTATACACTTTGTCG	Reverse flanking primer to check *mlaC* gene deletion in *E. coli* BW25113
pqiA_XhoI_R	CGAAGCCTCGAGGGACTCCTCATGCTCTGATT	Reverse primer for generation of a *pqiA* fragment for cloning into pET22b(+)
pqiB_NdeI_F	GGCGGCCATATGGAATCTAATAATGGGGAAGC	Forward primer for generation of a *pqiB* fragment for cloning into pET22b(+)
pqiB_TM_NdeI_F	GGCGGCCATATGCAGGGACCGGAAGTGACCCT	Forward primer for generation of a *pqiB* fragment without the first 120 bps for cloning into pET22b(+)
yebS_XhoI_R	CGAAGCCTCGAGGTCGTCGAAGCGGGCGTTTC	Reverse primer for generation of a *yebS* fragment for cloning into pET22b(+)
yebT_NdeI_F	GGCGGCCATATGAGTCAGGAAACGCCCGCTTC	Forward primer for generation of a *yebT* fragment for cloning into pET22b(+)
yebT_TM_NdeI_F	GGCGGCCATATGAGTTATCAGGACCGGGGTAA	Forward primer for generation of a *yebT* fragment without the first 120 bps for cloning into pET22b(+)
pqiA_GD_F	AAGCTTGGATCCAGGCCTGAAAAACGGGGCATAACGAAAGGAGCGCCTATGGACCGGTCAATTGGCTGGAG	Forward primer for amplification of kanamycin resistance cassette for *pqiAB* deletion in *S.* Typhimurium SL1344
pqiB_GD_R	CATATGGCTAGCTCACGACTAGCCATTTTTTCATTGTTTCGCCCTCTTAGG	Reverse primer for amplification of kanamycin resistance cassette for *pqiAB* deletion in *S.* Typhimurium SL1344
pqiA_STm_check_F	AATCGGACGTTATATCGGCG	Forward flanking primer to check *pqiAB* gene deletion in *S.* Typhimurium SL1344
pqiB_STm_check_R	GTTTAACGTATCCTGGGTGG	Reverse flanking primer to check *pqiAB* gene deletion in *S.* Typhimurium SL1344
yebS_GD_F	AAGCTTGGATCCAAGATTACATTATTCTAACACTTTGATTCACGATACATGGACCGGTCAATTGGCTGGAG	Forward primer for amplification of kanamycin resistance cassette for *yebST* deletion in *S.* Typhimurium SL1344
yebT_GD_R	CATATGGCTAGCTACGCCGGAGCCTGGTGTTTAACGTGGCAGAGCGGTGCCAATATCCTCCTTAGTTCC	Reverse primer for amplification of kanamycin resistance cassette for *yebST* deletion in *S.* Typhimurium SL1344
yebS_STm_check_F	CGTTGAACCTTATTCTGCGC	Forward flanking primer to check *yebST* gene deletion in *S.* Typhimurium SL1344
yebT_STm_check_R	TAAATGCTCAGCGGTACTGC	Reverse flanking primer to check *yebST* gene deletion in *S.* Typhimurium SL1344
mlaD_STm_DW_F	TTTCATTGCCGCTACTGA	Forward primer for generation of kanamycin cassette for construction of *mlaD* gene deletion is *S.* Typhimurium SL1344
mlaD_STm_DW_R	CGCTATCTGGTAAGTCTGACC	Reverse primer for generation of kanamycin cassette for construction of *mlaD* gene deletion is *S.* Typhimurium SL1344
mlaD_STm_check_F	CTCCAGTATGGAGATGATGG	Forward flanking primer to check *mlaD* gene deletion in *S.* Typhimurium SL1344
mlaD_STm_check_R	TACACGGATAGGCACAATG	Reverse flanking primer to check *mlaD* gene deletion in *S.* Typhimurium SL1344

Table 2.4. Thermocycling conditions used for Phusion PCR.

Cycles	Step	Temperature	Time
1	Initial denaturation	98°C	5 minutes
30	Denaturation	98°C	10 seconds
	Annealing	55°C	30 seconds
	Extension	72°C	30 seconds/kb
1	Final extension	72°C	5 minutes

2.2.4. Cloning

Purified PCR fragments and plasmids were digested with two restriction enzymes (most commonly NdeI and XhoI) using FastDigest enzymes (Thermo Scientific). Digested products were separated by agarose gel electrophoresis and extracted using Qiagen QIAquick Gel Extraction kit. The vector was treated with Antarctic Phosphatase (New England Biolabs) to prevent vector re-ligation. DNA was quantified using Nanodrop and the insert was ligated into the vector at a 3:1 ratio with 50 ng of vector DNA, using T4 ligase (NEB). After incubation at 16°C overnight, 5 µl of the ligation mixtures were transformed into NEB 5-alpha Competent *E. coli* (High Efficiency) and plated onto LA with appropriate antibiotics.

2.3. Bacterial transformation

2.3.1. Chemical competence

The desired strain was inoculated into 5 ml LB and grown overnight at 37°C. The cells were diluted 1:100 into 100 ml of fresh LB and grown to OD_{600} of 0.5. The cells were stored on ice for 30 min and centrifuged for 10 min at 6000 *g* and 4°C. The pellet was re-suspended in 12.5 ml of ice cold 0.1 M $CaCl_2$ and stored on ice for 30 min. The cells were centrifuged again and re-suspended in 1 ml of 0.1 M calcium chloride and stored on ice for 1 hour.

For transformation, 50 ng of vector DNA was mixed with 50 µl of competent cells in a 1.5 ml microcentrifuge tube. The mixture was heat shocked for 90 s in a 42°C water bath and placed on ice for 5 min. 1 ml of LB was added and the cultures were aerated at the appropriate temperature for 1 hour (usually 37°C, but 30°C for temperature sensitive vectors). The cells were pelleted at 6000 *g* for 5 min and re-suspended in 100 µl of LB. The transformation was plated onto LA with the appropriate antibiotic and incubated overnight.

2.3.2. Electroporation

An overnight culture was diluted 1: 100 into 100 ml of fresh LB and grown to an OD_{600} of 0.6. The cultures were placed on ice for 30 min and centrifuged at 6000 *g* for 10 min at 4°C.

The pellet was wash three times by resuspension and centrifugation with decreasing volumes of ice cold H_2O (100 ml, 50 ml and 10 ml), and were finally re-suspended in 1 ml of H_2O.

For transformation, 50 μl of cells were mixed with 50 ng of plasmid DNA or 300 ng of linear DNA. The mixture was transferred to a 2 mm electroporation cuvette (Cell Projects) and pulsed at 2200V, after which 1 ml SOC medium (New England Biolabs) was added immediately. The mixture was transferred to a test tube for recovery for 1 hour in a shaking incubator at either 37°C or 30°C. The transformation was then plated onto LA with the appropriate antibiotic and incubated overnight at 37°C or 30°C.

2.4. Chromosomal genetic engineering

2.4.1. Datsenko and Wanner method for gene inactivation

The gene of interest was replaced with a linear fragment containing an antibiotic resistance cassette, using λ Red recombination as described by Datsenko and Wanner (2000). The steps in this protocol are summarised in figure 2.1. In the first step, a linear fragment containing a kanamycin resistance (*aph*) cassette with flanking homology to the region up and downstream of the gene of interest was generated by PCR with the plasmid pKD4 as a template. The 50 bp at the 5′ end of the forward primer were homologous to the region upstream of the gene of interest (H1). The next 20 bp were homologous to 5′ end of the *aph* cassette on pKD4 (P1). The 50 bp at 5′ end of the reverse primer were homologous to the last seven codons of the gene of interest and the region downstream (H2). The next 20 bp of the reverse primer were homologous to the end of *aph* cassette on pKD4 (P2). Retaining the last seven codons prevented disruption of any overlapping downstream gene or DNA binding sites.

The strain of interest was made chemically competent and transformed with temperature sensitive plasmid pKD46, which contains an arabinose inducible λ Red recombinase, and incubated at 30°C on LA supplemented with carbenicillin. The strain containing pKD46 was made electrocompetent (growth at 30°C with carbenicillin, and 0.2%

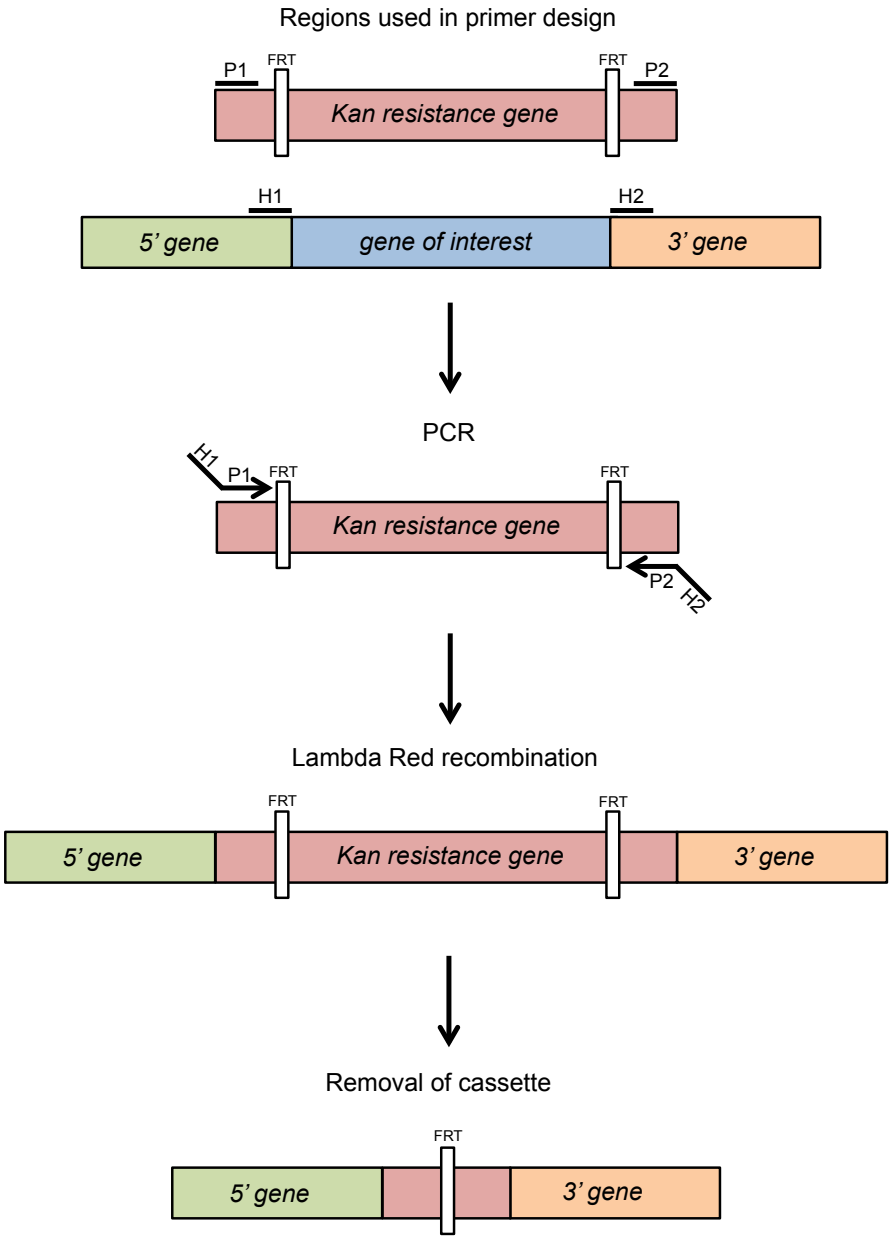


Figure 2.1. The main steps for gene inactivation using lambda red recombination. In the first step, a linear fragment containing a kanamycin resistance (*aph*) cassette with flanking homology to the region up (H1) and downstream (H2) of the gene of interest is generated. The target strain is transformed with the linear DNA fragment to recombine with the gene of interest on the chromosome.

L-arabinose for induction of recombination machinery expression) and transformed with the linear DNA fragment and recovered at 37°C. Half of the transformation was plated on LA supplemented with kanamycin and incubated overnight at 37°C. The other half was left overnight at room temperature for subsequent plating if there were no colonies on the first plate. Resultant colonies were re-streaked LA supplemented with kanamycin to purify the potential knockouts from dead WT cells. The colonies were checked by colony PCR using check primers homologous to between 100 and 500 bp either side of the gene of interest. A change in band size from WT to kanamycin cassette confirmed successful gene replacement.

2.4.2. P1 transduction

P1 transduction was used to transfer mutations made by λ Red recombination into *E. coli* strain BW25113, or into a strain constructed during this project. The method used is based on the protocol described by Thomason et al. (2007). An overnight 5 ml culture of the donor strain was diluted 1: 100 into 5 ml LB containing 0.2% glucose and 5 mM $CaCl_2$ and incubated with shaking for 45 min at 37°C. The culture was infected with 100 µl of recently prepared virulent P1 stock and incubation was continued until lysis occurred (~3 h). To kill any un-lysed bacteria, 200 µl of chloroform was added and incubated for a further 5 min. The liquid was centrifuged at 10 000 *g* for 10 min at 4°C and the supernatant (p1 lysate) was stored at 4°C until required.

Two 1.5 ml samples of an overnight culture of the recipient strain were pelleted by centrifugation in a micro centrifuge tube at maximum speed for 2 min at room temperature. The pellet was re-suspended in 750 µl of P1 salts (10 mM $CaCl_2$, 5 mM $MgSO_4$) and 100 µl of cells were mixed with 0, 1, 10 and 100 µl of P1 lysate in glass tubes and left at room temperature for 30 min. LB (1 ml) and 0.2 mM sodium citrate was added to each transduction and recovered at 37°C for 1 hour. The cells were then pelleted in micro centrifuge tubes at maximum speed for 2 min at room temperature. The pellet was re-suspended in 100 µl of LB and plated onto LA containing kanamycin and 5 mM sodium citrate and incubated overnight at 37°C. The plate that showed colonies from the least

amount of phage was selected and colonies were screened by colony PCR using check primers between 100 and 500 bp up either side of the gene of interest.

2.4.3. Removal of kanamycin cassette

To remove the kanamycin resistance gene, confirmed mutant strains were made chemically competent and transformed with the temperature sensitive plasmid pCP20, and incubated at 30°C with carbenicillin. The plasmid pCP20 contains the yeast FLP recombinase gene that recombines the two FLP sites flanking the kanamycin resistance gene, removing the gene and leaving a small scar on the chromosome. Colonies containing pCP20 were inoculated into 5 ml of LB for 6 h at 42°C to cure the bacteria of the temperature sensitive plasmid. A small volume of this culture was spread onto plain LA to give single colonies after growth overnight at 42°C. The colonies were patch plated onto LA, LA supplemented with carbenicillin and LA supplemented with kanamycin to check for loss of pCP20 and the kanamycin cassette. Colonies that grew only on plain LA were checked by colony PCR to confirm removal of the kanamycin resistance cassette.

2.5. Sequencing

2.5.1. Genome sequencing

Strains for sequencing were sent to MicrobesNG sequencing service, University of Birmingham, which uses an Illumina MiSeq or HiSeq platform. The service included SNP and short indel calling compared to the chosen reference strain and aligned reads were browsed manually using Artemis Comparison Tool to confirm SNPs, indels and gene deletions.

2.5.2. Plasmid sequencing

Plasmids were sequenced using the Functional Genomic Unit, University of Birmingham. In a 10 µl sequencing reaction there was 200-500 ng of vector DNA and 3.2 pmol of sequencing primer. The service uses the capillary sequencer ABI 3730.

2.6. Protein analysis

2.6.1. SDS-PAGE analysis

In this study all protein samples were analysed on 10% resolving and 6% acrylamide stacking gels made in 1 mm gel casts (BioRad). Per 10 ml of resolving gel there were 3.3 ml of 30% (w/v) acrylamide (Protogel), 2.5 ml of resolving buffer (Protogel), 4.2 ml of H_2O, 180 µl of 10% (w/v) ammonium persulphate (APS) and 24 µl of TEMED. Per 10 ml of stacking buffer there were 1.7 ml of 30% (w/v) acrylamide (Protogel), 2.5 ml of stacking buffer buffer (0.5 M Tris, pH 6.8), 5.8 ml of H_2O, 100 µl of 10% (w/v) ammonium persulphate (APS) and 14 µl of TEMED.

Protein samples were loaded in 2x Laemmli sample buffer (Sigma Aldrich). For whole cell lysates an equivalent of 1 ml of culture with an OD_{600} of 1 were pelleted in a microcentrifuge tube at maximum speed for 2 min. The pellet was re-suspended in 100 µl of Laemmli buffer and boiled for 15 min to lyse the cells. All other sample types were loaded in a 1: 1 ratio with the sample buffer without boiling. For 10 and 15 well gels a maximum of 20 µl and 10 µl were loaded, respectively. Gels were run at 120 V until the loading dye reached the bottom of the gel. Gels were stained with R-250 Coomassie Brilliant Blue (Thermo Fisher Scientific) for 30 min and destained with destain solution (10% (v/v) glacial acetic acid, 40% (v/v) methanol, 50% (v/v) H_2O).

2.6.2. Protein expression

For expression of a particular protein, the desired gene was cloned into the pET22b, which harbours a C-terminal polyhistine tag, and was transformed into chemically competent *E. coli* BL21 DE3. Transformants were incubated overnight in LB supplemented with carbenicillin and were subsequently diluted 1: 100 for further growth. The cells were grown to an OD_{600} of 0.6 and induced with IPTG (usually at a final concentration of 1 mM). Cultures were left for a further 3 h at 37°C (to produce protein for antibodies) or overnight at 18°C (to produce protein for crystallography).

2.6.3. Protein purification

Conditions for protein purification varied depending on whether the protein was soluble or insoluble (membrane located). For soluble protein, the induced cells were pelleted at 16000 *g* for 10 min at 4°C. The pellet cells were re-suspended (3 ml/g pellet) in binding buffer (50mM NaP pH 7, 500 mM NaCl and 50mM imidazole) and broken in a C3 cell disrupter until visible lysis had occurred. To pellet unbroken cells the suspension was centrifuged at 17400 *g* for 20 min at 4°C. The supernatant was bound to either a 1 ml (for antibodies) or 5 ml (for crystallography) HisTrap column (GE healthcare) overnight, using a peristaltic pump. After binding, the column was washed twice using binding buffer (5x column volume). The protein was eluted in 5 x 1 ml fractions (for 1 ml column) or 10 x 1 ml fractions (for 5 ml column) in elution buffer (50 mM NaP pH 7, 500 mM NaCl and 500 mM imidazole). At all stages, Complete EDTA-free Protease Inhibitor Cocktail Tablets (Roche) were used to prevent protein degradation. The flow-through, washes and elutions were analysed by SDS-PAGE to determine the success of the purification.

For membrane-localised proteins, the purification protocol was altered. Following the pelleting of unbroken cells after lysis, the supernatant was spun again at 48400 *g* for 1 hour at 4°C, to pellet cell membranes. The membrane was solubilised overnight shaking at 4°C in binding buffer and 0.5% n-Dodecyl β-D-maltoside (DDM) with 1 ml of buffer per 40 mg of membrane. The remaining steps were as for the soluble protein, except the wash and elution buffers contained 0.1% DDM.

2.6.4. Generation of polyclonal antibodies

For generation of antibodies, soluble protein was purified as described in the previous section. The protein was sent to Eurogentec for the Anti-protein 28-day Speedy polyclonal package, where antibodies are raised in rabbits. This protocol involved injection of 200 μg of protein into two rabbits on days 0, 7, 10 and 18, with final serum collect on day 28. The final serum was used to detect the desired protein by western blot.

2.6.5. Western blotting

Samples for western blotting were separated using SDS-PAGE. The protein was transferred to a nitrocellulose membrane using the iBlot 2 Dry Blotting System (Thermo fisher). Membranes were blocked in 5% (w/v) milk buffer (per l: 50 g skimmed milk, 2.42 g Tris base, 8 g NaCl, pH 8.4) for 1 hour at room temperature. The membrane was incubated overnight at 4°C in primary antibody (diluted in 5% (w/v) milk buffer at the required ratio). The membranes were washed three times for 5 min in TBS-T (per l: 1 ml Tween-20, 2.42 g Tris base, 8 g NaCl, pH 8.4). For secondary incubation, the HRP secondary antibody (Sigma Aldrich) was added at a ratio of 1: 15000 into 5% (w/v) milk buffer and incubated with the membrane for 1 hour. The membrane was washed a further 3 times in TBS-T. The western blot was developed using ECL Prime Western Blotting Detection Reagent (Amersham), and exposed to Hyperfilm ECL (Amersham) for between 5 s and 5 min.

2.7. Preparation of cellular fractions

2.7.1. Separation of soluble and insoluble cellular compartments

Two litres of cells were grown to an OD_{600} of 0.6-0.8. The cells were pelleted at 16000 *g* for 10 min at 4°C and re-suspended in 3 ml/g of phosphate buffered saline (PBS, per l: 8 g NaCl, 0.2 g KCl, 1.42 g Na_2HPO_4, 0.24g KH_2PO_4). The cells were broken in a C3 cell disrupter (at 15000 Psi) and unbroken cells were pelleted at 17400 *g* for 20 min at 4°C. The supernatant was spun again at 48400 *g* for 1 hour at 4°C to pellet cell membranes and leave the soluble cell components in the supernatant.

For determination of protein location by western blot, the pellet was re-suspended in 5 ml of PBS and the supernatant was concentrated down to 5 ml using a Vivaspin 500 concentrator (GE Healthcare, MWCO 5000). The protein concentration of each sample was determined by a Bradford Protein Assay. For the western blot, 10 μg and 50 μg of protein were analysed for the insoluble and soluble fractions, respectively, based on the estimated amount of protein found in the insoluble and soluble compartments of the cell (Ames, 1974).

2.7.2. Sarkosyl extraction for separation of inner and outer membranes

This protocol was adapted from Hobb et al., (2009). Membranes from 2 l of culture were isolated as described in the previous section and re-suspended in 80 ml of HEPES buffer (10 mM HEPES, pH 7.4) to wash the membranes. The membranes were re-pelleted at 10000 *g* for 1 hour at 4°C. The pellet was re-suspended in 40 ml HEPES buffer and 1% (w/v) sarkosyl (Sigma Aldrich) and incubated shaking at 37°C for 30 min. The sarkosyl-treated membranes were pelleted again at 10000 *g* for 1 hour at 4°C. The outer membrane pellet was re-suspended in 5 ml of HEPES buffer, and the inner membrane supernatant was concentrated down to 5 ml using a Vivaspin 500 concentrator (GE Healthcare, MWCO 5000). The success of the separation was determined by western blot, using known inner and outer membrane protein markers against 10 μg of protein per sample.

2.7.3. Short protocol for separation of inner and outer membrane using a sucrose gradient

The method was adapted from online protocol for isolation of the outer membrane (cmdr website, 2000). Isolation of membranes was done as described previously, with an additional step before cell disruption. This was an osmotic shock step to remove periplasmic contents using a protocol described by Magnusdottir et al., (2009). Whole cells were re-suspended in 5 ml/g of sucrose buffer (50 mM HEPES, 20% (w/v) sucrose, 1 mM EDTA pH 7.9) and re-pelleted at 7000 *g* for 30 min at 4°C. The pellet was re-suspended again in 5 ml/g of 5 mM $MgSO_4$ and stored on ice for 10 min. The cells were pelleted again at 4500 *g* for 20 min at 4°C. This pellet was re-suspended in 3 ml/g of PBS and lysed in a C3 cell disrupter. The cell membranes were isolated as described previously and re-suspended in 5 ml of 20% (w/v) sucrose.

The sucrose gradient was set up with 4 ml 70% (w/v) sucrose (bottom), 4 ml 60%(w/v) sucrose (middle) and 4 ml 20% (w/v) sucrose containing the isolated membranes (top) made in 14 ml Ultra-Clear Thinwall tubes (Beckman Coulter). A total of 1 mg of protein was loaded per gradient. The gradients were ultracentrifuged at 285000 *g* for 16 h

at 4°C in a SW 40 Ti rotor. After centrifugation, the tube was pierced at the bottom and fractions of ~500 µl were obtained. The fractions were analysed by SDS-PAGE and western blot with known membrane markers to determine the success of inner and outer membrane separation.

2.7.4. Long protocol for separation of inner and outer membrane using a sucrose gradient

This protocol was adapted from Osborn and Munson, (1974) and Dalebroux et al., (2015). Before cell disruption and isolation of membranes spheroplasts were prepared. The cell pellet (from 2 l of cells at an OD_{600} of 0.6-0.8) was re-suspended in 10 ml of sucrose-Tris buffer (0.75 M sucrose, 10 mM Tris, pH 7.8). The mixture was transferred to an Erlenmeyer flask where 500 µl of 2 mg/ml lysozyme was added. After a 2 minute incubation on ice, 20 ml of ice cold 1.5 mM EDTA was slowly added over 10 min using a peristaltic pump, with gentle stirring. The spheroplasts were broken using a C3 cell disrupter, and the membranes were isolated as described previously. To wash the membranes, the pellet was re-suspended in 20 ml of sucrose-Tris-EDTA buffer 1 (0.25 M sucrose, 3.3 mM Tris, pH 7.8, 1 mM EDTA) and re-pelleted at 165 000 *g* for 1 hour at 4°C. The membranes were re-suspended again in 10 ml of sucrose-Tris-EDTA buffer 2 (20% (w/v) sucrose, 0.5 mM EDTA, 10 mM Tris, pH 7.8) and were ready for separation by sucrose gradient.

For the gradient, all sucrose was made up in EDTA-Tris buffer (0.5 mM EDTA, 10 mM Tris, pH 7.8). The gradient was made up in 38.5 ml Ultra-Clear Thinwall tubes (Beckman Coulter) with 10 ml of 73% (w/v) sucrose (bottom), 18 ml of 53% (w/v) sucrose (middle) and 10 ml of 20% (w/v) sucrose including the membrane sample (top). The gradient was centrifuged 141 000 *g* for 40 h in a SW 28 Ti rotor at 4°C. To obtain the inner membrane a pipette was used to withdraw the membrane through the top of the gradient from the 20%-53% boundary. To obtain the outer membrane the tube was pierced at the bottom and collected by gravity flow from the 53%-73% boundary. To assess the success of

separation, the isolated fractions were analysed by western blot using known membrane markers.

2.8. X-ray Crystallography of PqiB

2.8.1. Size exclusion to obtain cleaner purified protein

Purifed PqiB was separated from any contaminants by size exclusion using a Superdex 200 10/300 GL column on an AKTA Pure system. This gel filtration system elutes proteins based on sized. The buffer used throughout was 200 mM NaCl and 20 mM HEPES, pH 8 (plus 0.5% DDM for insoluble protein). The system was set up to elute in 2 ml fractions and measure the protein concentration of each fraction by absorbance at 280 nm. The fractions that contained high amounts of protein were analysed by SDS-PAGE and clean fractions of the correct molecular weight were pooled and concentrated for crystallography.

2.8.2. Crystal tray set up

After gel filtration, protein was concentrated using a Vivaspin 500 concentrator (GE Healthcare, MWCO 30000), to ~20 mg/ml. The following 96 condition screens were used to test for crystal formation: JCSGand, ProPlex, Morpheus HT-96, Midas HT-96, PACT premier (Molecular Dimensions) and PEGRx (Hampton Research). Sitting drops were set up at a 300: 300 nl protein: condition ratio using the mosquito Crystal automated liquid handler (TTP Labtech), with a 30 µl reservoir of each condition. The crystal trays were checked at regular intervals for crystal formation. Any condition that formed a crystal was scaled up manually in a 24 well plate with hanging drops of 2 : 2 µl protein: condition ratio, over a 1 ml reservoir. These scale-up trays were designed to cover a small range of pH's and precipitant concentration around the original condition in an attempt to repeat and improve crystal formation. Poor quality crystals and microcrystals were used for seeding crystal growth. The crystals were broken up by harsh vortexing in ~20 µl of the solution it formed in. This stock was used to seed back into the same condition manually (24 well) or

into a new screen using the mosquito. The drops were set up at a 3: 2: 1 ratio of protein: condition: seed stock in a 4 μl drop for manual trays and a 600 nl drop for mosquito trays.

2.8.3. Testing the diffraction of crystals

Crystals were picked with appropriate sized loops under a 10x microscope, frozen in liquid nitrogen and transported to Diamond Light Source by members of the Lovering research group, University of Birmingham. Crystals were picked with and without soaking in bromine to help overcome the phasing problem (see 5.1).

2.9. Lipidomics techniques

2.9.1. Extraction of lipids from purified protein

To extract lipid, 2 ml of chloroform and 1 ml of methanol was added to 2 ml of 1 mg/ml protein in a glass tube. The mixture was vortexed for 2 min and incubated at 50^{o}C for 30 min. The mixture was vortexed for another 2 min and centrifuged at 1000 RPM for 10 min in an IEC tabletop centrifuge forming a two-phase system (top=aqueous, bottom=organic). The lower phase was collected using a glass Pasteur pipette and transferred to a fresh glass tube. The tubes were placed in a heat block at 100^{o}C without lids until all liquid had evaporated. The lipids were re-dissolved in 200 μl of chloroform ready for thin layer chromatography.

2.9.2. One-directional thin layer chromatography

TLC plates (Silica gel 60 – Merck Millipore) were cut to 10 cm in length and a width that allowed 1 cm between each sample and the edge of the plate. A total of 10 μl of each sample was loaded with 5 μl glass microcapillary tubes (Sigma Aldrich) 1 cm from the bottom of the TLC plate. The plate was placed in a solvent system designed for separation of phospholipids by head group polarity: 65:25:4 chloroform: methanol: water. The silica plate remained in the solvent until the solvent front was ~1 cm from the top of the plate. The plate was left to dry thoroughly for 30 min and stained with phosphomolybdic acid. The

stained plated was heated with a hair-dryer to activate phosphomolybdic acid until lipid species became visible.

For identification of lipid species, two plates were run side by side. One was stained to identify the location of the lipids and one remained unstained for mass spectrometry. The unstained TLC was sent to the Centre for Chemical and Materials Analysis, School of Chemistry, University of Birmingham for identification of the lipids species by mass spectrometry. The masses were analysed directly from the TLC plates using the combined Plate Express™ and express**ion®** compact mass spectrometer system from Advion. PE, PG and CL standards were also analysed alongside to aid with identification of the masses.

2.9.3. Extraction of membrane lipids

For analysis of lipid composition of the inner and outer membranes, lipids were extracted from samples obtained from sucrose gradient protocol 2 (see 2.7.4). For 1.5 ml of isolated membrane sample, 5.7 ml of 1:2 chlorform: methanol was added and vortexed. An additional 1.875 ml of chloroform was added and vortexed again followed by a further 1.875 ml of water and a further vortex. The mixture was centrifuged at 1000 RPM in an IEC table-top centrifuge for 5 min, forming a two-phase system. The bottom phase was collected using a glass Pasteur pipette and transferred to a new tube.

To clean the bottom layer "authentic upper phase" was prepared. This involved repeating the same steps for lipid extraction but with water in place of sample. After the centrifugation the upper phase was collected. This authentic upper phase was mixed back with the lipid extract at a 2.25: 1 ratio and vortexed and centrifuged as before. The bottom phase was collected and transferred into a clean glass tube. The liquid was evaporated off at 100 ^{o}C and the lipids were re-dissolved in 200 μl of chloroform ready for TLC.

2.9.4. Phospholipid quantification using ammonium ferrothiocyanate

An ammonium ferrothiocyanate assay was used to normalise loading of phospholipid samples onto 2D TLC plates. This assay was used as described by Stewart (1980).

2.9.5. Two-directional thin-layer chromatography

For visualisation of membrane lipid content, lipid extracts were analysed using 2D-TLC on 7 x 7 cm silica plates. At total of 10 µl of each sample was loaded with 5 µl glass microcapillary tubes in the bottom left corner of the plate, 1 cm the edges. For direction one, the plate was placed in the same solvent system used previously: 65:25:4 chloroform: methanol: water, until the solvent front was 1 cm from the top of the plate. The plate was allowed to dry thoroughly before being rotated 90° anticlockwise and placed in the second solvent system: 80:12:15:4 chloroform: methanol: acetic acid: water. This solvent system is mainly designed to separate phosphotidylglycerol and cardiolipin, which co-migrate in the first system. After thorough drying the plate was sprayed with phosphomolybdic acid and heated until lipids were clearly visualised.

2.10. *Salmonella* techniques

2.10.1. Gene doctoring

Gene doctoring was used to construct gene deletions in *S.* Typhimurium. The principle for creating the deletion is similar, except the recombination fragment is transformed on a vector rather than a linear fragment (Lee et al., 2009). In the first step, a linear fragment containing a kanamycin resistance (*aph*) cassette with flanking homology to the regions up and downstream of the gene of interest was generated by PCR with the plasmid pDoc-K as a template. The plasmid pDoc-K carries a kanamycin resistance cassette that is flanked by FLP sites. The primers were designed so the fragment could be ligated into pDoc-C using the HindIII and BamHI restriction sites. After the HindIII restriction site, the 39 bp at the 5′ end of the forward primer were homologous to the region upstream of the gene of interest. The next 20 bp were homologous to 5′ end of the *aph* cassette on pDoc-K. After the BamHI restriction site, the 50 bp at 5′ end of the reverse primer were homologous to the last seven codons of the gene of interest and the region downstream. The next 20 bp of the reverse primer were homologous to the end of *aph* cassette on pDoc-K. The generated fragment was ligated into pDoc-C using HindIII and BamHI restriction sites, and

transformed and checked as described previously. Features of the pDoc-C plasmid include a multiple cloning site that is flanked by two I-SceI sites, the *sacB* gene (for sucrose sensitive) and an ampicillin resistance gene. The *sacB* gene is important for the determination of genuine recombinants later in the protocol.

For construction of the gene deletion, *Salmonella* Typhimurium SL1344 was electroporated with pDoc-C containing the PCR fragment and pACBSCE and plated onto LA supplemented kanamycin, ampicillin and chloramphenicol (LA_{CAK}). The plasmid pACBSCE carries the λ-Red recombination machinery and I-SceI. The transformants were plated onto LA_{CAK} with and without 5% (w/v) sucrose to ensure the *sacB* gene on pDoc-C was still intact. A sucrose sensitive colony was inoculated into 1 ml LB_{CAK} with 0.5% glucose (to prevent expression of *I-SceI* and λ-Red genes on pACBSCE) and grown for 2 h at 37°C. The cells were pelleted by centrifugation at 6000 *g* for 5 min and re-suspended in LB with 0.5% arabinose to induce expression of the *I-SceI* and λ-Red genes on pACBSCE. The I-SceI endonuclease cleaves the I-SceI sites either side of the cloning site in pDoc-C, linearising the kanamycin resistance cassette for recombination on the chromosome. It also cleaves the I-SceI sites in pACBSCE to encourage plasmid loss. The culture was grown at 37°C for 4-5 h and pelleted again at 6000 *g* for 5 min. The cells were re-suspended in 100 μl of LB and plated on LA with kanamycin and 5% (w/v) sucrose. The *sacB* gene on p-Doc-C allows the distinction of kanamycin resistant colonies that have had successful chromosomal recombination (i.e. are sucrose resistant and lack pDoc-C) from those that have not had successful recombination (i.e. are sucrose sensitive and kanamycin resistant due to the presence of pDoc-C). Kanamycin and sucrose resistant colonies were screened for loss of pDoc-C and pACBSCE by patch plating onto LA, LA with ampicillin and LA with chloramphenicol. Ampicillin and chloramphenicol sensitive colonies were checked for gene insertion mutations by colony PCR using check primers.

2.10.2. P22 transduction

In this study P22 phage was used to transfer kanamycin cassette insertion mutations to a clean background of *S.* Typhimurium. For generation of the donor phage stock, the donor strain was grown overnight in 5 ml LB with kanamycin. The next day the overnight was diluted 1:100 in 5 ml of LB containing 10 mM magnesium sulphate and 5 mM calcium chloride and was incubated at 37°C for 30 min. Five microlitres of P22 phage stock (previously generated from the parent strain) was added to the culture, which was then incubated overnight at 37°C. Following overnight incubation, 1ml of chloroform was added to the culture and vortexed for 15 s, to kill the donor bacteria. This was centrifuged at 6000 *g* for 10 min at 4°C. The clear supernatant was transferred to a glass universal containing 200 μl chloroform for storage.

For transduction, the recipient strain was grown overnight in 5 ml LB. To pellet the cells, the culture was centrifuged at 6000 *g* for 5 min at room temperature. The pellet was re-suspended in 1 ml of LB containing 10 mM magnesium sulphate and 5 mM calcium chloride. Separate transductions were prepared with 200 μl of recipient bacteria and 5, 10, 50 or 100 μl of the donor phage stock. The mixtures were incubated for 15 min at 37°C before the addition of 100 μl sodium citrate and 1 ml LB. After a further 45 min of incubation the cells were pelleted at 6000 *g* for 5 min at room temperature. The pellet was washed 3 times in PBS before a final resuspension in 100 μl LB. The cells were and plated onto LA with kanamycin and incubated at 37°C for 48 h. Any colonies were checked for gene insertion mutations by colony PCR using check primers.

2.10.3. Preparation of Lipopolysaccharide

In order to establish an infection in mice, *Salmonella* Typhimurium requires intact lipopolysaccharide (LPS). To check for the presence of LPS, the strain of interest was grown overnight in 5 ml LB. After overnight incubation, 250 μl of culture with an OD_{600} of 4 (or the equivalent volume for the correct number of cells) was centrifuged at 14000 *g* for 10 min. The pellet was re-suspended in 2x laemmli sample buffer (Sigma) and incubated for 4 min

at 100°C. The suspension was incubated at -80°C for 5 min, then defrosted and incubated again at 100°C for a further 4 min. The samples were centrifuged at 14000 *g* for 2 min and 80ul of the supernatant was removed and combined with 5 μl of 5 mg/ml proteinase K and incubated at 60 °C for 1 hour. To deactivate proteinase K, the samples were incubated at 95 °C for 5 min. The LPS extract was separated by SDS-PAGE and visualised using the Silver Quest Staining Kit (Invitrogen).

2.10.4. Serum bactericidal assay

Serum bactericidal assay (SBA) was used to check if different *S.* Typhimurium strains were sensitive to serum from healthy humans. After overnight incubation, 1.5 ml of culture with an OD_{600} of 0.6 (or equivalent) was centrifuged at 6000 *g* for 5 min at room temperature. The pellet was washed twice and re-suspended in 1 ml PBS to make the strain stock inoculum. The stock inoculum was diluted 1:10 by transfer of 100 μl into 900 μl of PBS. In four fresh tubes, 10 μl of the 1: 10 dilution was added to of 90 μl serum and incubated for 45, 90 and 180 min at 37°C. At each time point a tube was collected and the contents were diluted down to 10^{-3} by mixing 10 μl of the tube contents with 90 μl of PBS. Ten microliters of each dilution was pipetted as drops in triplicate on LA plates. The original stock was also diluted down 10^{-8} and plated. After all the time points were collected, the plates were incubated overnight at 37°C. The number of colonies was counted and the values were adjusted to CFUs/ml for 10^{o}. The data were plotted as a log decrease in colonies relative to the strain stock inoculum.

2.10.5. Adhesion and invasion of HeLa cells

HeLa cells were taken from liquid nitrogen stocks and placed into a 37°C water bath to thaw. The cells were removed and placed in a 75 ml tissue culture flask with 10 ml of complete DMEM media (DMEM supplemented with 2 mM L-glutamine, 10% (v/v) fetal bovine serum (FBS) and 1% (w/v) penicillin-streptomycin). At all stages, tissue culture media were pre-warmed to 37°C and mammalian cells were incubated in 5% CO_2 The next

day the medium was removed and the cells were washed once with 10ml PBS and replace with 10 ml fresh complete DMEM medium. The cells were left for approximately 3 days until the cells were 80% confluent.

To prepare for infection assays, medium was removed from the confluent cells and the cells were washed once with 10 ml PBS. To detach the cells 1 ml of trypsin was added and incubated for 5 mins at 37°C. After incubation, 10 ml of complete DMEM was added to the cells and pipetted up and down to remove clumps of cells. A small volume of cells was counted using a haemocytometer and the number of cells/ml was calculated. The cells were adjusted to a concentration of 1 x 10^5 cells/ml and 1 ml of cells were added to each well of a 24 well plate. The cells were incubated overnight before infection studies.

The following day, bacterial cells were adjusted from an overnight culture to a concentration of 5 x 10^6 cells/ml in plain DMEM medium. The number of bacterial cells was calculated based on the estimation that 1 ml of culture with an OD_{600} of 1 contains 1 x 10^9 bacterial cells. For infection, the complete DMEM was removed from the HeLa cells and each well was washed twice with 1 ml of PBS. For each bacterial strain, 1 ml of bacterial cells (in plain DMEM) was added in triplicate to the wells containing macrophages. To initiate infection, the plates were centrifuged at 600 *g* for 5 min at 37°C. The plates were incubated for 30 min at 37°C in 5% CO_2 to allow infection. After 30 min the media was changed to complete DMEM with 100 μg/ml gentamycin to kill external bacteria. This was considered time 0 for invasion experiments. To prevent damage of the HeLa cells from long-term exposure to a high concentration of gentamycin, the media was changed again to plain DMEM with 10μg/ml gentamicin after 1 hour of invasion. To end the experiment after 0, 2 or 5 h of invasion, the cells were washed twice with 1 ml of PBS. To lyse the HeLa cells, 1 ml of PBS and 0.5% Triton-X 100 was added to each well and pipetted up and down. The lysed cells were diluted down 10^{-3} in a 96 well plated. LA plates were split into quarters and 10 μl of each dilution (10^0 to 10^{-3}) was plated and incubated overnight at 37°C. The following day the plates were counted and the CFUs/ml of the 10^0 stocks were calculated.

2.10.6. Preparation of L-cell media for differentiation of bone marrow cells to bone marrow derived macrophages

L-cell media is used to differentiate murine bone marrow cells to bone marrow derived macrophages.. To prepare L-cell media, a 1 ml aliquot of L-cells was thawed in a water bath at 37°C. Once thawed, the cells were re-suspended in 1 ml of complete RPMI media (RPMI and 10% (v/v) FBS and 1% (w/v) penicillin-streptomycin) and transferred to a 15 ml centrifuge tube with 9 ml of complete RPMI media (cRPMI). The cells were centrifuged at 1000 *g* for 10 min at 4°C. The media was gently removed and the cells were re-suspended in 10 ml of cRPMI and transferred to a 25 ml tissue culture flask. A further 15 ml of cRPMI was added to the culture flask to make a final volume of 25 ml and the flask was incubated overnight at 37°C. The following day, the media was removed and dead cells were washed off with 10 ml of cRPMI. 25 ml of fresh cRPMI was added and the cells were incubated at 37°C until confluent (approx. 2 days).

Once confluent, the medium was removed and the cells were washed with 10 ml of PBS. The cells were detached from the flask by the addition of 2 ml of trypsin and incubated for 5 min at 37°C with gentle agitation. The cells were mixed in 10 ml of cRPMI by pipetting and transferred to a 50 ml centrifuge tube. A small volume of cells was counted using a haemocytometer, and the number of cells/ml was calculated. A total of 2.5 x 10^5 cells were seeded into a 175 ml tissue culture flask with 50 ml of cRPMI. The cells were incubated until confluent (approx. 5 days). Once confluent, the media was removed and filtered through a 0.22 μM filter into a 50 ml tube. The media was stored at -80°C until required.

2.10.7. Preparation of bone marrow derived macrophages

To obtain bone marrow derived macrophages, the bone marrow was isolated from mice aged 6-8 weeks. The bones of the hind legs were removed and placed into a petri dish with 5 ml of cRPMI. All excess muscle and skin was removed and each leg was cut in two at the joint using a scalpel. To remove bone marrow, the ends of each bone were cut and the bone marrow was rinsed through using cRPMI and a needle and syringe. The bones were rinsed

until they were translucent to remove all of the bone marrow. The isolated bone marrow was transferred to a 50 ml centrifuge tube and centrifuged at 600 *g* for 10 min at room temperature. The pellet was re-suspended in 1 ml of red blood cell lysis buffer (Sigma) and incubated at room temperature for 5 min. 10 ml of cRPMI was added and the white blood cells were pelleted at 600 *g* for 10 min at room temperature. The cells were re-suspended in 10 ml of cRPMI and passed through a cell strainer into a fresh 50 ml tube. The cells were counted using a haemocytometer and 5 x 10^6 cells were seeded into a petri dish with 10 ml of cRPMI and 10% (v/v) L-cell medium. One petri dish provided enough cells for 5 invasion experiments. The cells were incubated overnight at 37°C in 5% CO_2. The following day an additional 10 ml of cRPMI and 10% (v/v) L-cell medium was added to each petri dish. After 3 days, 10 ml of media was removed and replaced with 10 ml of fresh cRPMI and 10% (v/v) L-cell medium. After a further 3 days the cells were ready the invasion experiments.

2.10.8. Invasion of bone marrow derived macrophages

For invasion experiments, 15 ml of medium was removed from each petri dish of recently prepared bone marrow derived macrophages. In the remaining 5 ml of medium, the cells were removed from the bottom of the petri dish using a cell scraper and transferred to a 50 ml centrifuge tube. The cells from all of the petri dishes were centrifuged at 600 *g* for 10 min at 37°C and re-suspended in 5 ml of cRPMI. The cells were counted and diluted to a concentration of 5 x 10^5/ml using cRPMI. For each invasion experiment 200 µl were placed in triplicate into a 96-well tissue culture plate and left overnight at 37°C in 5% CO_2.

The following day, bacterial cells were adjusted from an overnight culture to a concentration of 5 x 10^6 cells/ml in plain RPMI medium. The number of bacterial cells was calculated based on the estimation that 1 ml of culture with an OD_{600} of 1 contains 1 x 10^9 bacterial cells. For infection, the cRPMI was removed from the macrophages and each well was washed twice with 200 µl of PBS. For each bacterial strain, 200 µl of bacterial cells (in plain RPMI) was added in triplicate to the wells containing macrophages. To initiate infection, the 96 well plates were centrifuged at 600 *g* for 5 min at 37°C. The plates were

incubated for 30 min at 37°C in 5% CO_2 to allow infection. After 30 min the media was changed to plain RPMI with 100 μg/ml gentamycin to kill external bacteria. This was considered time 0 for invasion experiments. After 1 hour the media was changed to plain RPMI with 10μg/ml gentamicin. To end the experiment after 2, 5 or 24 h of invasion, the cells were washed twice with 200 μl of PBS. To lyse the mammalian cells, 40 μl of PBS and 1% (w/v) Triton-X 100 was added to each well and incubated at 37°C for 10 min. After incubation, an additional 160 μl of PBS was added to each well and the suspensions were diluted down to 10^{-3}. LA plates were split into quarters and 10 μl of each dilution (10^{0} to 10^{-3}) was plated and incubated overnight at 37°C. The following day the plates were counted and the CFUs/ml of the 10^{0} stocks were calculated.

2.10.9. Intraperitoneal infection of mice with Salmonella

To study the pathogenesis of different *Salmonella* strains, C57/Bl6 mice were infection via intraperitoneal (IP) infection. Overnight cultures of each strain were inoculated 1:20 into 10 ml of fresh LB and grown to and OD_{600} of 1. The cells were centrifuged at 6000 *g* for 5 min and the pellets were washed twice with PBS. After a final re-suspension the concentration of cells was adjusted to a concentration of 5000 cells/ml by a series of dilutions. For injection, each mouse was injected with 200 μl of bacteria (1000 bacterial cells). For pilot experiments 2 mice were infected per *Salmonella* strain, and this was scaled up to 4 mice for full experiments.

After 3 days of infection, the mice were culled and the livers and spleens were isolated. The organs were mashed through a cell strainer using a sterile syringe and rinsed through with 5 ml of PBS into a 50 ml tube. The mixture was diluted down to 10^{-3} and 10 μl of each dilution were plated and spread onto a quarter of an LA plate. Each organ was plated out in duplicate. After overnight incubation at 37°C, *Salmonella* colonies were counted and the number of CFUs/organ was calculated.

2.11. Bioinformatics

The bioinformatics analyses were carried out in collaboration with Nathaniel J. Davies, University of Leicester. Nathaniel's input included advice on software use, making custom Perl scripts where software was unavailable, and construction of heat maps.

2.11.1. MCE sequence retrieval and architecture definition

All MCE domain-containing sequences were retrieved by scanning the UniProtKB database (version 2015_03) for matches to the PFAM MCE HMM (accession PF02470.15) using HMMER (hmmsearch 3.1b2) (Finn et al., 2015; UniProt Consortium, 2015). Originally, the PFAM gathering bit threshold was used which resulted in some known MCE domain containing proteins having too few MCE domains. Therefore HMMER default significance values were used and the thresholds were judged appropriate through manual inspection of known MCE domain-containing *E. coli* proteins. All MCE domain containing protein sequences were compiled and put again into HMMER to scan for all PFAM domains (Pfam database 27.0), using the PFAM gathering bit score thresholds. As per the 'no overlap' rule in PFAM (Finn et al., 2016), in the case of overlapping predictions with members of the same clan only the most significant hit was retained. These resulting predicted domains were sorted by start position and were used to define the architectures.

2.11.2. Phylogenetic distribution

To provide a percentage measure of prevalence for each MCE architecture in each taxonomic rank, all of the species with complete proteomes were examined for the absence or presence of each MCE architecture. To avoid bias, one proteome sequence per species was selected randomly. These results were then summed for each species/genera/class/phylum to calculate the percentages. Heat maps were made using a custom Perl script.

2.11.3. Clustering

As the full set of MCE-domain containing proteins was too large for cluster analysis, proteins were reduced to a representative set using the CD-HIT suite at a cut-off of 50% protein identity (Huang, 2010). This resulted in a computationally tractable set of 1,734 proteins. These proteins were used in all-against-all searches in BLASTP with various e-value thresholds. Information about each protein (phylum, architecture, etc) was incorporated into the BLAST results based on the previous architecture designations and the information from Uniprot. Cluster diagrams were constructed using Cytoscape (v.3.3.0), where each node represents a single protein sequence and each line (or edge) represents a match below the e value threshold. The nodes were coloured by either architecture, phylum or class.

2.11.4. Gene neighbourhoods

For each MCE-domain containing protein in Bacteria, the Ensembl gene IDs provided by Uniprot were cross-referenced in conjunction with the NCBI taxonomic identifier to retrieve the sequence database for each organism in Ensembl (using the Ensembl Perl API)(Cunningham et al., 2015; Sayers et al., 2009). A section of the genome 10 kbp either side of each MCE gene was extracted and gene information was retrieved for all genes in this region. Domain architectures were predicted for the proteins encoded by these genes by scanning for PFAM domains (PFAM database 27.0) using HMMER (hmmsearch 3.1b2). Where several genomes for a particular species were found, a single representative genome was chosen randomly so that each species was only represented once in the neighbourhood maps. Sets of adjacent MCE domain-containing genes were grouped together, sorted in ascending order by start position, oriented by the strand of the first gene in the group, and analysed as a single neighbourhood based on the 5'-most gene.

In order for groups of conserved neighbourhoods to be displayed together, gene neighbourhoods were first clustered using a nearest neighbour joining approach. The similarity measure used in the clustering method gives greater weight to genes closer to the

centre of the neighbourhood, based on the assumption that gene positions further away from the centre of the neighbourhood are less likely to be conserved. The analysis was done with a maximum of 10 genes up and downstream. This clustering method was used to constructs neighbourhood diagrams for each protein architecture and/or phylum, where similar neighbourhoods were clustered together. The most common neighbourhoods were manually selected from these diagrams. To calculate a percentage for a particular gene neighbourhood architecture, the number of genes with that architecture was divided by the total number of genes in that category (phylum and/or protein architecture).

2.11.5. Secondary structure prediction

PSIPRED was used to predict secondary structures for a chosen set of proteins (Jones, 1999).

CHAPTER 3

BIOINFORMATIC ANALYSIS OF THE DISTRIBUTION, GENOMIC NEIGHBOURHOODS AND EVOLUTION OF MCE DOMAINS

3.1. Introduction

An MCE domain is defined by an 81 amino acid PFAM hidden Markov model (see figure 1.12). This model can be used to identify MCE domains in a given protein and with the aid of other bioinformatic resources all MCE domain-containing proteins can be analysed. A limited bioinformatics analysis of MCE domains already exists, focusing on Actinobacteria (Casali and Riley, 2007). This study from 2007 looked closely at the distribution and operon types across *Actinomycetales*, with a brief insight into the genomic neighourhoods of genes that encode MCE domain containing proteins in Gram-negative bacteria. This study did not look at their evolution outside Actinobacteria nor did it distinguish between the different architectures of the proteins. In the 9 years since the publication of this paper, the decreased cost of sequencing has resulted in a far greater phylogenetic depth and breadth of the genomic data upon which such studies can be conducted.

The aim of this chapter was therefore to conduct a bioinformatic study of MCE domain containing proteins, utilising new genomic data and expanding the focus to the bacterial kingdom as a whole. Such a study could answer several unanswered questions. First, what are the main protein architectures (types of MCE domain containing protein) and in which bacteria are they found? Secondly, are these different types of protein associated with different sets of neighbouring genes, and does this give an insight into the different functions of these proteins? Thirdly, how conserved are MCE domains across bacteria and how have they evolved?

3.2. Results

3.2.1. The top MCE domain containing protein architectures

To identify the top protein architectures, the PFAM HMM was scanned against all UniProtKB proteins using HMMscan. A total of 17,282 MCE domain containing proteins were identified. These proteins were scanned further for other PFAM domains in order to construct protein architectures. A total of 38 architectures were detected, with the top four architectures found in over 97% of the proteins (Figure 3.1). For a full list of protein

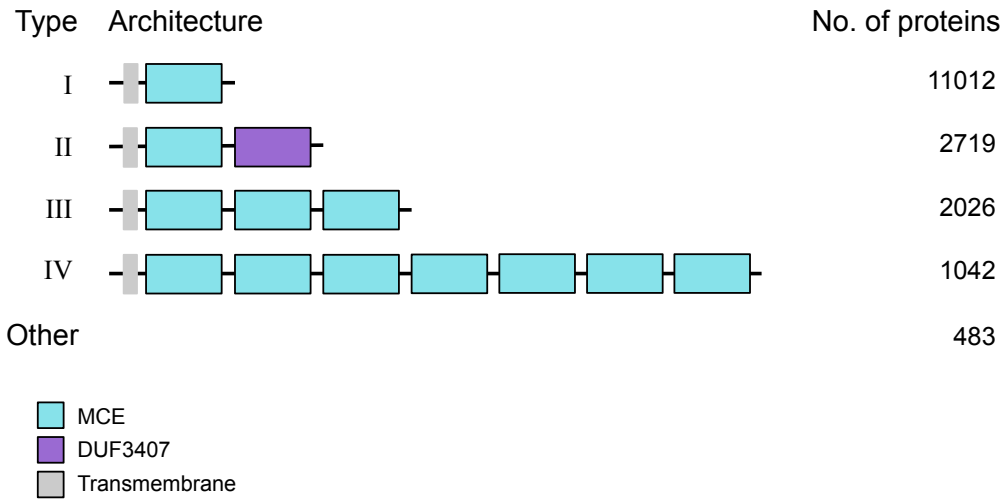


Figure 3.1. The top four MCE domain containing protein architectures, ordered by the number of proteins they are found in. "Other" is the sum of 32 other protein architectures and the full list can be found in appendix i.

architectures see appendix i. There was an N-terminal TM region in nearly all of the MCE domain containing proteins analysed. The most common architecture, found in over 11,000 proteins, was a single MCE domain (type I). Type II proteins contained a single MCE domain followed by a domain of unknown function (DUF) 3407 and were the second most common protein architecture. DUF3407, also known as Mce4_CUP1 (Cholesterol Uptake Porter) domain, is Actinobacteria-specific and is associated with a single MCE domain in 99% of its occurrences (Finn et al., 2016). The 3rd and 4th most common architectures, type III and IV proteins, contained multiple MCE domains with three and seven MCE domains in tandem, respectively. These top protein architectures were used as a foundation for the rest of the bioinformatic study.

3.2.2. Phylogenetic distribution of MCE domain containing proteins

All species with fully sequenced genomes were analysed to determine the distribution and prevalence of the top protein architectures. To remove bias, the analysis was limited to one proteome per species. The results were presented as heat maps that reveal the distribution of the top four MCE architectures in the Metazoa, Viridiplantae and Bacteria kingdoms (as defined by NCBI) (Figures 3.2 and 3.3).

For bacteria, proteins containing MCE domains were found in 24 of the 31 phyla (Figure 3.2). Type I proteins were by far the most common, and were usually more abundant than any other architecture within a phylum. The only exception was in Actinobacteria, where type I and II proteins were equally as prevalent. Type III and IV proteins were specific to Proteobacteria, along with all proteins with greater than 2 MCE domains (see appendix ii).

A number of MCE domains were identified in eukaryotic proteomes (Figure 3.3). Type I proteins were abundant in Viridiplantae and a small number of MCE domain-containing protein of other architectures were identified in animal proteomes. While the Viridiplantae type I proteins appeared to be genuine, all but one of the animal proteins appeared to be due to bacterial contamination. The one exception was a protein found in

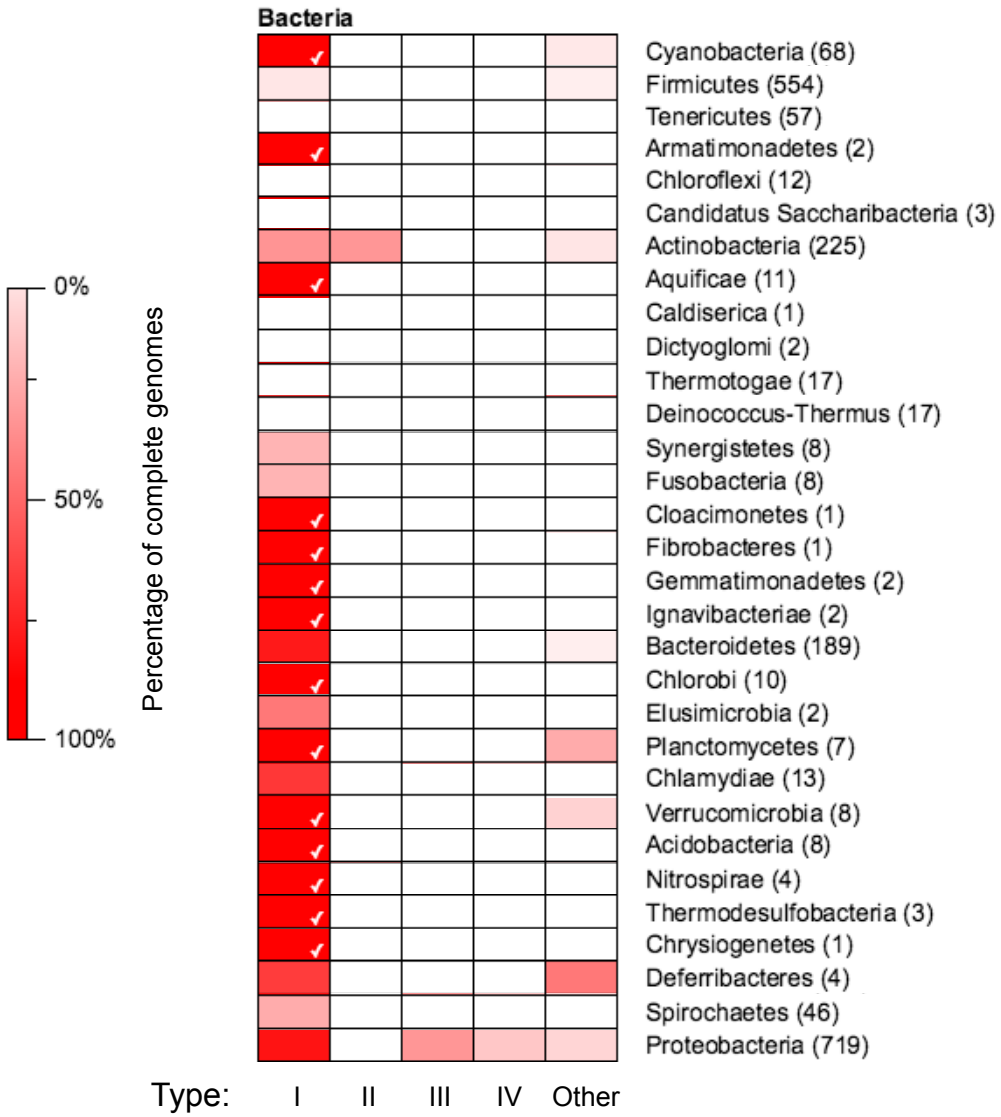


Figure 3.2. A heat map showing the distribution of MCE domain containing proteins across bacterial phyla. The heat map is separated by architecture type and phylum, with the number of species in brackets. A white tick indicates that a given architecture is found in 100% of species in the phylum that have a full proteome sequence.

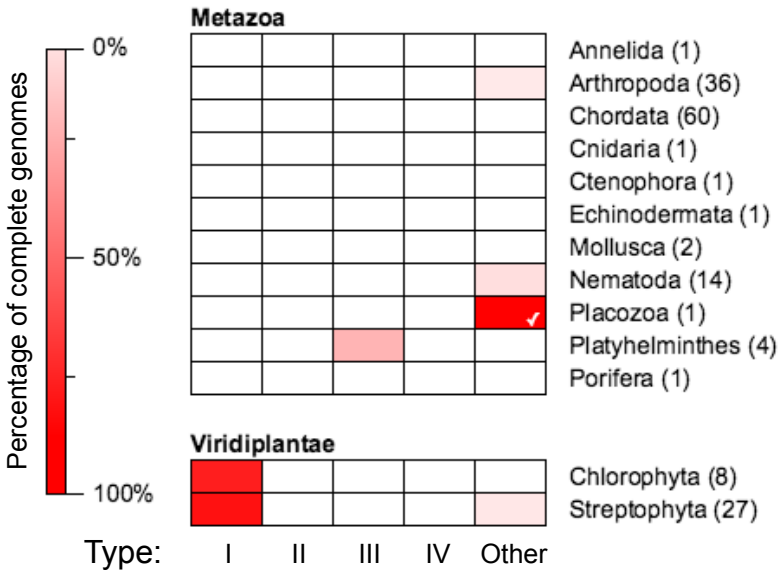


Figure 3.3. A heat map showing the distribution of MCE domain containing proteins across Metazoa and Viridiplantae phyla. The heat map is separated by architecture type and phylum, with the number of species in brackets.

Trichoplax adhaerens, a simple animal with an unusually large mitochondrial genome found in the Placozoa phylum (Dellaporta et al., 2006; Signorovitch et al., 2007). This protein is similar to a mitochondrial NADH dehydrogenase subunit (NDUFA12) in other animal species and is encoded by a gene that contains introns, making a bacterial contamination origin unlikely. In this protein, the MCE domain is found between the domains NDUFA12 and DUF155, a conserved bacterial domain of unknown function.

Overall, this analysis revealed the widespread distribution of the different protein architectures in bacterial phyla as well as a genuine unstudied MCE domain-containing protein in an animal species.

3.2.3. Distribution of MCE architectures in Proteobacteria, Gamma-proteobacteria and Enterobacteriaceae

As the aim of this project was to understand the function of MCE domains in *E. coli*, the distribution of MCE architectures in Proteobacteria, Gammaproteobacteria and Enterobacteriaceae were studied more closely. Type I and III proteins were found in all five classes of Proteobacteria and type I proteins were the most abundant (Figure 3.4). Type IV proteins were only found in Delta, Epsilon and Gammaproteobacteria and were completely absent from Alpha and Betaproteobacteria.

Closer analysis of Gammaproteobacteria revealed that MCE domain containing proteins were found in all of the families (Figure 3.5). Type III proteins were much more widespread than type IV, the latter of which were only found in 7 of the 31 families, including Enterobacteriaceae. In Enterobacteriaceae, type III and IV proteins were almost as prevalent type I proteins. In over 50% of the Enterobacteriaceae genera, including *Escherichia* and *Salmonella,* type I, III and IV proteins were found in all species analysed, indicating that multi-domain proteins are particularly important in this family (Figure 3.6).

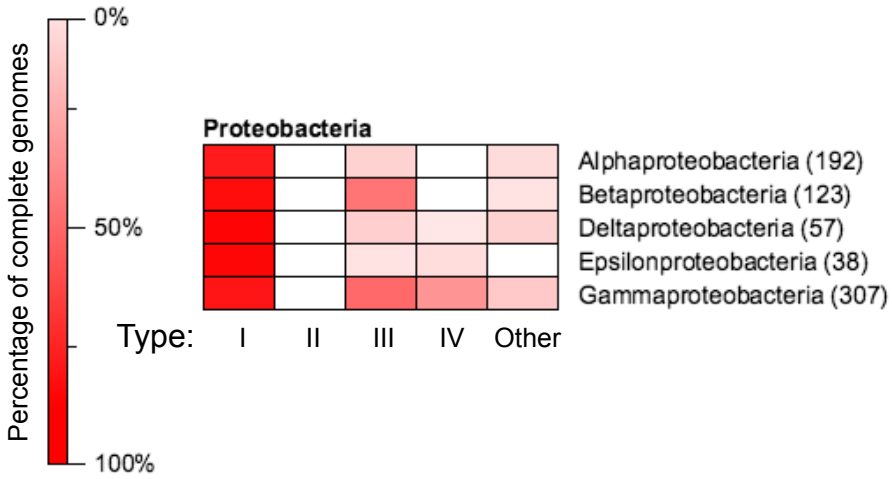


Figure 3.4. A heat map showing the distribution of MCE domain containing proteins across Proteobacteria. The heat map is separated by architecture type and class, with the number of species in brackets.

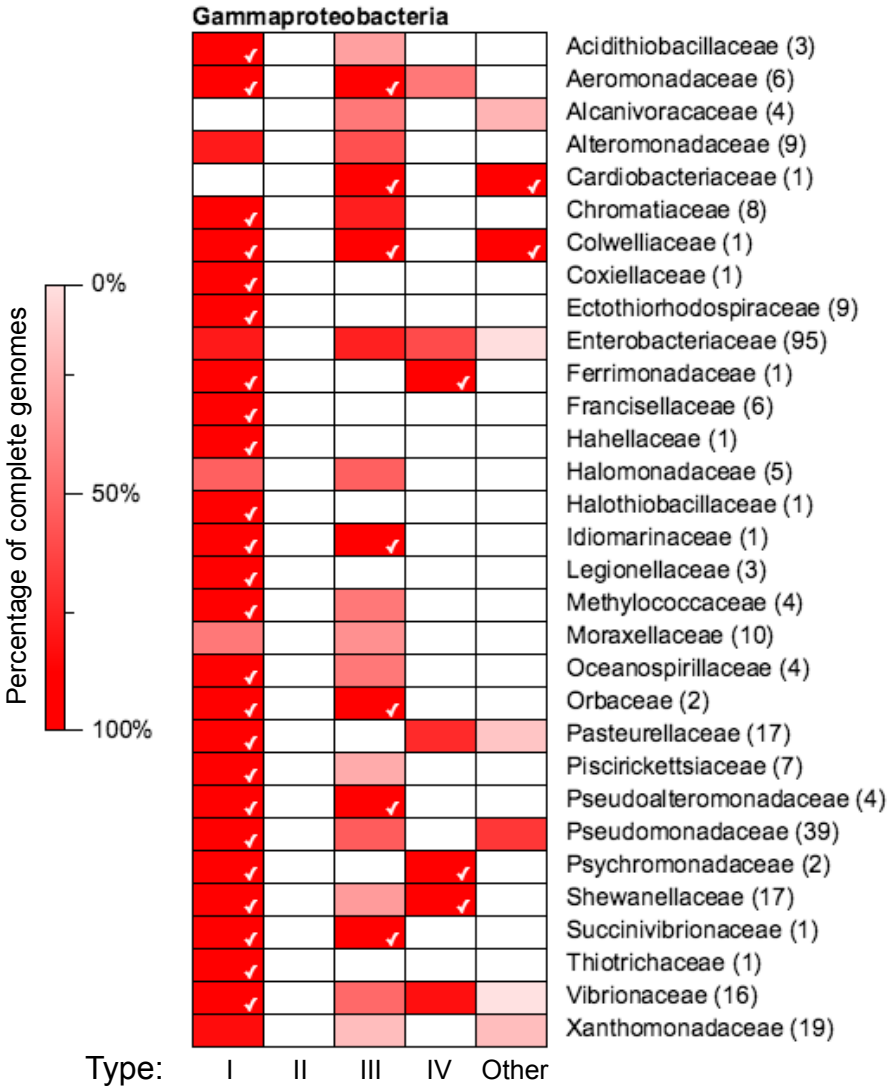


Figure 3.5. A heat map showing the distribution of MCE domain containing proteins across Gammaproteobacteria. The heat map is separated by architecture type and family, with the number of species in brackets.

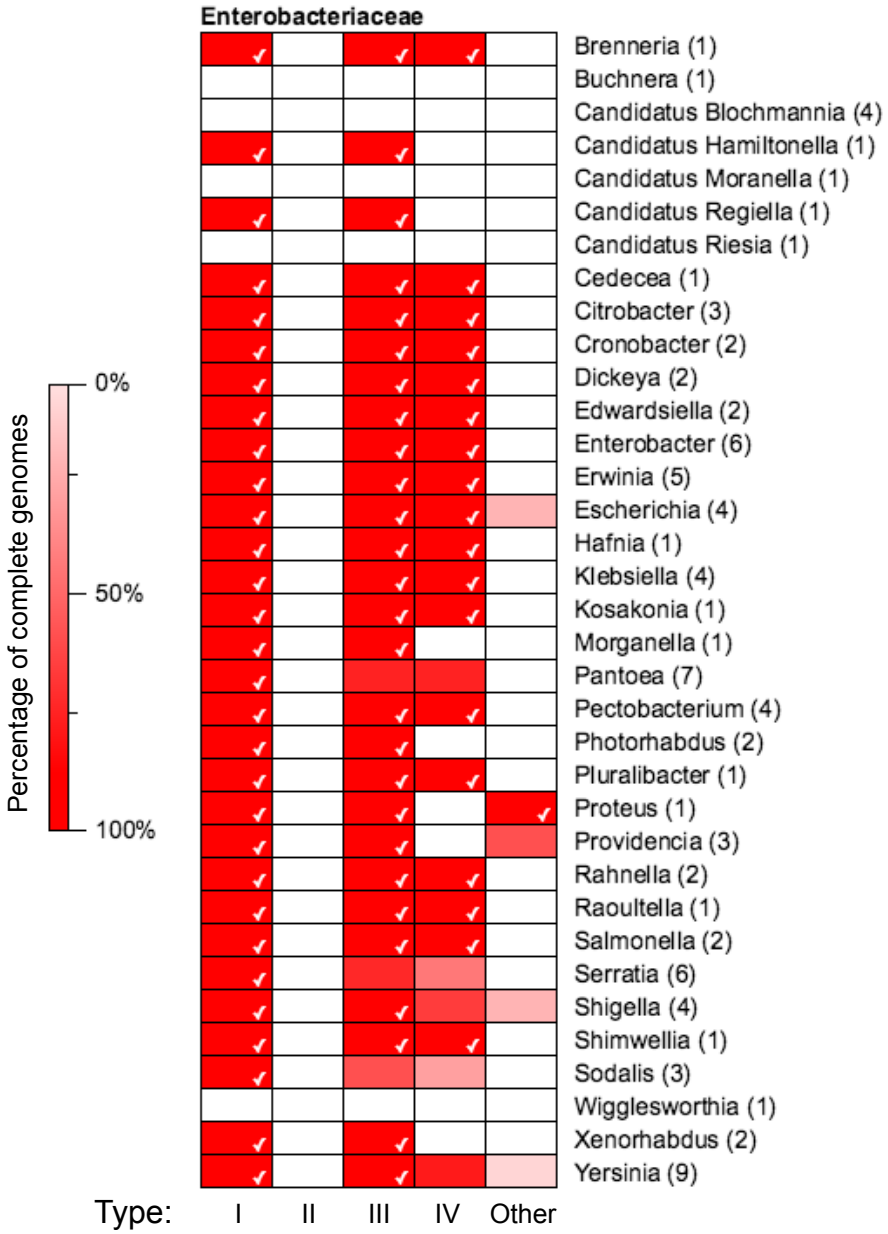


Figure 3.6. A heat map showing the distribution of MCE domain containing proteins across Enterobacteriaceae. The heat map is separated by architecture type and genus, with the number of species in brackets.

3.2.4. Clustering of MCE domain containing proteins

To help understand the relationships between MCE domain containing proteins, protein-protein similarity clusters were constructed. The cluster database CD-HIT was used to select a computationally tractable set of 1,734 proteins that were used in all-against-all BLASTP searches, with an e-value cut-off of 1e-15. The results were visualised using Cytoscape and the networks were coloured by protein architecture or phylum (Figure 3.7). The cluster diagrams revealed that MCE domain containing proteins generally clustered within phyla. Nearly all Actinobacteria proteins formed a single cluster, containing type I and II proteins. Type I proteins from most other phyla clustered loosely together, including proteins from Cyanobacteria, Bacteroidetes, and Proteobacteria. All of the plant proteins clustered closely with Cyanobacteria, and the single *Trichoplax* protein clustered with Proteobacteria. All multi-domain proteins clustered together within a large cluster dominated by Proteobacteria. The positions of type III and IV proteins in this cluster suggest that they are a functionally divergent population that arose from a type I Proteobacterial protein in a single event.

As MCE domain containing proteins in Proteobacteria were found in two distinct clusters, the diagram was re-coloured to determine whether the distinction was class-dependent (Figure 3.8). All five major classes were found in both clusters, suggesting that the divergence of these protein types into distinct clusters occurred before the divergence of the classes. Overall, this analysis provides an insight into the evolution of the different protein types in Proteobacteria.

3.2.5. The genomic neighbourhoods of genes that encode MCE domains

In this study, a genomic neighbourhood is defined as the region surrounding the gene of interest on the genome. The neighbourhoods of genes that encode type I – IV proteins were analysed to identify proteins that might be functionally linked to MCE domains. The 10 kb regions up and downstream of genes that encode MCE domains were translated and scanned for protein domains. The neighbourhoods of type I and II protein-encoding genes

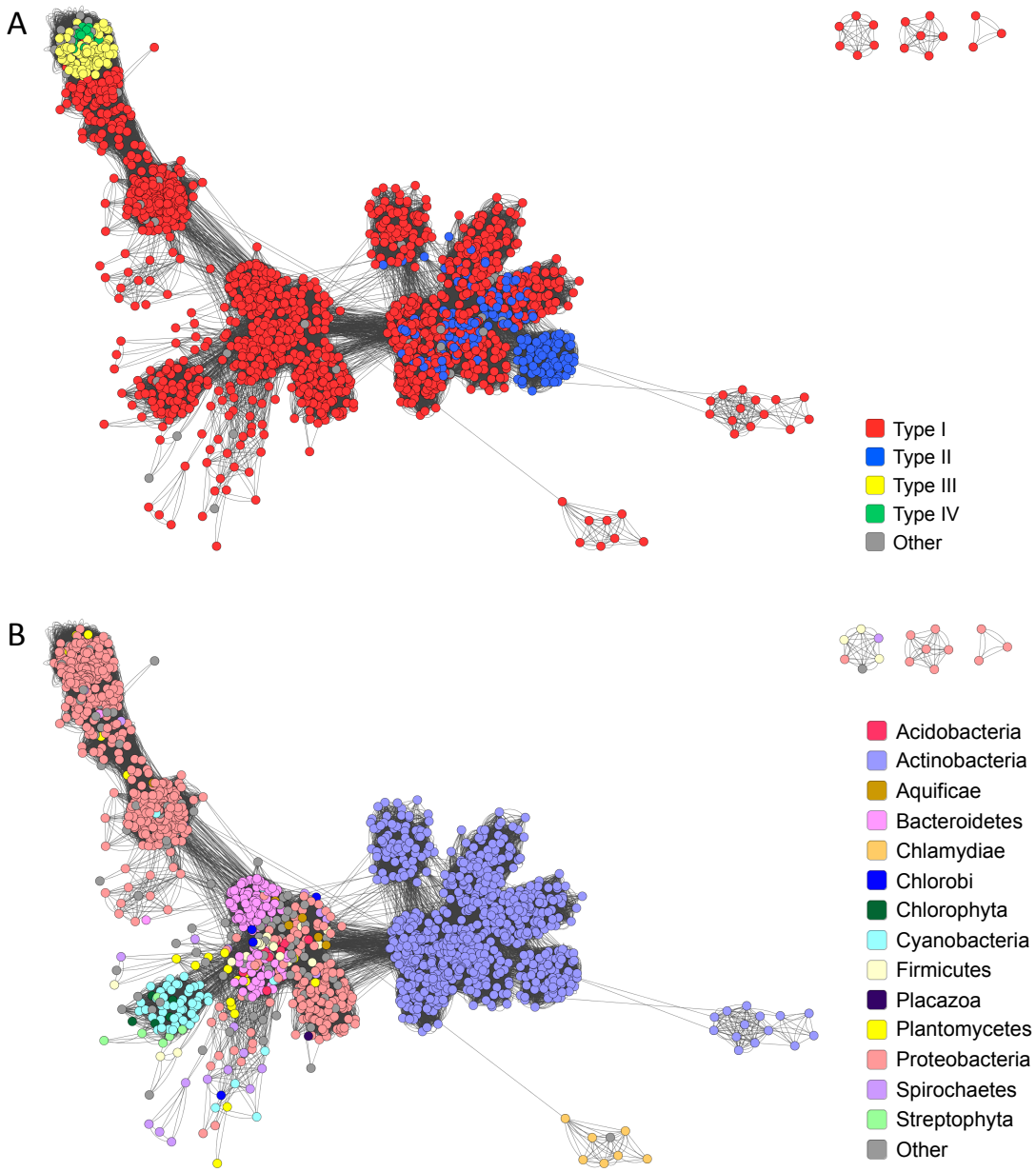


Figure 3.7. A cluster diagram of MCE domain containing proteins. Each node represents a set of proteins that are coloured by **A.** protein type and **B.** phylum.

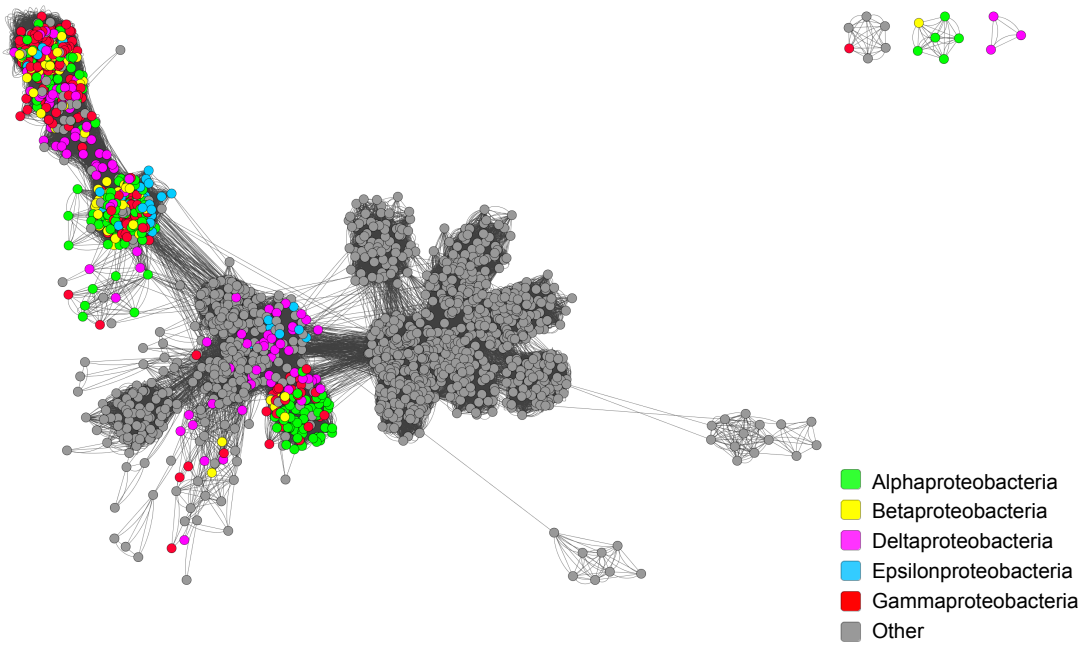


Figure 3.8. A cluster diagram of MCE domain containing proteins. Each node represents a set of proteins that are coloured by classes of Proteobacteria.

in Actinobacteria were remarkably similar, and therefore their analyses were combined. In Proteobacteria, the neighbourhoods for the genes that encoded type I, III and IV proteins were analysed separately. The neighbourhoods of all other phyla were analysed together. For each architecture and/or phylum the top MCE neighbourhoods were selected and their prevalence was calculated as a percentage (Figure 3.9).

Sixty-three percent of the *mce* operons in Actinobacteria consisted of 6 *mce* genes with 2 ABC transporter permease genes upstream, supporting a role as an ABC transporter (Figure 3.9i) (Casali and Riley, 2007). In around 20% of cases, this operon was associated with an ATP-binding gene, a component essential for a functional ABC transporter (Higgins et al., 1986). A similar ABC transporter operon structure was present in other bacterial phyla, including Cyanobacteria, Spirochaetes and Firmicutes, but with only one MCE repeat per operon (Figure 3.9v).

In Proteobacteria, type I genes were found with the same permease and ATP-binding domains upstream as Actinobacteria, 59% of the time (Figure 3.9ii). The most common domain encoded downstream in Proteobacteria was DUF330, found in 29% of the Proteobacteria type I neighbourhoods. In Beta- and Gammaproteobacteria, this DUF330 gene was replaced by genes that encode the periplasmic Tol domain, related to toluene resistance (Kim et al., 1998) and the cytoplasmic NTP binding STAS domain (Aravind and Koonin, 2000). This operon encodes the Mla pathway in *E. coli*, with the ATPase, permease, MCE, Tol and STAS domains found in MlaF, MlaE, MlaD, MlaC and MlaB, respectively. In Betaproteobacteria, such as *Neisseria meningitidis*, the gene that encodes outer membrane protein MlaA (VacJ) was found in the same operon (Caspi et al., 2016).

In Proteobacteria, genes that encode multi domain proteins were nearly always associated with upstream genes that encode two PqiA domains (Figure 3.9iii and iv). A PqiA domain has similarity to Proton_antipo_M, an antiporter domain involved in bidirectional membrane transport that requires energy from ATP hydrolysis but does not directly bind it (Finn et al., 2016). Type III neighbourhoods usually included a DUF330-encoding gene downstream.

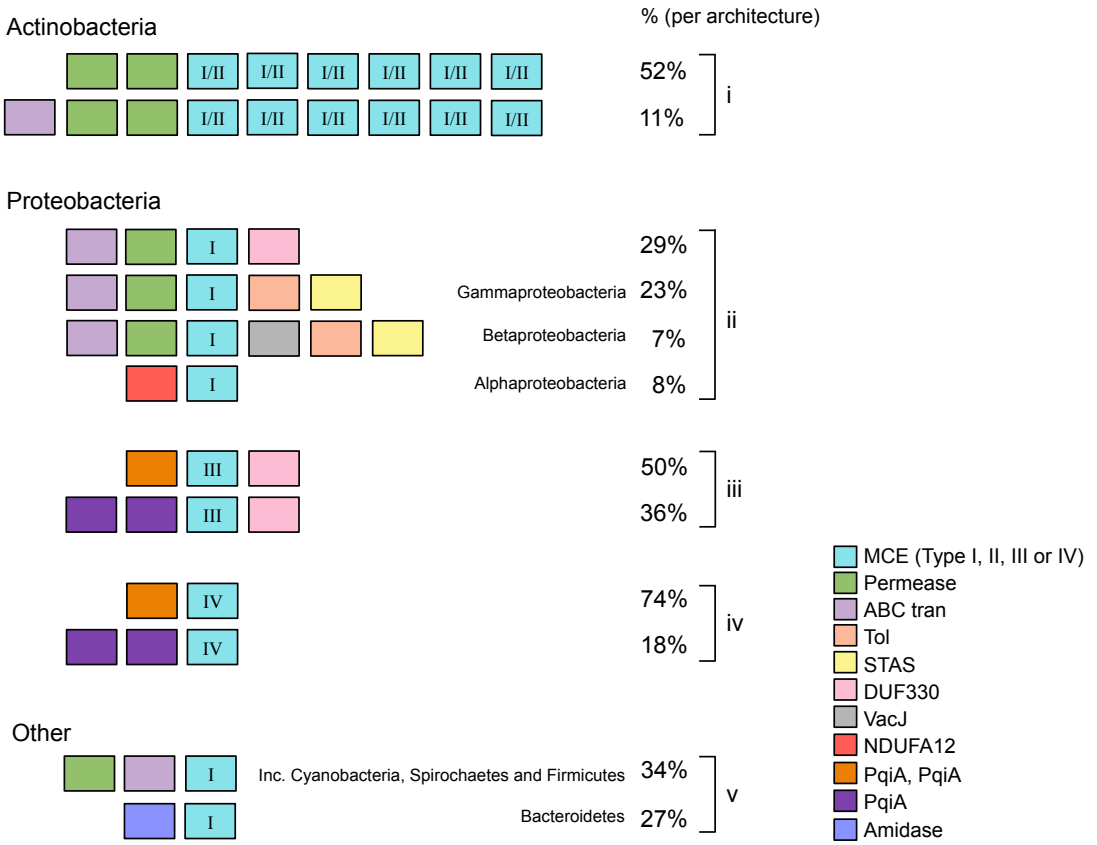


Figure 3.9. The most common MCE gene neighbourhoods separated by phylum and/or protein type. The percentages represent the prevalence of that given neighbourhood type within a category. All Actinobacteria neighbourhoods were analysed together. Proteobacteria were split into separate analyses based on protein architecture. All other phyla were combined.

However type IV encoding genes were not associated with a DUF330 gene, highlighting a distinction between these two multi-domain proteins.

In the neighbourhoods that encode type I proteins in Alphaproteobacteria, it was common to find upstream genes that encode proteins with an NADH dehydrogenase subunit domain (NDUFA12), which is involved in ATP generation in mitochondria (UniProt Consortium, 2007). It is not clear if these genes are functionally related in bacteria. The same two domains were found together in the *T. adhaerens* protein suggesting a gene fusion event might have occurred following the formation of mitochondria from Alphaproteobacteria (Andersson et al., 1998). In Bacteroidetes, a gene that encodes a protein with an amidase 3 domain was found upstream of a type I MCE gene. The amidase 3 domain is involved in peptidoglycan synthesis, but there is no evidence that this protein is functionally linked to the MCE domain containing protein or encoded in the same operon (Caspi et al., 2016). Overall, this analysis revealed that neighbourhoods of the genes are well conserved within phyla and protein architecture types. The neighbourhoods of genes that encode MCE domain containing proteins usually resemble transportation machinery.

3.2.6. Three operon architectures in E. coli

As the MCE operons in *E. coli* were the focus of this study, the predicted operon architectures in these organisms were determined. InterPro defines 3 proteins in *E. coli* named MlaD, PqiB and YebT, which are type I, III and IV proteins, respectively (Hunter et al., 2009). The *mlaD* gene is part of the *mlaFEDCB* operon, which has been reported here as the second most common gene neighbourhood for type I proteins in Proteobacteria. To obtain the same information for *pqiB* and *yebT*, their operons were predicted using ProOpDB and EcoCyc, using the *E. coli* K-12 MG1655 genome (Karp et al., 2002; Taboada et al., 2012). The most likely operon structures were as follows: *pqiAB-ymbA* and *yebST* (Figure 3.10). The domain architecture for each protein encoded in these operons was predicted using HMMER. The *pqiAB-ymbA* operon encodes proteins with two PqiA domains (in PqiA), three MCE domains (in PqiB) and one DUF330 domain (in YmbA).

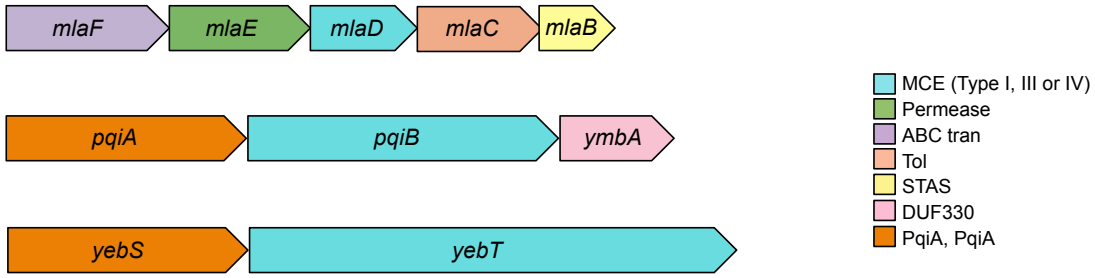


Figure 3.10. The three MCE operons in *E. coli*. Each gene is coloured by the domains that they encode.

The *yebST* operon in *E. coli* encoded YebS (which contains two PqiA domains), and YebT (which contains seven MCE domains). Therefore, the *pqiAB-ymbA* and *yebST* operons resembled the most common type III and type IV MCE neighbourhoods, respectively.

3.2.7. Conserved secondary structure of MCE domains

To determine whether the secondary structures of MCE domains are conserved across all life, proteins from *M. tuberculosis, E. coli, A. thaliana* and *T. adhaerens* were analysed using PSIPRED (Figure 3.11). Six beta strands were found in conserved locations in all of the MCE domains analysed. In multi-domain proteins there were some additional beta or alpha sheets in some of the MCE domains. Overall, these predictions showed a great conservation of the secondary structure of MCE domains across life.

3.3. Discussion

Many studies have shown that MCE domain containing proteins function in lipid trafficking (Cantrell et al., 2013; Klepp et al., 2012; Mohn et al., 2008; Pandey and Sassetti, 2008). The PFAM HMM defines highly conserved hydrophobic residues, which could be important for the binding of hydrophobic substrates such as lipids. The results presented in this chapter show that the most common type of MCE domain containing protein consists of a single MCE domain and no other predicted PFAM domains (type I). This type of protein is widespread in bacteria and includes MlaD, the MCE component of the retrograde phospholipid transport pathway identified in *E. coli*, and multiple predicted lipid transport proteins in Actinobacteria (e.g. Mce4B from predicted cholesterol uptake pathway in *M. tuberculosis*) (Malinverni and Silhavy, 2009; Pandey and Sassetti, 2008). In Actinobacteria, an MCE domain is commonly found in the same protein as DUF3407 (type II), an architecture that is specific to this phylum. This type of protein is often encoded by the same operons as type I proteins, indicating that the two protein types have similar functions in Actinobacteria. The next most abundant types of protein are the multi domain proteins, with three and seven MCE domains found in type III and IV proteins, respectively.

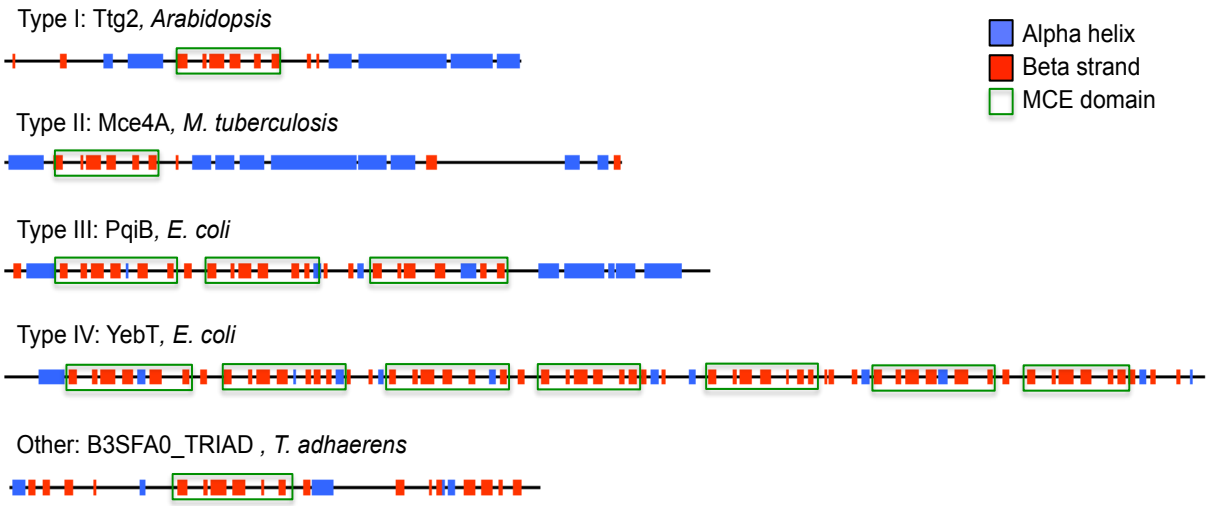


Figure 3.11. The predicted secondary structures of various MCE domain containing proteins, as predicted by Psipred.

These types of protein are specific to Proteobacteria, being particularly prominent in Enterobacteriaceae, indicating that they may be important for survival in the gut.

MCE domain containing proteins were also present in some Eukaryotes, where type I proteins were particularly prominent in Viridiplantae. In *Arabidopsis thaliana* this protein is found in the inner membrane of the chloroplast and is involved in the trafficking of phosphatidic acid from the outer membrane (Awai et al., 2006). A genuine MCE domain containing protein was found in the mitochondria of the animal *T. adhaerens*, which has a particularly large mitochondrial genome (Dellaporta et al., 2006; Signorovitch et al., 2007). It is feasible that this protein is only present due to a lack of evolutionary pressure on *T. adhaerens* to reduce its mitochondrial genome. Alternatively, the protein might be important for the existence of *T. adhaerens* in its marine habitat (Pearse and Voigt, 2007).

Type I – IV proteins make up over 97% of all identified MCE domain containing proteins, all of which contain an odd number of MCE domains. Given that multi-MCE-domain proteins with an even number of MCE domains are exceedingly rare, it may be that having an odd number of domains reflects an evolutionary constraint related to the function of these proteins. The clustering analysis indicated that MCE domains evolved with no major horizontal gene transfer events. Their existence in early branching lineages of bacteria (such as Cyanobacteria and Actinobacteria) suggests that these proteins were likely to have been present in early bacterial species (Hug et al., 2016). All examples of eukaryotic MCE domains are ultimately bacterial in origin, including those present in chloroplasts and the mitochondria of *T. adhaerens*. The chloroplast proteins clustered closely with Cyanobacteria and the *T. adhaerens* protein with Proteobacteria, as would be expected given the bacterial origins of chloroplasts and mitochondria, respectively (Andersson et al., 1998; Bonen and Doolittle, 1975). Overall, the single MCE domain type I architecture has remained remarkably conserved throughout bacterial evolution. The two major exceptions include the evolution of type II proteins in Actinobacteria, and multi-domain proteins in Proteobacteria. The close clustering of type III and IV proteins in Proteobacteria indicate

that multi-MCE domain containing proteins evolved from a single type I protein to fulfil a function specific to this lineage.

Previous bioinformatic research indicates that MCE domain containing proteins resemble the substrate-binding constituents of ABC transporters (Casali and Riley, 2007). The neighbourhood analysis lends support to this view, with the most common neighbourhood type across all bacteria resembling an ABC transporter. A typical ABC transporter core consists of two ATPase domains and two permease domains (Higgins et al., 1986). In Actinobacteria, it is common to find six MCE genes and two permease genes in an operon, but only occasionally does the operon encode an ATPase. It has however been shown in *Mycobacterim* species that MCE-related ATPases are often found elsewhere in the genome (Casali and Riley, 2007) rather than being strictly tied to the ABC operon. Two MCE-associated (mas) genes are found downstream in most of the studied MCE operons in Actinobacteria, but these were not identified in this analysis as they encode proteins without a conserved domain. This operon structure is specific to Actinobacteria and could be linked to the types of lipids these organisms require and/or the complex cell wall structures of some species. In *Mycobacterium tuberculosis* there are four MCE operons that have been reported to traffick lipids including palmitic acid and cholesterol (Forrellad et al., 2014; Mohn et al., 2008; Pandey and Sassetti, 2008).

The type I neighbourhoods of Cyanobacteria, Firmicutes and Spirochaetes typically encoded a single ATPase, permease and MCE domain. This type of operon also encodes the phosphatidic acid trafficking pathway in *Arabidopsis thaliana* (Awai et al., 2006). When considering the origin of chloroplasts, it is likely that these transporters have similar functions in both plants and Cyanobacteria. In Proteobacteria, type I neighbourhoods encode the same 3 protein domains but with conserved genes downstream of the MCE-encoding gene. The most conserved downstream gene encodes a protein that contains the ABC auxiliary domain DUF330. In *Spingobium japonicum*, this ABC transporter is essential for the uptake and utilisation of γ-hexachlorocyclohexane in *S. japonicum*, an insecticide used as its sole carbon source (Endo et al., 2007). In other Proteobacteria type I

neighbourhoods, the downstream genes encode proteins that contain STAS and Tol domains. In *E. coli*, these proteins make up the phospholipid trafficking Mla pathway (MlaBCDEF) (Malinverni and Silhavy, 2009). This operon has been studied in other Proteobacteria, including *Pseudomonas putida*, where it is required for toluene resistance and *Shigella flexneri,* where it is needed for intracellular spreading (Kim et al., 1998; Hong et al., 1998).

In Proteobacteria, the neighbourhoods of genes that encode multi-domain proteins do not clearly resemble an ABC transporter. Although genes that encode DUF330 are often found downstream of type III protein-encoding genes, genes that encode two PqiA domains replace the upstream ATPase and permease genes in both type III and IV neighbourhoods. As the PqiA domain resembles an NADH antiporter domain it is possible that these operons encode a different type of transporter that lacks direct binding of ATP (i.e. not strictly ABC) (Finn et al., 2016; Hunter et al., 2009). This suggests that the mechanism of action of type III and IV transport might differ from type I/II proteins.

The operons that encode the two multi-domain proteins in *E. coli* (*pqiAB-ymbA* and *yebST*) were found to resemble the most common type III and IV gene neighbourhoods, respectively. The tight clustering and conservation of type III and IV proteins in Proteobacteria indicates that they fulfil an important and possibly unique function in this phylum. In the next chapter, phenotypes of *pqiAB* and *yebST* deletion strains and properties of PqiB and YebT are investigated in order to shed light on the function of these conserved domains in *E. coli*.

CHAPTER 4

IDENTIFICATION OF NOVEL PHENOTYPES, PROTEIN LOCALISATION AND LIPIDOMICS PROVIDES AN INSIGHT INTO THE FUNCTION OF MCE DOMAINS IN *E. COLI*

4.1. Introduction

Bioinformatic analysis has confirmed the presence of 3 MCE domain-containing proteins in *E. coli*: MlaD, PqiB and YebT. There are 1, 3 and 7 MCE domains in MlaD, PqiB and YebT, respectively. MlaD is reported to be part of a phospholipid trafficking pathway, but the functions of PqiB and YebT remain unknown. The genes *pqiB* and *yebT* are found directly downstream of *pqiA* and *yebS*, respectively. The gene *ymbA* is downstream of *pqiAB* and is predicted to be in the same operon (Petersen et al., 2011; Taboada et al., 2012). In the previous chapter, bioinformatic analysis revealed that both multi-MCE domain operons resemble a type of transporter.

The genes *pqiA* and *pqiB* were named following the observation that they were "paraquat induced" and therefore might be involved in protection against reactive oxygen species (ROS) (Koh and Roe, 1996). The *pqiAB* genes are regulated by SoxRS in *E. coli*, a regulon that directly governs the response to superoxide radicals (Koh and Roe, 1995; Koh and Roe, 1996). ROS play a key part in the immune response to invading microorganisms, particularly in phagocytes. The *pqiAB* operon is also induced by membrane stress from overproduction of fatty acids and by cyclic-di-GMP, a compound that regulates virulence, adherence and motility in prokaryotes (Mendez-Ortiz et al., 2006). Furthermore, both *pqiAB* and *yebS* are regulated by σ^s, the sigma factor that regulates starvation and stationary phase responses (Weber et al., 2005). The best-studied multi MCE-domain containing protein is Multivalent Adhesion Molecule- 7 (MAM-7) from *Vibrio parahaemolyticus*, which is a homologue of YebT. MAM-7 is reported as an outer membrane adhesin that binds to mammalian cells via phosphatidic acid and fibronectin (Krachler et al., 2011; Krachler and Orth, 2011). As a whole, the literature on PqiB and YebT signifies a role in virulence and protection again stress. However, no specific function has yet been identified for these operons in *E. coli*.

Of all the MCE domain containing proteins in *E. coli*, only the location of MlaD has been determined experimentally. MlaD has been characterised as an inner membrane periplasmic-facing protein with an N-terminal transmembrane (TM) domain (Malinverni

and Silhavy, 2009). This type I protein is part of the Mla pathway in *E. coli*, which is required for phospholipid trafficking from the outer membrane back into the cell. Strains in which individual components of the *mlaFEDBC* operon have been deleted accumulate phospholipids on the cell surface suggesting that the pathway is involved in removing unwanted phospholipids from the outer leaflet of the outer membrane (Malinverni and Silhavy, 2009). However, no research has thus far shown whether disruption of the Mla pathway results in altered lipid composition of the inner and outer membranes, or whether the Mla proteins directly interact with phospholipids.

To address what is unknown about MCE domain containing proteins, the aims of this chapter were as follows: 1) phenotypically characterise mutants of *pqiAB* and *yebST* in *E. coli* to determine the function of the two operons; 2) determine the locations of PqiB and YebT; 3) investigate the lipid-binding properties of MlaD, PqiB and YebT and 4) determine whether MCE mutants in *E. coli* have altered membrane phospholipid compositions.

4.2. Results

4.2.1. Construction of a pqiAB *deletion in* E. coli *K-12 BW25113*

In this study all gene deletions were constructed in *E. coli* K-12 BW25113, the parent strain for the Keio collection of single gene deletions that is closely related to the genome-sequenced strain W3110 (Baba et al., 2006). The *pqiAB* genes span positions 1012423 to 1015321 on the *E. coli* K-12 W3110 genome, the parental strain of BW25113. The genes are located between *uup* and *ymbA*. To investigate the function of *pqiAB*, deletions were constructed in *E. coli* K-12 BW25113 using λ Red recombination. First, a linear fragment for recombination was generated by PCR using the plasmid pKD4 as a template and pqiA_DW_F and pqiB_DW_R primers. The primers were designed from the *E. coli* K-12 W3110 genome to generate a kanamycin resistance cassette flanked by 50 bp of the sequence found directly upstream and downstream of *pqiAB*.

E. coli strain BW25113 was transformed with pKD46, a temperature sensitive vector harbouring the λ Red recombination machinery. The transformant was electroporated with

the *pqiAB* linear fragment to allow recombination. Kanamycin resistant colonies were checked by colony PCR for deletion of the *pqiAB* genes using pqiA_check_F and pqiB_check_R primers, which annealed to the flanking regions of *pqiAB* (Figure 4.1A). The PCR included BW251113 WT as a control, and a successful deletion was shown by a decrease in PCR product from ~3000 bp to ~2000 bp (Figure 4.1B, lanes 2 and 3). The confirmed *pqiAB::aph* deletion was transduced into a clean background of BW25113 using P1 bacteriophage to leave behind any second site mutations caused by λ Red recombination. Successful transduction was confirmed by colony PCR using the same check primers.

To remove the kanamycin resistance cassette, the *pqiAB::aph* mutant was transformed with the temperature sensitive plasmid pCP20, which encodes FLP recombinase. Successful removal of the kanamycin resistance gene resulted in a decrease in PCR product size from ~2000 bp to ~500 bp (Figure 4.1B, lanes 3, 4 and 5).

4.2.2. Construction of a yebST *deletion in* E. coli *K-12 BW25113*

The *yebST* genes span positions 1917972 to 1921857 on the W3110 chromosome, located between *msrC* and *rsmF*. To investigate the function of *yebST*, deletions were made in BW25113 using the same protocol described for construction of the *pqiAB* mutants. A linear *yebST* fragment generated by PCR using the primer pair yebS_DW_F and yebT_DW_R was electroporated into *E. coli* strain BW25113, which had been previously transformed with pKD46. Kanamycin resistant colonies were checked by colony PCR using the primer pair yebS_check_F and yebT_check_R (Figure 4.1A). A decrease from ~4500 bp to ~2000 bp showed successful deletion of *yebST* (Figure 4.1B, lanes 6 and 7). The deletion was transferred into a clean BW25113 parent background using P1 transduction. For construction of a kanamycin sensitive *yebST* deletion, the *yebST::aph* mutant was transformed with pCP20. Colony PCR confirmed the removal of the cassette by a decrease in size from ~2000 bp to ~ 500 bp (Figure 4.1B, lanes 7, 8 and 9).

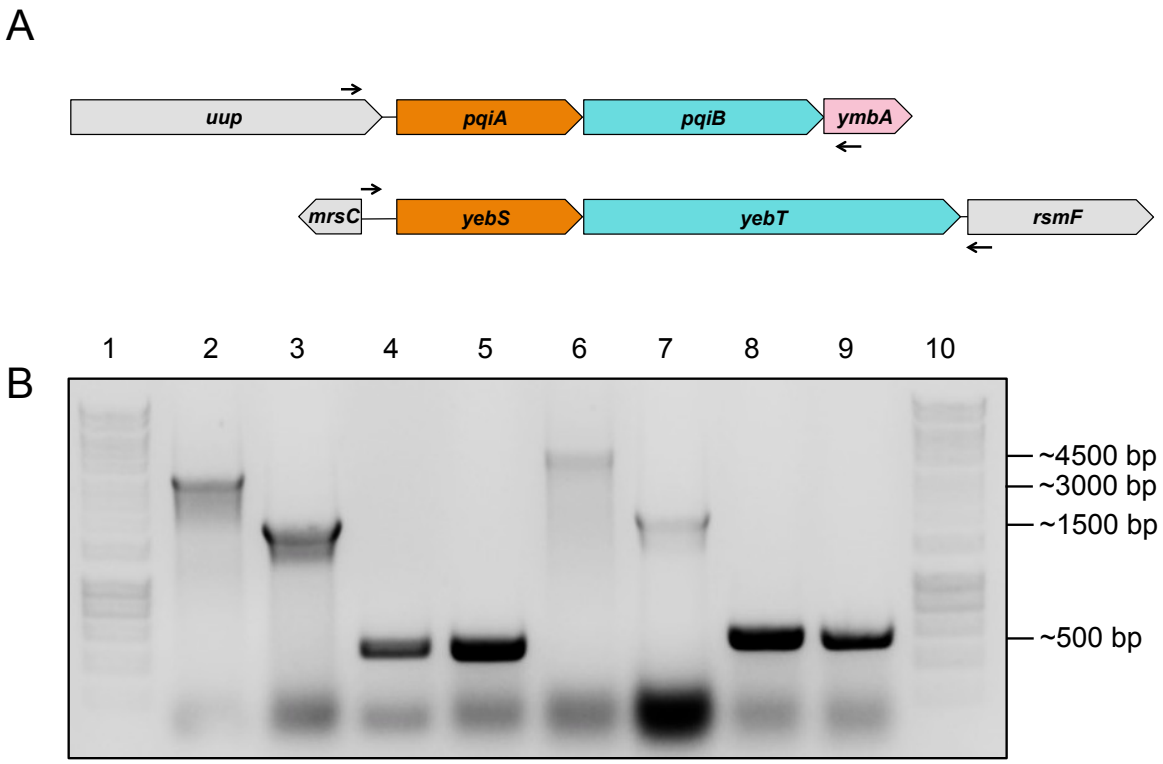


Figure 4.1. Gel electrophoresis of PCR products confirming the disruption of *pqiAB* and *yebST* in *E. coli* K-12 BW25113. **A.** Positions of *pqiAB* and *yebST* check primers. **B.** Results from PCR using check primers: lanes 2 to 5 show PCR products obtained using *pqiAB* check primers, lane 2: WT, lane 3: *pqiAB::aph,* lanes 4 and 5: Δ*pqiAB* deletion strain. Lanes 6 to 9 show PCR products generated using *yebST* check primers, lane 6: WT, lane 7: *yebST::aph*, lanes 8 and 9: Δ*yebST* deletion strain.

4.2.3. Construction of a pqiAB yebST *double deletion in* E. coli *K-12 BW25113*

To construct a *pqiAB yebST* double deletion, the Δ*yebST* mutant was transduced with the same P1 phage lysate that had been used to construct the *pqiAB::aph* strain. Kanamycin resistant colonies were checked with both sets of check primers by colony PCR. To remove the kanamycin resistance cassette, the Δ*pqiAB yebST::aph* mutant was transformed with pCP20 and a kanamycin sensitive double deletion was constructed. The strain was further checked by colony PCR. All of the kanamycin sensitive strains were further checked by whole genome sequencing and confirmed the gene deletions and the absence of secondary mutations.

4.2.4. Biolog screen of deletion mutants

The parent strain BW25113, the *pqiAB* deletion mutant, the Δ*yebST* mutant and the double mutant were sent to Biolog Inc., USA, for phenotypic microarray analysis. This service provides growth kinetics of the desired strain in approximately 2000 different conditions, measured relative to the parent strain. In the results provided by Biolog, the Δ*pqiAB* mutant was reported to be more sensitive to the lipophilic chelator clioquinol and to the zwitterionic detergent lauryl sulfobetaine (LSB) than the parent strain (Figure 4.2A). The Δ*yebST* mutant was also more sensitive to clioquinol and to the antibiotics tetracycline and penimepicycline (Figure 4.2B). The double deletion strain was sensitive to LSB but more resistant to the antibiotic azlocillin (Figure 4.2C). The full results from the Biolog analyses can be found in appendix iii.

4.2.5. Preliminary attempts to confirm the Biolog data

The Biolog analysis included results for a concentration range for each of these five chemicals as follows: clioquinol (1-100 μg/ml), LSB (0.02-2% (w/v)), tetracycline (0.2-50 μg/ml), penimepicycline (1.5-50 μg/ml) and azlocillin (10-250 μg/ml). In initial attempts

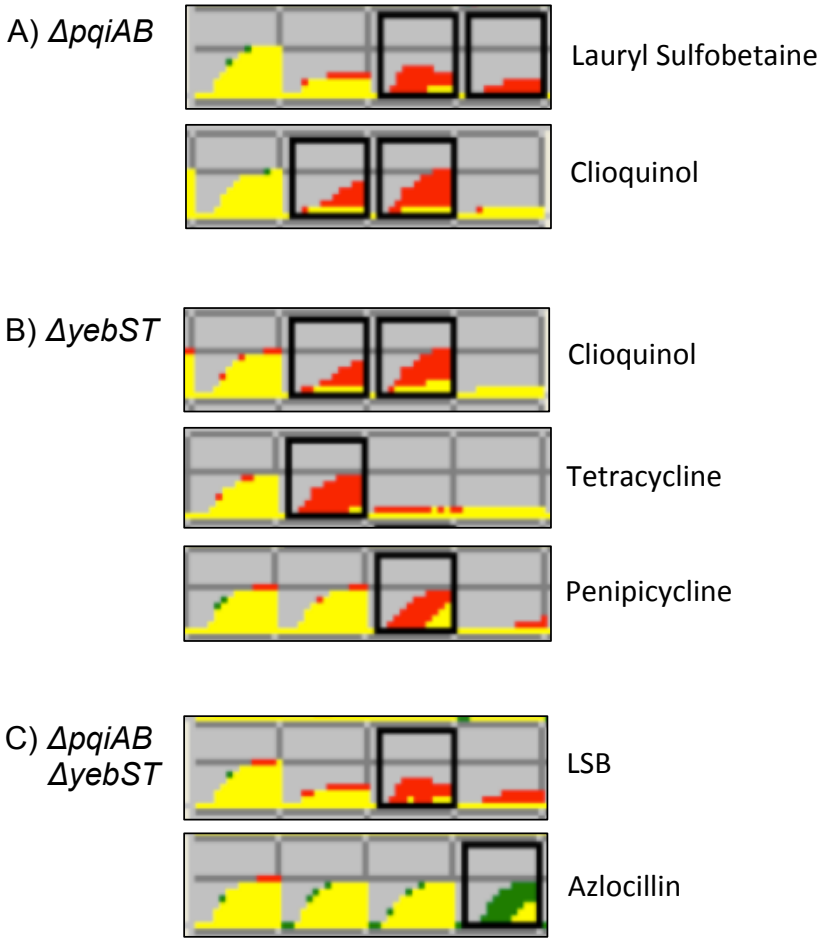


Figure 4.2. Results of the Biolog screen for the Δ*pqiAB*, Δ*yebST* and Δ*pqiAB* Δ*yebST* mutants vs the parent. Red represents growth of the parent and green represents growth of the mutants. Each row shows growth in a particular chemical at four different concentrations. **A.** Growth of the Δ*pqiAB* mutant relative to the parent strain on lauryl sulfobetaine (0.02-2% (w/v)) and clioquinol (1-100 µg/ml). **B.** Growth of the Δ*yebST* mutant relative to the parent strain on clioquinol, tetracycline (0.2-50 µg/ml) and penimepicycline (1.5-50 µg/ml). **C.** Growth of the Δ*pqiAB* Δ*yebST* mutant relative to the parent strain on lauryl sulfobetaine and azlocillin (10-250 µg/ml).

to confirm the Biolog results, overnight cultures of all four strains were inoculated into a 96-well microtitre plate containing 200 μl of LB supplemented with the same range of concentrations of these chemicals. The strains were grown from a starting OD_{600} of 0.05 for 16 hours at 37°C in a FLUOstar Optima plate reader. No growth curve was obtained for clioquinol due to insolubility of the compound. There were no substantial growth differences between the 4 strains when cultured in the presence of tetracycline, penimepicycline, or azlocillin. Therefore, it was concluded that the Biolog data were not reproducible (Figure 4.3 A to C). In contrast to the other chemicals tested, LSB inhibited growth of both the Δ*pqiAB* mutant and the double mutant (Figure 4.4). Although the Δ*yebST* mutant also initially appeared to have a minor growth defect, after 7 hours the OD_{600} of these cultures was the same as the parent. Subsequent experiments focused on the effect of LSB on these strains on a larger scale.

4.2.6. Confirmation of the effect of lauryl sulfobetaine on growth of the four strains

The LSB phenotype was explored further through growth of the strains in 25 ml of LB supplemented with 1% (w/v) LSB, in shake flasks. All four strains grew the same in LB (Figure 4.5A). In LB supplemented with 1% (w/v) LSB, the phenotype initially matched the phenotype shown by Biolog, with decreased growth for the Δ*pqiAB* mutant and the double mutant, but no growth inhibition for the Δ*yebST* single mutant (Figure 4.5B). However after 3 to 4 hours all of the strains lysed rapidly, a phenomenon often observed during bacterial growth in liquids containing detergents (Kramer, 1980). These preliminary experiments confirmed the LSB phenotype, but indicated that liquid broth might not be the best method to assay the effect of this detergent on bacterial growth.

4.2.7. Construction of mlaD *gene deletions*

The *mlaD* gene encodes a protein with a single MCE domain. Having established a phenotype for the Δ*pqiAB* and double deletions strains, it was appropriate to test the

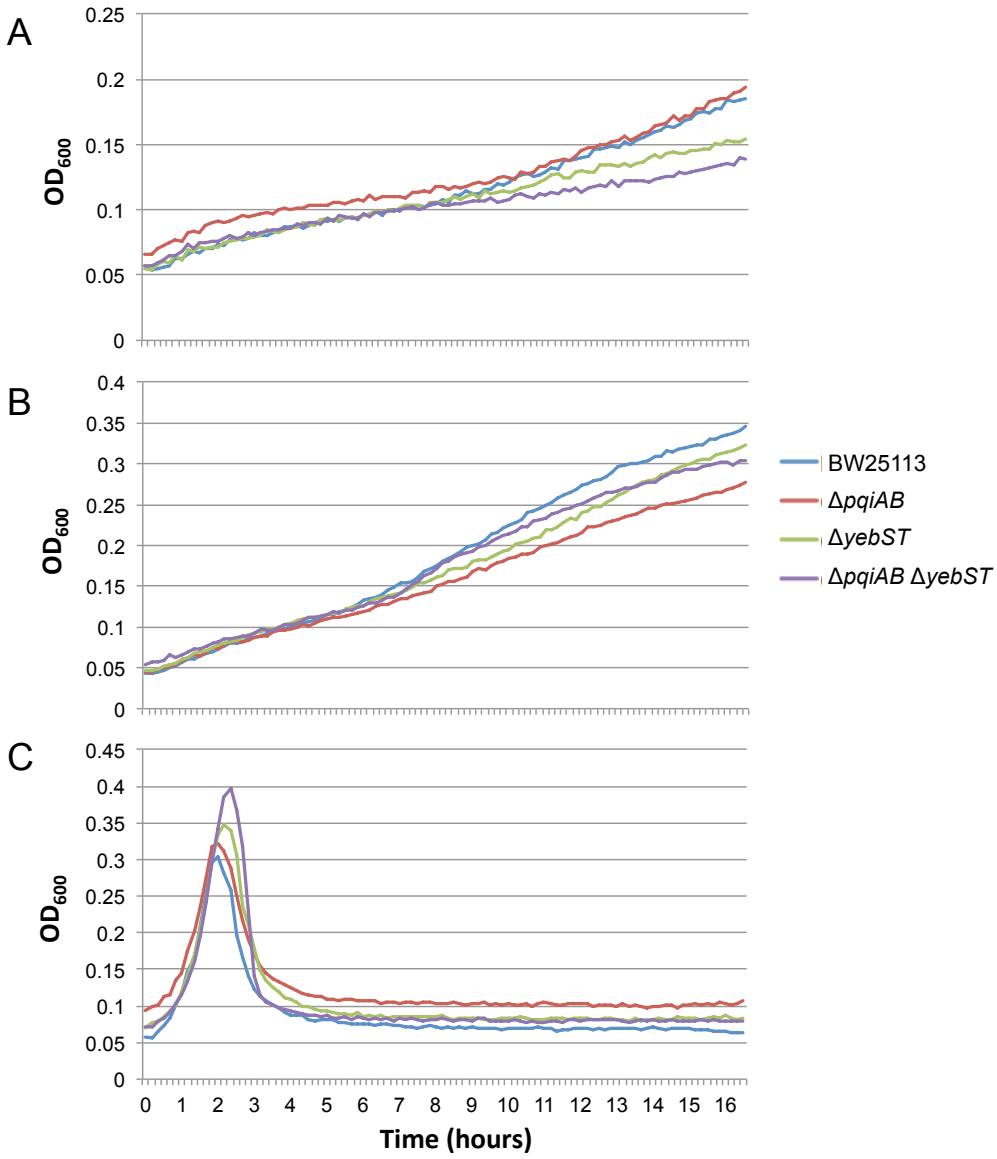


Figure 4.3. Growth of *E. coli* strain BW25113, the Δ*pqiAB* mutant, the Δ*yebST* mutant and the double mutant in the presence of tetracycline, penimepycycline and azlocillin. Each strain was challenged with the range of concentrations used in the Biolog assay. The graphs represent one concentration from each of these ranges **A**. Growth in 10 μg/ml tetracycline, **B**. Growth in 10 μg/ml penimepycline, **C**. Growth in 100 μg/ml azlocillin. Strains were grown in a microtitre plate in the FLUOstar Optima and readings were taken every 5 minutes.

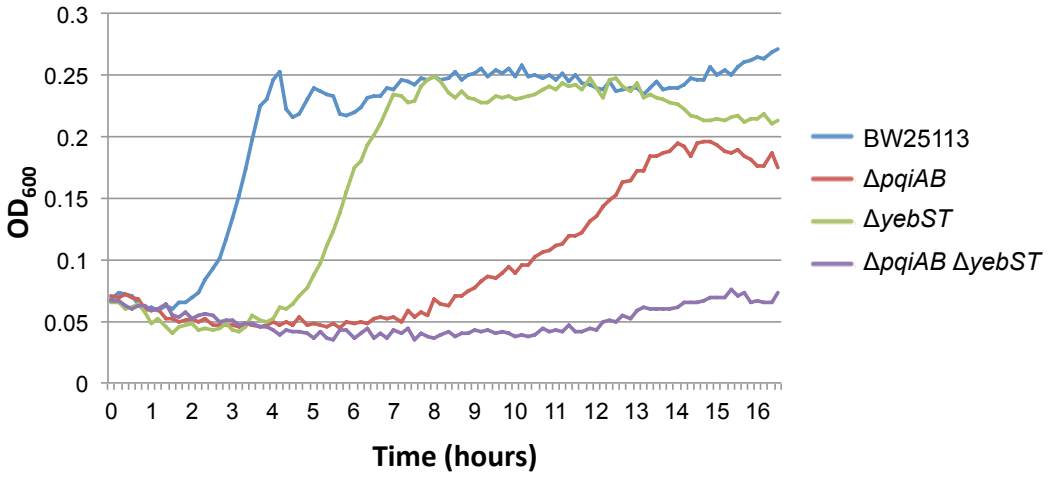


Figure 4.4. Growth curves of BW25113 and the three mutants in 1% (w/v) lauryl sulfobetaine in a microtitre plate. Strains were grown in the FLUOstar Optima and readings were taken every 5 minutes.

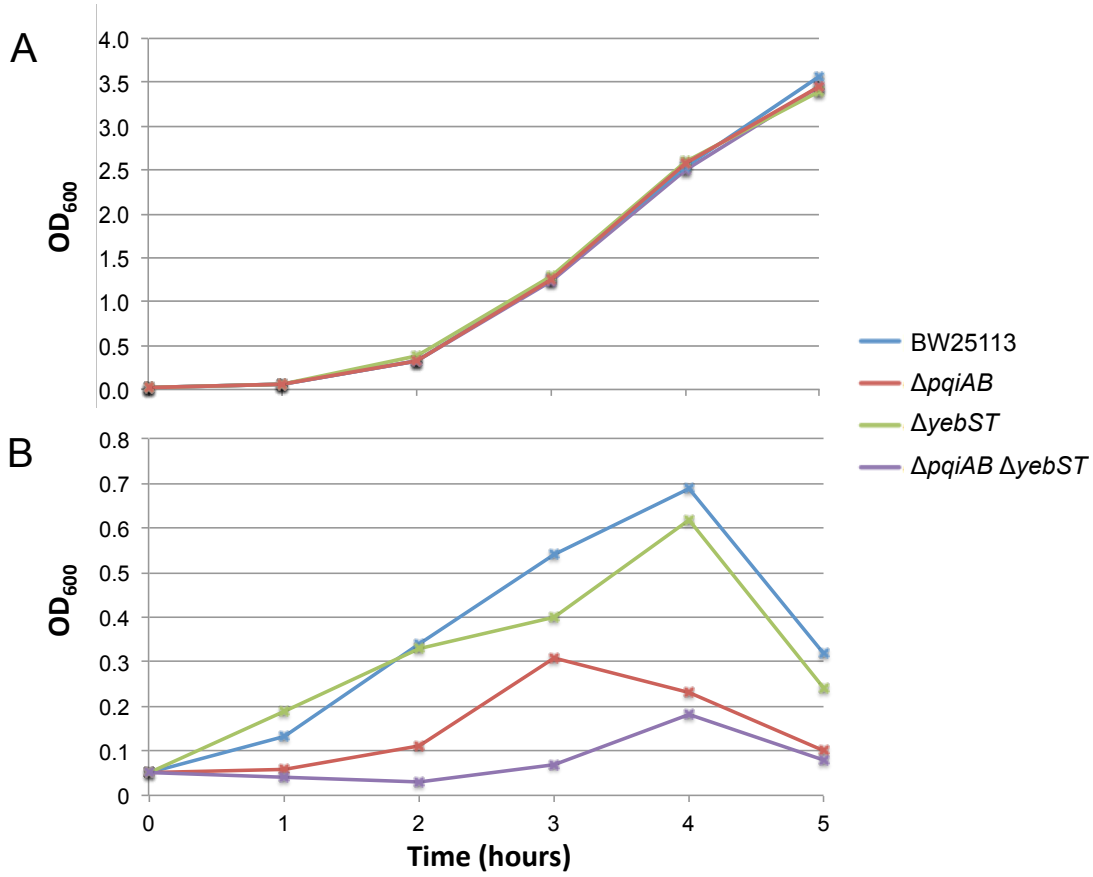


Figure 4.5. Growth curves of BW25113 and the Δ*pqiAB*, Δ*yebST* and Δ*pqiAB* Δ*yebST* mutants in **A.** LB and **B.** LB supplemented with 1% (w/v) lauryl sulfobetaine in shake flasks.

phenotype of a strain defective in *mlaD*. The *mlaD* gene spans positions 3338316 to 3337765 on the W3110 genome as part of the *mlaFEDCB* operon. Two copies of an *mlaD::aph* insertion mutant were already available in the Keio library (Baba et al. 2006). This library was made in *E. coli* strain BW25113 using lambda red recombination (Datsenko and Wanner, 2000). The two *mlaD* mutants were checked by colony PCR using mlaD_check_F and mlaD_check_R primers (Figure 4.6A). An increase in the size of the PCR product from ~1000 bp to ~2000 bp indicated that the *mlaD* gene had been replaced with a kanamycin resistance cassette. These data indicate that both strains in the Keio library harbour deletions in *mlaD* (Figure 4.6B, lanes 2, 3 and 4).

The *mlaD::aph* mutation from one of these strains was transduced from the Keio library strain into the BW25113 parent strain and the Δ*pqiAB,* Δ*yebST* and Δ*pqiAB* Δ*yebST* mutants and was checked again by colony PCR. This produced 4 new kanamycin resistant strains: *mlaD::aph,* Δ*pqiAB mlaD::aph,* Δ*yebST mlaD::aph* and Δ*pqiAB* Δ*yebST mlaD::aph*. To construct kanamycin sensitive derivatives, all 4 strains were transformed with pCP20 to remove the kanamycin resistance cassette and several colonies were checked by colony PCR. This resulted in the construction of 4 new kanamycin sensitive strains: Δ*mlaD,* Δ*mlaD ΔpqiAB*, Δ*mlaD* Δ*yebST* and Δ*mlaD ΔpqiAB* Δ*yebST*.

4.2.8. Screening of all seven deletions on lauryl sulfobetaine

Due to the cell lysis observed during growth in LB with 1% (w/v) LSB, the phenotypes were tested on agar plates. The parent strain and all seven mutants were diluted from overnight cultures to an OD_{600} of 1 and diluted serially down to 10^{-5}. The dilutions were plated onto square plates with LA, LA supplemented with 1% (w/v) LSB, and LA supplemented with 4% (w/v) LSB. All strains grew equally well on LA (Figure 4.7A). On LA supplemented with 1% (w/v) LSB, growth of the Δ*mlaD* and Δ*pqiAB* mutants was observed down to a dilution of 10^{-3}, whilst BW25113 grew down to a dilution of 10^{-5} (Figure 4.7B). Growth of the Δ*yebST* mutant was not inhibited. The growth of the Δ*mlaD* Δ*pqiAB* mutant was further reduced to 10^{0}, showing a much greater growth defect than either of individual deletion strains.

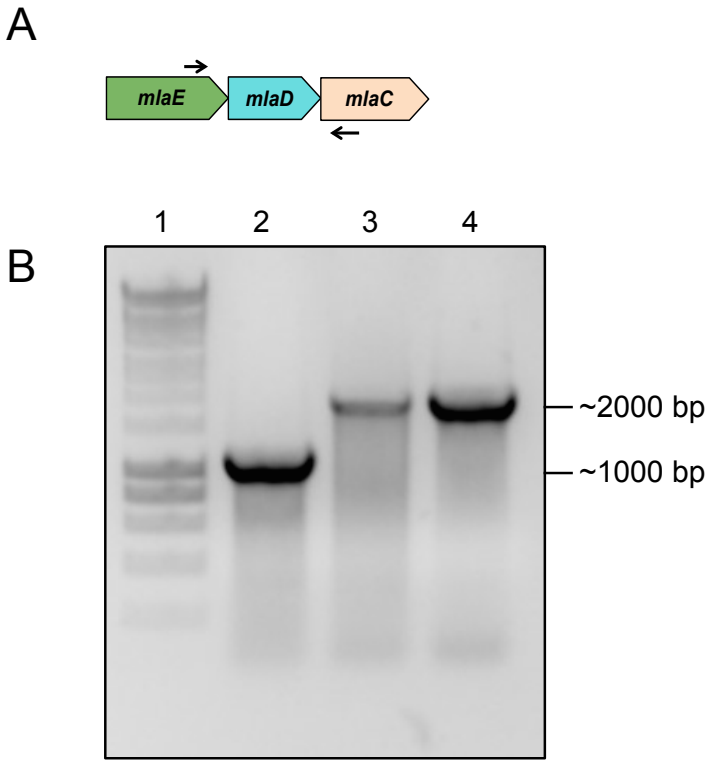


Figure 4.6. Gel electrophoresis of PCR products confirming the disruption of *mlaD* in *E. coli* K-12 BW25113 from the Keio library. **A.** Position of *mlaD* check primers. **B.** Results from PCR using check primers: Lane1: ladder, lane 2: BW25113, lanes 3 and 4: *mlaD::aph*.

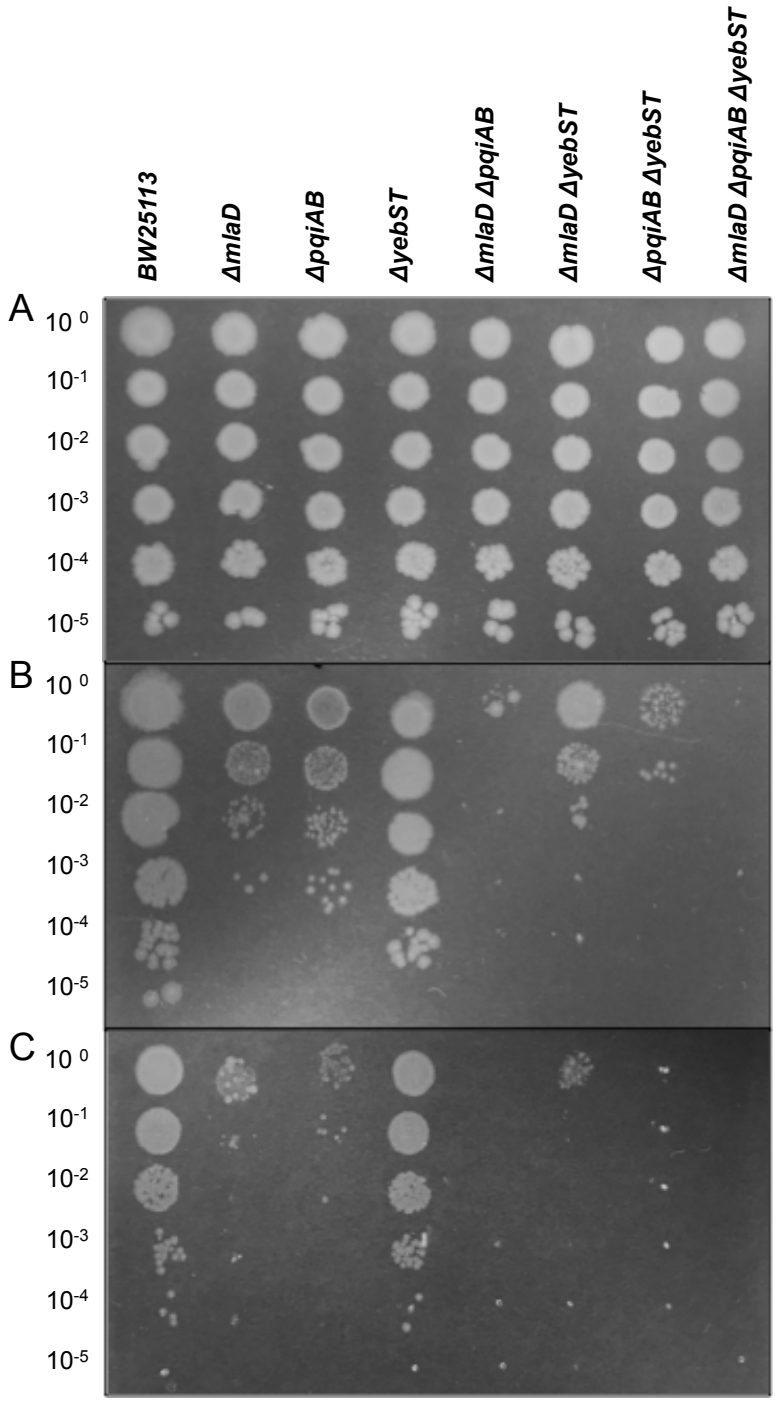


Figure 4.7. Growth of the parent and all *mlaD, pqiAB* and *yebST* mutant combination on lauryl sulfobetaine (LSB). The strains were plated onto **A**. LA, **B**. LA supplemented with 1% (w/v) LSB and **C**. LA supplemented with 4% (w/v) LSB.

Growth of the Δ*pqiAB* Δ*yebST* mutant was only observed down to 10^{-1} dilution, showing a greater growth defect than the Δ*pqiAB* mutant. The Δ*mlaD* Δ*yebST* mutant only grew marginally worse than the Δ*mlaD* mutant. No growth was observed for the triple mutant on agar supplemented with 1% (w/v) LSB. On agar supplemented with 4% (w/v) LSB, the same pattern of phenotypes were observed but the sensitive mutants appeared to have an even greater growth defect when compared to the parent strain (Figure 4.7C).

Overall, the phenotypes of the *pqiAB* mutants were clearer on agar plates than in broth. The plates revealed similar sensitivity phenotypes for the Δ*mlaD* and Δ*pqiAB* mutants, and a clear additive phenotype for the Δ*mlaD* Δ*pqiAB* mutant. In agreement with the Biolog results, the Δ*yebST* mutant was not sensitive to LSB, but increased the sensitivity of a Δ*pqiAB* mutant. The largest growth defect was seen for the triple deletion strain, which suggested that increased sensitivity on LSB occurs with increased loss of MCE domain encoding genes.

4.2.9. Complementation of the lauryl sulfobetaine phenotypes

In order to confirm the phenotypes observed in the previous section, complementation plasmids were constructed in the plasmid pET17b. The *pqiAB*, *yebST,* and *mlaD* genes were amplified using the following primer pairs: 1) pqiA_NdeI_F and pqiB_comp_R, 2) yebS_NdeI_F and yebT_comp_R, and 3) mlaD_comp_F and mlaD_comp_R. The *pqiAB*, *yebST* and *mlaD* fragments were purified and digested with NdeI and XhoI restriction enzymes. The fragments were ligated into pET17b digested with the same enzymes and transformed into 5-alpha *E. coli* competent cells. Several colonies were screened by PCR using T7_F and T7_R primers. Plasmids that contained an insert of the correct size were sequenced using the same primers. The 3 plasmid constructs were designated pET17b-*mlaD*, pET17b-*pqiAB* and pET17b-*yebST*.

For complementation, pET17b-*mlaD*, pET17b-*pqiAB* and pET17b-*yebST* were transformed into the Δ*mlaD*, Δ*pqiAB* and Δ*pqiAB* Δ*yebST* mutants, respectively. To allow growth on carbenicillin, the pET17b plasmid was transformed into the same strains and

BW25113. The strains were prepared as described previously and plated onto agar plates supplemented with carbenicillin and 1% (w/v) LSB (Figure 4.8). The plasmids pET17b-*mlaD* and pET17b-*pqiAB* fully complemented the phenotypes of the Δ*mlaD* and Δ*pqiAB* mutants, respectively, on LSB. The plasmid pET17b-*yebST* partially complemented the phenotype of the Δ*pqiAB* Δ*yebST* mutant, back to the level of the Δ*pqiAB* mutant, confirming that the additive Δ*pqiAB* Δ*yebST* phenotype was due to the deletion in *yebST*. Overall, these results confirm that the LSB phenotypes observed can be complemented by reintroduction of the deleted gene(s) into the deletion strains.

4.2.10. Screen of mutants on other sulfobetaines

To investigate whether detergents similar to LSB also inhibited growth of the mutants, other sulfobetaine detergents were tested. LSB is also known as sulfobetaine (SB) 3-12 as it has a 3 and a 12-carbon chain that meet at an ammonium ion. Sulfobetaine 3-10 (SB 3-10) and sulfobetaine 3-14 (SB 3-14) differ from LSB (or SB 3-12) by their hydrocarbon chain length only. Where LSB has a 12-long hydrocarbon chain, SB 3-10 and SB 3-14 have 10 and 14 carbons, respectively (Figure 4.9). The parent and all 7 deletions strains were plated as described previously onto square agar plates with 1% (w/v) SB3-10 and SB3-14. On SB 3-10, an LSB-like phenotype was observed for the *pqiAB* mutants, including the Δ*pqiAB* Δ*yebST* additive phenotype. However the additive phenotype was less extreme than the phenotype observed on LSB (Figure 4.10A and B). Deletions in *mlaD* had no effect on the ability of the strains to grow. For SB 3-14, there were no observed growth defects for any of the deletion strains (Figure 4.10C). Increasing the concentration to 4% (w/v) SB3-10 resulted in very faint growth that was difficult to interpret, and no growth inhibition of any strains was observed on LA supplemented with 4% (w/v) SB 3-14 (see appendix iv). In conclusion, the *pqiAB* mutants were sensitive to both SB-10 and LSB. However, further introduction of a *yebST* deletion had more of an additive effect on LSB than it did on SB-10.

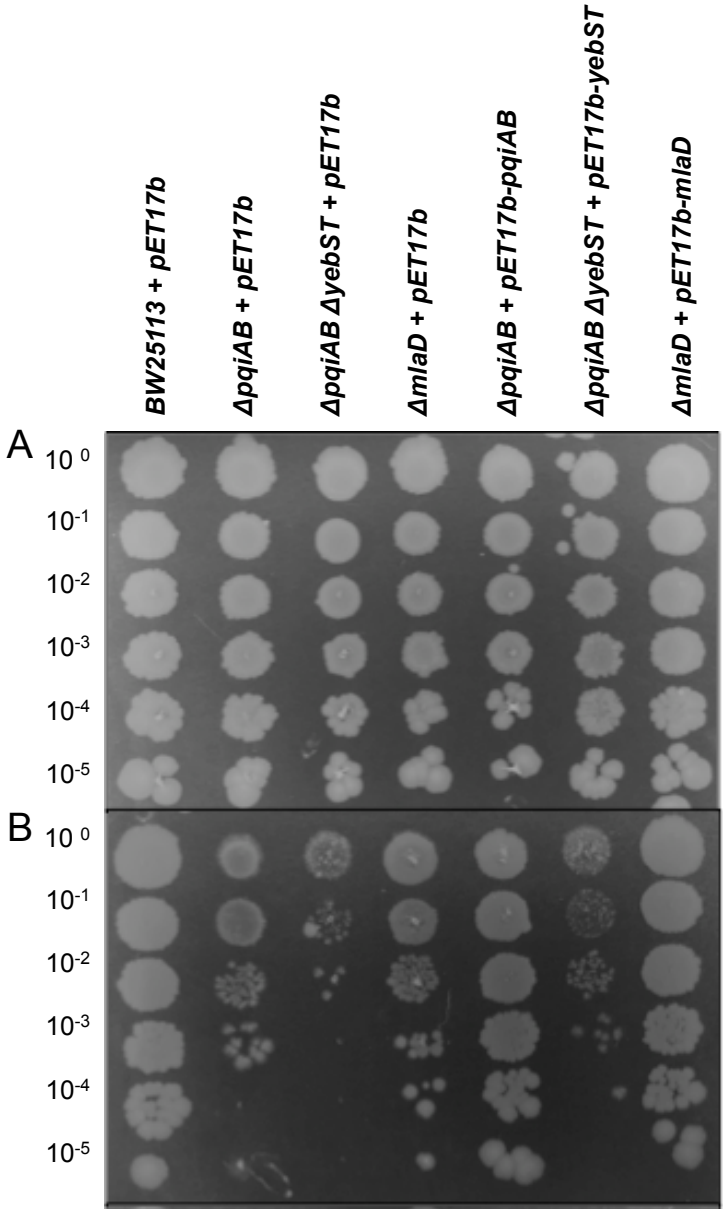


Figure 4.8. Complementation of the Δ*pqiAB,* Δ*pqiAB* Δ*yebST* and Δ*mlaD* mutant phenotypes on lauryl sulfobetaine. The strains were plated onto **A**. plain LA, **B**. LA supplemented with 1% (w/v) LSB and **C**. LA supplemented with 4% (w/v) LSB.

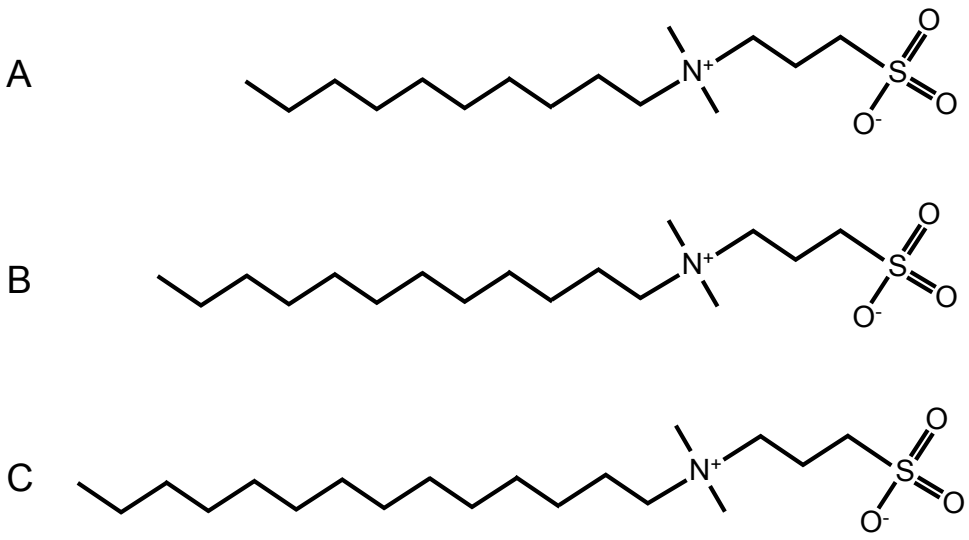


Figure 4.9. The structures of the three sulfobetaines used in this study: **A.** sulfobetaine 3-10, **B.** lauryl sulfobetaine and **C.** sulfobetaine 3-14.

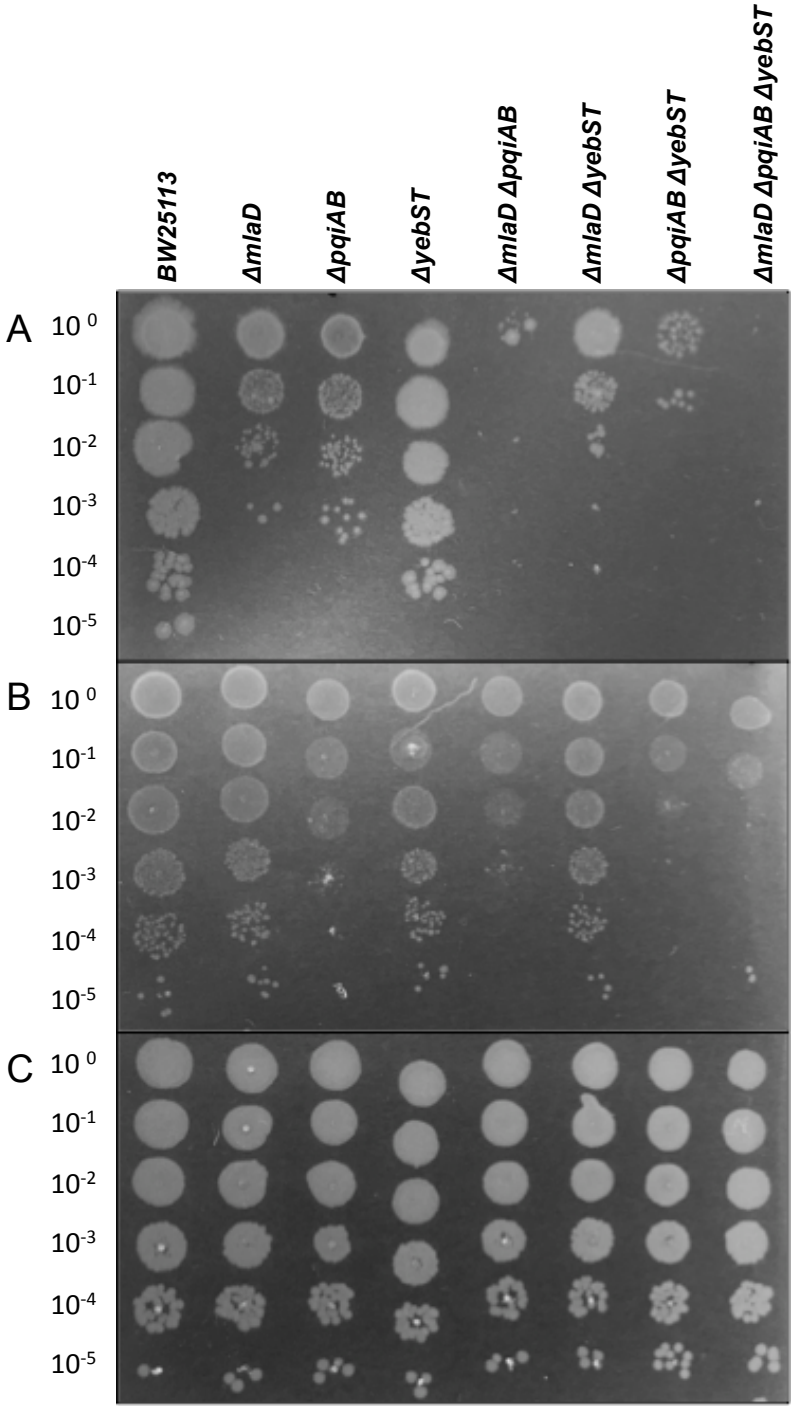


Figure 4.10. Growth of the parent and all *mlaD, pqiAB* and *yebST* mutant combination strains on sulfobetaines with different chain lengths. The strains were plated onto **A**. LA supplemented with 1% (w/v) LSB (taken from figure 4.6 for comparison), **B**. LA supplemented with 1% (w/v) SB3-10 and **C**. LA supplemented with 1% (w/v) SB3- 14.

4.2.11. Screening of MCE mutants for sensitivity to SDS or vancomycin

Increased sensitivity to detergents can indicate a loss of outer membrane integrity. Two chemicals commonly used to screen bacteria for outer membrane defects are sodium dodecyl sulphate (SDS) and vancomycin. Previously, decreased growth was observed for an *mlaD* deletion strain on agar plates supplemented with both SDS and EDTA (Malinverni and Silhavy, 2009). To confirm this phenotype and screen for new phenotypes, the BW25113 parent and mutants were plated onto LA supplemented with 1% (w/v) SDS, 0.5% SDS and 1.1mM EDTA, or 300 μg/ml vancomycin. SDS alone did not inhibit growth of any of the 8 strains (Figure 4.11A). On SDS and EDTA combined, no growth was observed for any strain with a deletion in *mlaD*, but growth of all of the other strains was similar to that of the parent strain (Figure 4.11B). No inhibition was observed for any strain on 1.1 mM EDTA, indicating that the phenotypes observed for the *mlaD* mutants were due to SDS and EDTA combined and not EDTA alone (Figure 4.11C).

On vancomycin, growth of the parent strain was limited to a dilution of 10^{-1}, while growth of all of the *mlaD* mutants was increased to a dilution of 10^{-4} or 10^{-5} (Figure 4.12). Similarly the Δ*pqiAB* Δ*yebST* mutant grew to a dilution of 10^{-5}, whilst there was no growth of the Δ*pqiAB* and Δ*yebST* mutants at dilutions greater than 10^{-1} and 10^{-2}, respectively. This reveals yet another additive phenotype for the Δ*pqiAB* Δ*yebST* mutant. Overall these results reveal a novel vancomycin resistance phenotype for the Δ*pqiAB* Δ*yebST* mutant and all of the *mlaD* mutants, whilst confirming the SDS+EDTA sensitivity phenotype of the Δ*mlaD* mutant.

4.2.12. Screening of single pqiA, pqiB and ymbA mutants for sensitivity to lauryl sulfobetaine

The *pqiAB* genes are predicted to be in an operon with the downstream gene *ymbA*. So far in this study, the phenotypes of *pqiAB* mutants have been investigated, rather than the phenotypes of strains lacking *pqiA*, *pqiB* or *ymbA*. To address this, individual *pqiA::aph*, *pqiB::aph* and *ymbA::aph* insertions were transduced from the Keio library into a clean

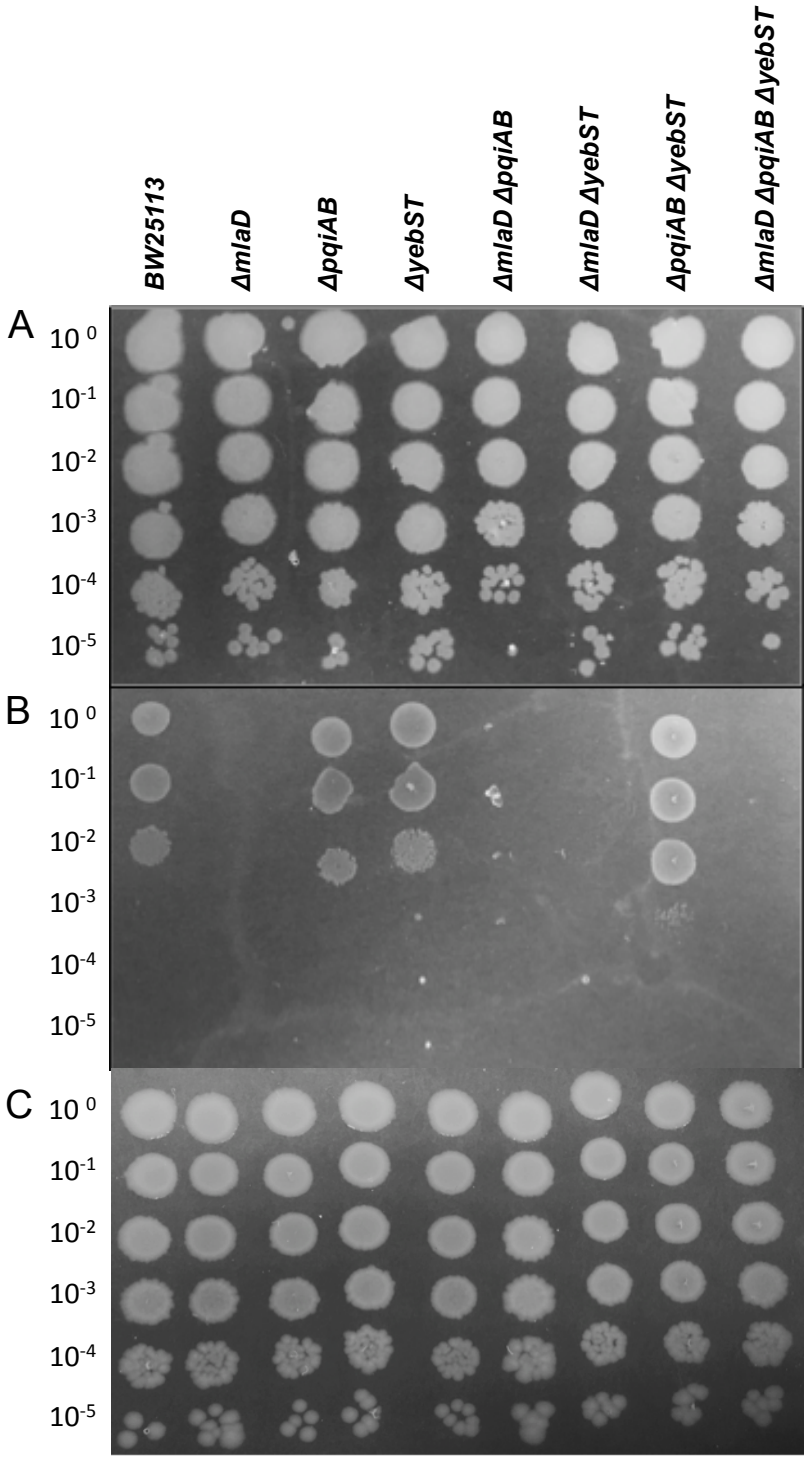


Figure 4.11. Growth of the parent and all *mlaD, pqiAB* and *yebST* mutants on SDS and/or EDTA. The strains were plated onto **A**. LA supplemented with 1 % (w/v) SDS, **B**. LA supplemented with 0.5% (w/v) SDS and 1.1 mM EDTA and **C**. LA supplemented with 1.1 mM EDTA.

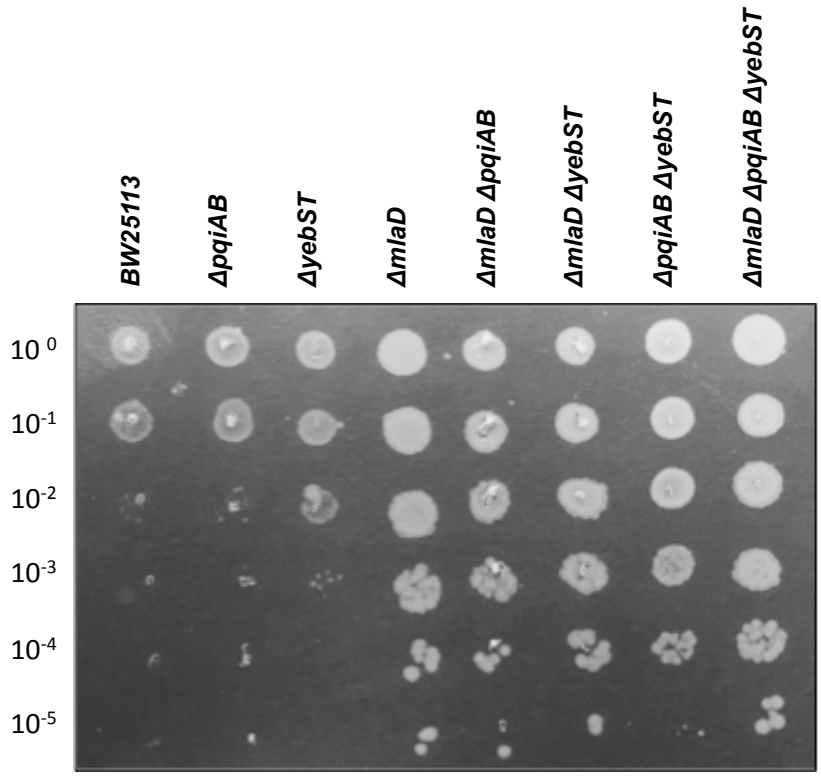


Figure 4.12. The growth the parent and all *mlaD, pqiAB* and *yebST* mutants on vancomycin. The strains were plated onto LA supplemented with 300 μg/ml vancomycin.

BW25113 background. The same primers used to check the *pqiAB* mutants were used to screen potential *pqiA::aph* and *pqiB::aph* mutants, and the *ymbA::aph* mutant was checked with ymbA_check_F and ymbA_check_R primers. The *pqiA::aph*, *pqiB::aph* and *ymbA::aph* mutants were screened for their ability to grow on agar plates in the presence of 1 and 4% (w/v) LSB. The growth of all three deletions strains was observed down to a dilution of 10^{-3} on 1% (w/v) LSB, while the parent strain grew to a dilution of 10^{-5} (Figure 4.13B). On 4% (w/v) LSB, the *pqiA::aph* mutant was less sensitive than the *pqiB::aph* and *ymbA::aph* mutants, but more sensitive than the parent strain (Figure 4.13C). These results indicate that both PqiA and PqiB are required for resistance to LSB. Furthermore, observing the same phenotype for the *ymbA::aph* mutant strain suggests that the three genes are functionally related and supports the prediction that *ymbA* is in the same operon as *pqiAB*.

4.2.13. Screening of mlaC deletion strains for sensitivity to lauryl sulfobetaine

Results presented so far have established that a strain defective in *mlaD* is sensitive to LSB. As the *mlaD* gene is part of a five-gene operon, it was of interest to determine whether the same phenotype would result from a mutation in other genes from this operon. The *mlaC::aph* mutation was therefore transduced from the Keio library into the BW25113 parent and the Δ*pqiAB* and Δ*pqiAB* Δ*yebST* mutants. The primers used at all stages to check *mlaC* mutants were mlaC_check_F and mlaC_check_R.

Growth of these 3 new strains on 1% (w/v) LSB was compared with the equivalent *mlaD* mutants (Figure 4.14). The growth of *mlaC::aph* single and double mutants was similar to that of the corresponding *mlaD* deletion strains, while no growth was observed for either of the triple mutants. As the phenotypes of the *mlaC* and *mlaD* deletion strains were similar, it is possible that all five genes of the *mlaFEDCB* operon are required for LSB resistance.

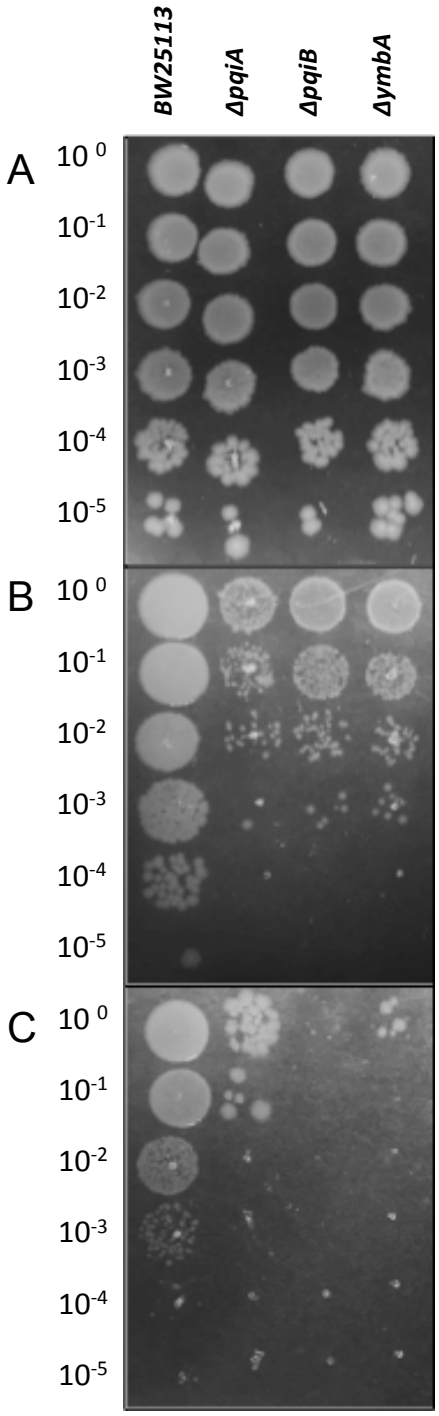


Figure 4.13. The growth of the parent and individual *pqiA::aph*, *pqiB::aph* and *ymbA::aph* mutants on lauryl sulfobetaine. The strains were plated onto **A**. LA, **B**. LA supplemented with 1% (w/v) LSB and **C**. LA supplemented with 4 % (w/v) LSB.

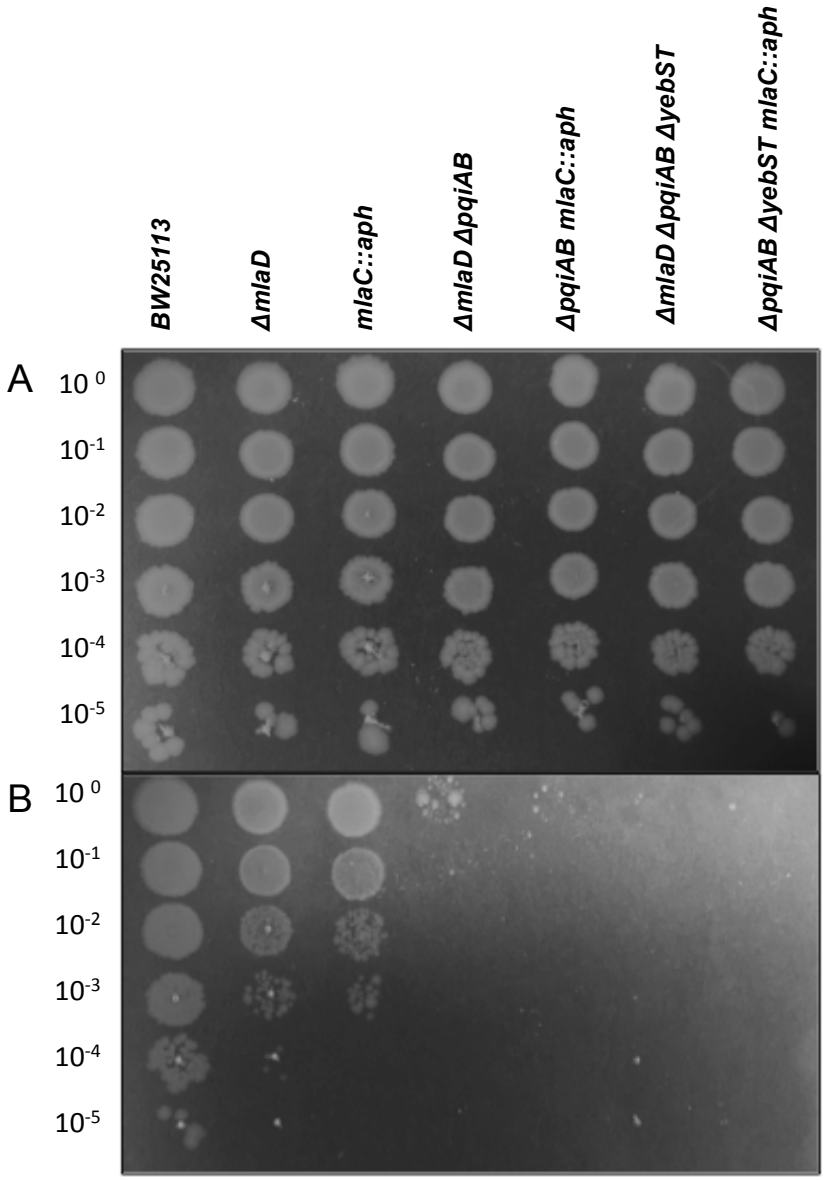


Figure 4.14. The growth of the *mlaC* mutants vs the *mlaD* mutants on lauryl sulfobetaine. The strains were plated onto **A**. LA, **B**. LA supplemented with 1% (w/v) LSB and **C**. LA supplemented with 4% (w/v) LSB.

4.2.14. Determination of whether truncated PqiB and YebT proteins can complement the lauryl sulfobetaine sensitivity phenotype of the pqiAB deletion strain

Aside from 4 additional MCE domains in YebT and the large C-terminal alpha helix in PqiB, the secondary structures of these two proteins were predicted to be very similar (see 3.11). The first aim of these experiments was to identify segments of PqiB necessary for its function. The second aim was to determine whether a "*pqiAB*-like" *yebST* operon could complement the phenotype of the Δ*pqiAB* mutant.

Three truncated operons were synthesised and cloned by Genscript into the restriction sites NdeI and XhoI of the plasmid pET17b. The constructs were as follows: 1) pET17b-*pqiAB*_3, which encodes PqiB without the C-terminal alpha helix; 2) pET17b-*pqiAB*_2α, which encodes PqiB that lacks the most C-terminal MCE domain but has the C-terminal alpha helix still intact; and 3) pET17b- *yebST*_3α, which encodes a truncated YebT protein with the first three MCE domains from YebT and the alpha helix from PqiB added to the C terminus. The constructs were transformed into the Δ*pqiAB* mutant and screened on LA supplemented with 1 and 4% (w/v) LSB (Figure 4.15). In contrast to the vector pET17b-*pqiAB*, which fully complemented the Δ*pqiAB* mutant, none of the other constructs restored any resistance to LSB. This suggests that the C-terminal alpha helix and the C-terminal MCE domain are essential for PqiB function, and that a *"pqiAB*-like" *yebST* operon cannot function as *pqiAB*.

4.2.15. Predictions of PqiA, PqiB, YebS and YebT subcellular locations

Online bioinformatics resources were used to predict the cellular locations of PqiA, PqiB, YebS and YebT. The transmembrane hidden Markov model (TMHMM) server was used to predict transmembrane regions in the proteins and PslPred was used to predict the subcellular locations (Bhasin et al., 2005; Krogh et al., 2001). TMHMM predicted that both PqiA and YebS span the membrane 8 times, with the N- and C-termini located inside the membrane (Figure 4.16A). The N-terminal residues of PqiB and YebT were predicted to

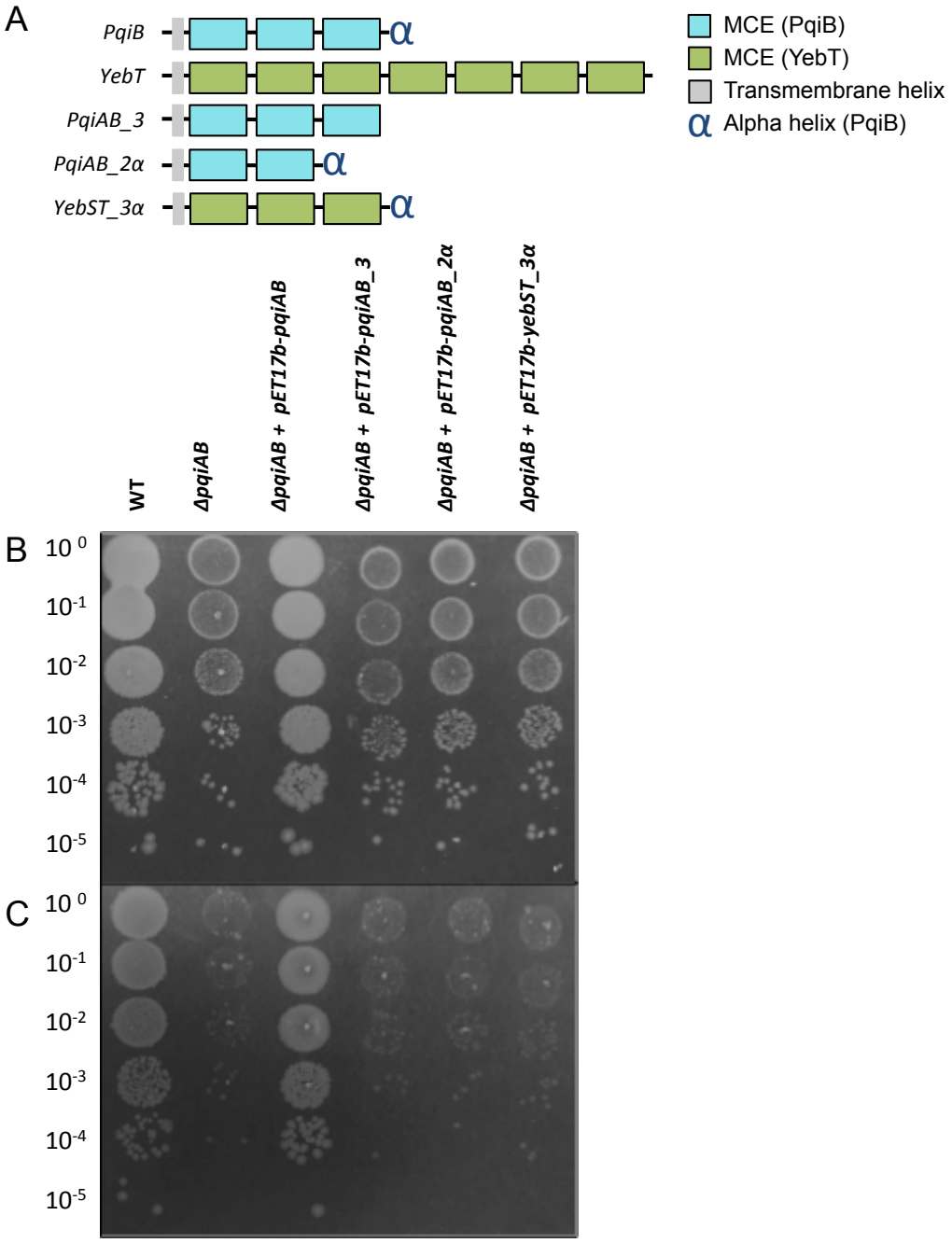
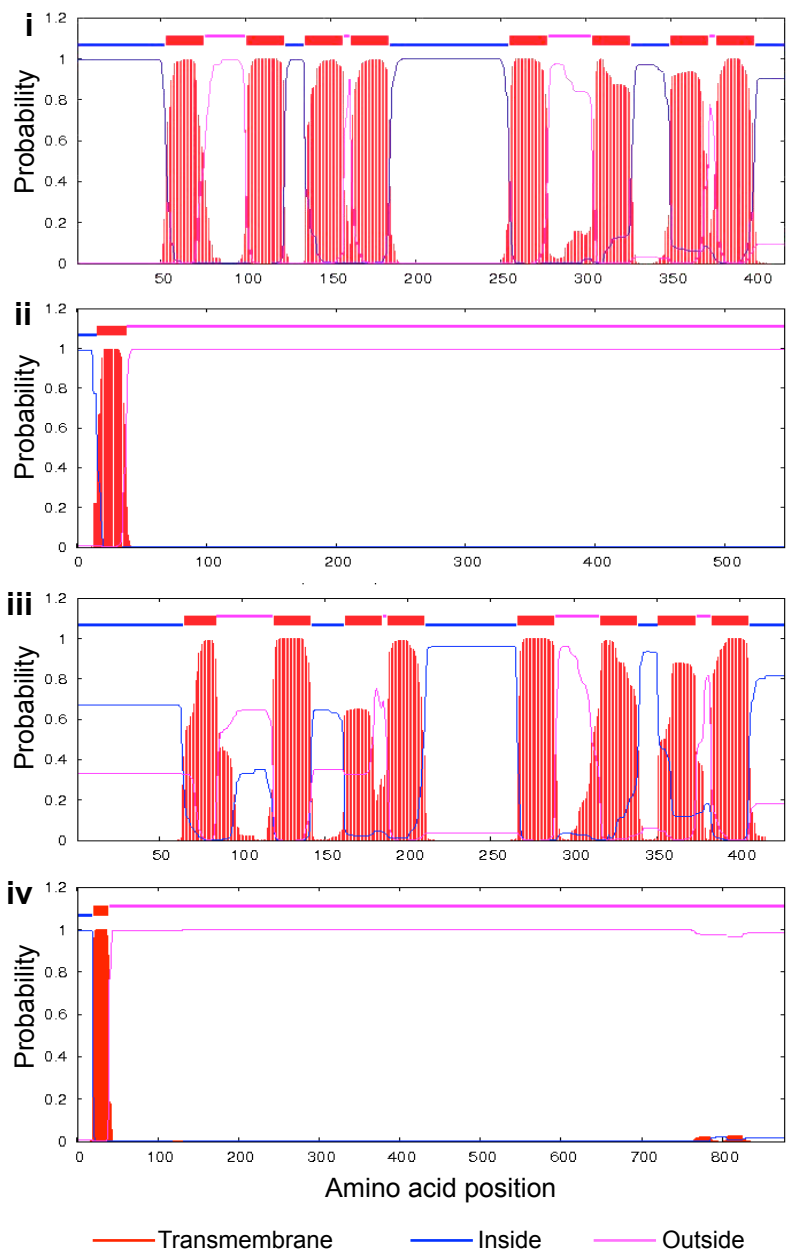


Figure 4.15. The growth of the Δ*pqiAB* mutant with various complementation vectors on lauryl sulfobetaine. **A.** The architectures of PqiB and YebT, **B.** growth of the strains on LA supplemented with 1% (w/v) LSB and **C.** growth of the strains on LA supplemented with 4% (w/v) LSB.

A



B

Protein	Pslpred location
PqiA	Inner membrane: 98.1%
PqiB	Inconclusive
YebS	Inner membrane: 98.1%
YebT	Inconclusive

Figure 4.16. Predictions of the locations of PqiA, PqiB, YebS and YebT using online bioinformatic resources. **A.** The prediction of transmembrane helices using TMHMM. **B.** Subcellular localisation predictions using Pslpred.

span the membrane once with the C-termini located outside of the membrane. PqiA and YebS were predicted by Psipred to locate to the cytoplasmic membrane (Figure 4.16B). The location predictions of PqiB and YebT were inconclusive. These results suggest that all four proteins are membrane localised and that PqiA and YebS are integral inner membrane proteins.

4.2.16. Expression of PqiA, PqiB, YebS and YebT

In order to determine the cellular location of PqiA, PqiB, YebS and YebT experimentally, antibodies specific to each protein were required. Each gene was cloned into pET22b (+), an expression vector with an IPTG-inducible T7 promoter and a polyhistine C-terminal histidine tag. To improve protein solubility, the first 120 bps were removed from *pqiB* and *yebT*, which encode the predicted TM regions. The following primer pairs were used to amplify the genes by PCR: pqiA_NdeI_F and pqiA_XhoI_R; pqiB_TM_NdeI_F and pqiB_XhoI_R; yebS_NdeI_F and yebS_XhoI_R and yebT_TM_NdeI_F and yebT_XhoI_R. The amplified genes were cloned into pET22b using NdeI and XhoI restriction sites and transformed and checked as described previously.

For expression, the pET22b-*pqiA*, pET22b-*pqiB*ΔTM, pET22b-*yebS* and pET22b-*yebT*ΔTM constructs were transformed into the expression strain BL21 DE3. To check for expression, cells were grown to an OD_{600} of 0.6 in 10 ml of LB and induced with 1 mM IPTG for 3 hours at 37°C. PqiA and YebS could not be produced, even when a range of IPTG concentrations were used for induction for various times and at different temperatures. The production of PqiB and YebT was successful and the subsequent purifications are described in the next section.

4.2.17. Purification of PqiB and YebT for antibodies

Without the N-terminal TM regions, the molecular weights of PqiB and YebT are 56 kDa and 85 kDa, respectively. For purification, the cultures expressing PqiB and YebT were scaled up to 1 l and induced under the same conditions. After 3 hours of induction, the cells

were pelleted and lysed and the lysate was centrifuged to remove unbroken cells. The cell lysate was bound to a 1 ml nickel column overnight. The column was washed twice 200 mM imidazole and the proteins were eluted in 1 ml fractions using 500 mM imidazole. The flow-through, washes and eluted fractions from each purification were analysed by SDS-PAGE, revealing purified PqiB and YebT with some minor contaminants (Figure 4.17A and B). A large amount of PqiB and YebT were still present in each flow-through, indicating that both proteins were well expressed. The eluted samples of the purified proteins were pooled and adjusted to a concentration of 0.5 mg/ml. The proteins were sent to Eurogentec to raise polyclonal antibodies in rabbits.

4.2.18. Experimental confirmation that PqiB and YebT localise to the membrane

To confirm the bioinformatic predictions that PqiB and YebT are membrane localised, the soluble (periplasm/cytoplasm) and insoluble (membranes) cellular fractions were isolated. Two litre cultures of the parent strain and Δ*pqiAB* Δ*yebST* mutant were grown to an OD_{600} of 0.6-0.8 and the cells were pelleted and lysed. The cell lysates were centrifuged to pellet the cell membranes, leaving the components of the cytoplasm and periplasm in the supernatant. The pellets were re-suspended and analysed by western blotting alongside the supernatant with anti-PqiB and anti-YebT antibodies. Proteins of the correct molecular weight were observed in the membrane fraction (Figure 4.18). The absence of these bands in the Δ*pqiAB* Δ*yebST* mutant confirmed that the detected proteins were PqiB and YebT. These results confirm that PqiB and YebT are located in the cell membrane, as predicted by TMHMM.

4.2.19. Sarkosyl extraction for the separation of inner and outer membranes

Separate outer and inner membranes were required to identify the specific locations of PqiB and YebT in *E. coli*. Sarkosyl is an ionic detergent that can be used for the solublisation of

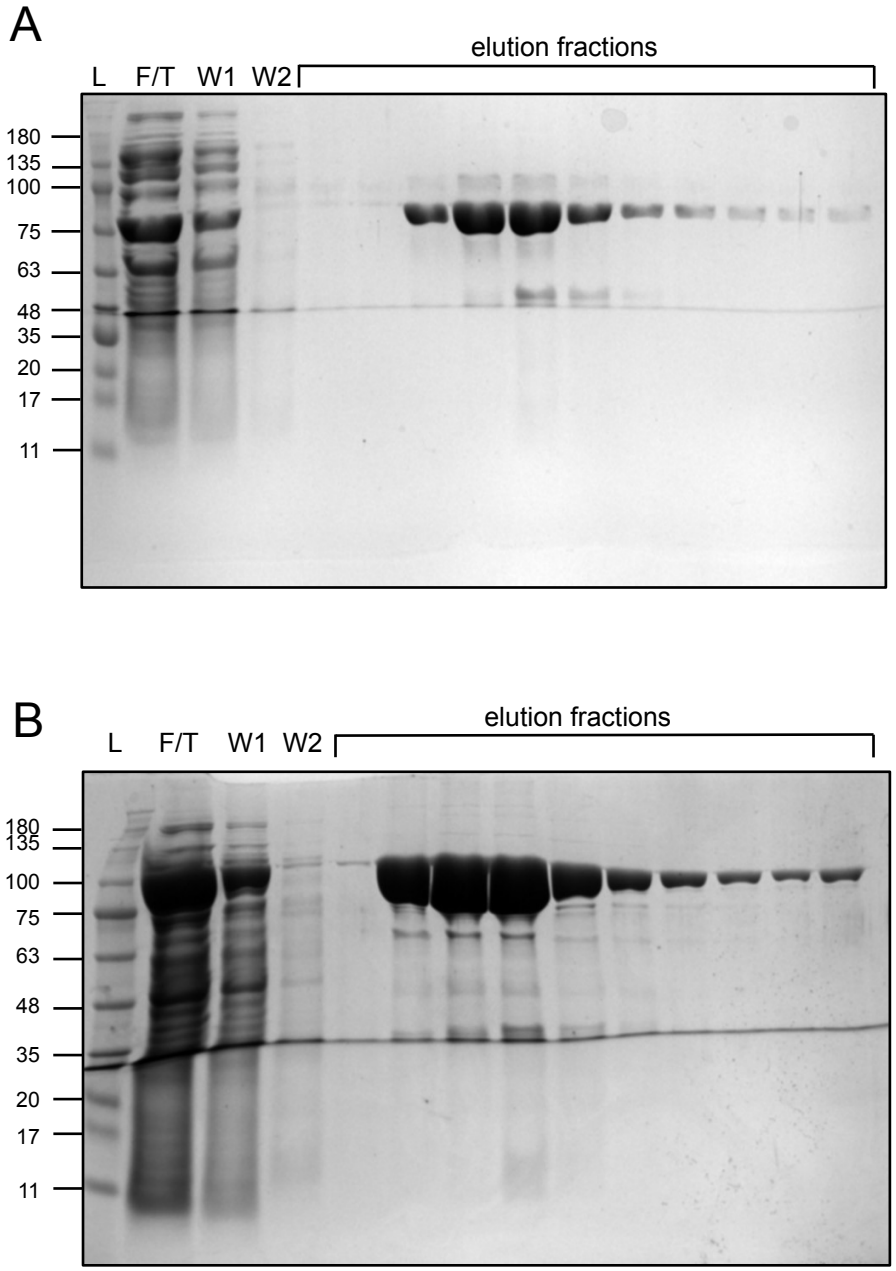


Figure 4.17. SDS-PAGE analysis of the purification of PqiB and YebT for use in the generation of polyclonal antibodies. The proteins were his-tagged and purified using a nickel column. **A.** The purification of PqiB without the N-terminal TM domain. **B.** The purification of YebT without the TM domain. L=molecular weight ladder, F/T=flow through, W1=wash 1 and W2=wash 2.

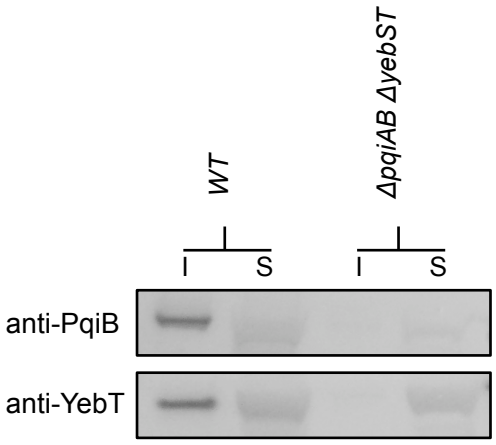


Figure 4.18. Western blot analysis of the soluble (S) and insoluble (I) cell components to confirm the membrane localisation of PqiB and YebT. The fractions of BW25113 and the Δ*pqiAB* Δ*yebST* mutant were analysed using anti-PqiB and anti-YebT antibodies.

the inner membrane (Hobb et al., 2009). To test the effectiveness of this method, the membranes of *E. coli* K-12 BW25113 were isolated as described previously and the inner membrane was solubilised overnight in 1% (w/v) sarkosyl. The insoluble outer membrane was pelleted by ultracentrifugation, leaving the soluble inner membrane in the supernatant. The success of the separation was determined by western blot analysis using antibodies against the known inner and outer membrane proteins AcrB and OmpF. The western blots revealed that the outer membrane fraction did not contain detectable inner membrane proteins but that inner membrane fraction was contaminated with outer membrane proteins (Figure 4.19A). This method is therefore only suitable for the isolation of pure outer membranes, and was not used for the determination of the locations of PqiB and YebT.

4.2.20. Separation of inner and outer membranes by sucrose gradient centrifugation

In an attempt to achieve better separation, the inner and outer membranes were separated by sucrose gradient centrifugation. The cell envelopes of *E. coli* K-12 BW25113 were isolated as described previously (see 4.2.18) and re-suspended in 20% (w/v) sucrose. The membranes made up the top layer of the sucrose gradient and the middle and bottom layers consisted of 60% (w/v) and 70% (w/v) sucrose, respectively. The gradient was ultra-centrifuged for 16 hours to allow membrane separation. Following centrifugation, the cytoplasmic membrane fraction formed a discrete band at the 20-60% (w/v) sucrose boundary. The outer membrane fraction formed a second band at the 60-70% (w/v) boundary. The gradient was collected in ~0.5 ml fractions and analysed by SDS-PAGE and western blotting. SDS-PAGE revealed that the majority of the proteins were found in fractions 7 and 8, which contained the outer membranes, and in fractions 13 and 14, which included the inner membranes (Figure 4.19B). For western blot analysis, the outer membrane marker was changed from OmpF to BamA because BamA had similar expression levels to AcrB (Li et al., 2014). The western blots indicated that the inner membrane

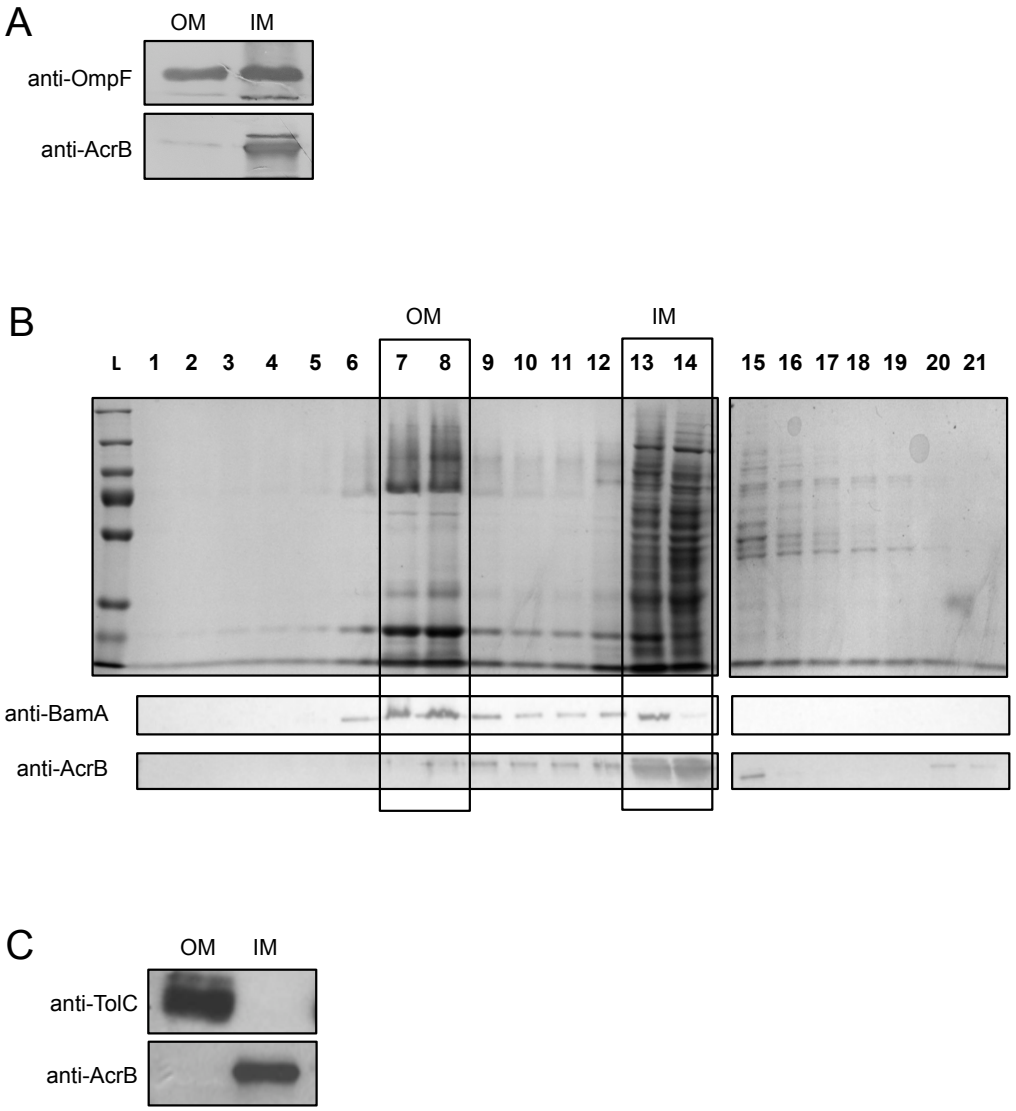


Figure 4.19. Western blot and SDS-PAGE analysis to determine the success of various methods for inner and outer membrane separation in BW25113. OM= outer membrane fraction(s) and IM=inner membrane fraction(s). **A.** Western blot analysis using anti-OmpF and anti-AcrB to assess membrane separation using sarkosyl extraction **B.** Coomassie stained SDS-PAGE gel and western blot (using anti-BamA and anti-AcrB antibodies) of a fractionated sucrose gradient following 16 hours of centrifugation. **C.** Western blot analysis (using anti-TolC and anti-AcrB) of a sucrose gradient following spheroplasting and 40 hours of centrifugation.

fractions were less contaminated with outer membrane proteins than those isolated in the sarkosyl extraction method, but were still contaminated with outer membrane proteins. Furthermore, the outer membrane fractions were contaminated with inner membrane proteins. Therefore, this method was better for the isolation of inner membranes than the sarkosyl extraction method, but worse for the isolation of outer membranes.

4.2.21. Separation of inner and outer membranes by spheroplast formation followed by sucrose gradient centrifugation

An alternative method for preparing samples for sucrose gradient centrifugation was tested in an attempt to decrease cross-membrane contamination. This method incorporated three major changes into the protocol that are published to improve separation (Dalebroux et al., 2015; Osborn and Munson, 1974). First, prior to the isolation of the total cell membranes, the whole cells were converted into spheroplasts by incubating them with lysozyme and EDTA. Secondly, the volume of the middle sucrose layer was increased to increase the separation of the inner and outer membranes in the gradient. Thirdly, the ultracentrifugation step was increased to 40 hours, allowing more time for the membranes to reach equilibrium within the gradient. Following centrifugation, the inner and outer membranes formed very distinct bands that were taken as individual fractions. For western blot analysis the outer membrane marker was changed to TolC to partner its inner membrane counterpart, AcrB. The western blots revealed that the separation of the inner and outer membranes was absolute, with no visual cross-contamination in either membrane fraction (Figure 4.19C). In conclusion, these alterations to the sucrose gradient method resulted in the efficient separation of the inner and outer membranes of *E. coli*.

4.2.22. Sucrose gradient and western blot analysis to determine membrane locations of PqiB and YebT

To determine whether PqiB and YebT are found in the inner or outer membrane, the membranes of BW25113 were separated using the optimised sucrose gradient protocol. As

controls, the membranes of the Δ*pqiAB,* Δ*yebST* and Δ*pqiAB* Δ*yebST* mutants were separated alongside. The fractions were analysed by western blotting using antibodies against TolC, AcrB, PqiB and YebT. The anti-TolC and anti-AcrB western blots confirmed the clean separation of the inner and outer membranes (Figure 4.20). The anti-PqiB and anti-YebT western blots detected proteins of the correct sizes in the BW25113 inner membrane fraction. The proteins were absent in the corresponding deletion strains, confirming that the detected proteins were PqiB and YebT. These results revealed that both PqiB and YebT are inner membrane proteins of *E. coli*.

4.2.23. Thin-layer chromatography to identify phospholipids bound to purified MlaD, PqiB and YebT

Despite the reported role of the Mla pathway in phospholipid trafficking, no study has looked at whether the components of this pathway directly bind phospholipids. To test whether MlaD, PqiB and YebT in *E. coli* bind phospholipids, all 3 proteins were purified without the N-terminal TM domain. PqiB and YebT were purified as described for antibody production and purified MlaD protein was provided by Dr Tim Knowles (School of Biosciences, University of Birmingham). Lipids were extracted by the addition of 3 ml of 2:1 chloroform:methanol to 2 ml of 1 mg/ml protein (or buffer only as a control). After centrifugation, the mixture separated into an upper aqueous phase and a lower organic phase. The organic phases were analysed by one-directional thin layer chromatography (TLC) in a 65:25:4 chloroform:methanol:water solvent system alongside PE, PG and CL standards. Staining with phosphomolybdic acid revealed the presence of two major lipid species bound to MlaD, PqiB and YebT. No lipids were visible in the buffer only (Figure 4.21). When compared to the standards the upper major lipid species migrated to a similar position as the PE standard, and the lower major lipid species to a similar position as the PG and CL standards.

To identify the phospholipid species, the TLC plates were prepared again with and without phosphomolybdic acid staining. The retardation factor values of each lipid species

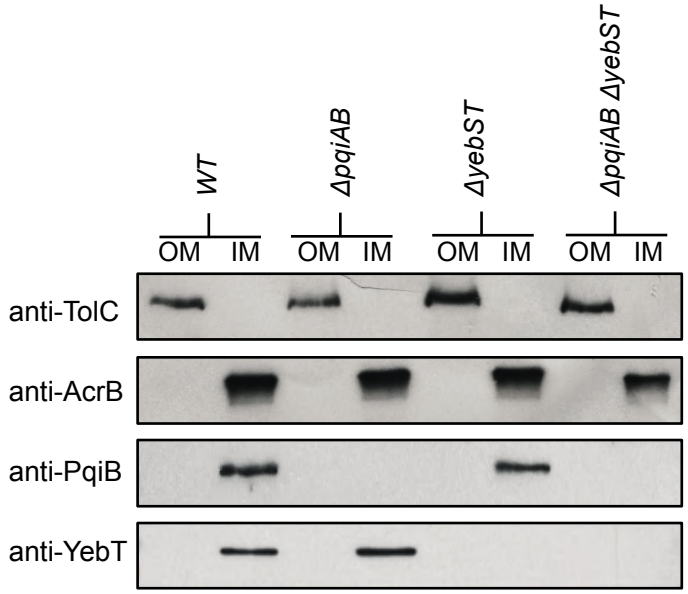


Figure 4.20. Western blot analysis of the outer membrane (OM) and inner membrane (IM) fractions to identify the presence of PqiB and YebT in the inner membrane. The membranes of the BW25113 parent strain and the Δ*pqiAB,* Δ*yebST* and Δ*pqiAB* Δ*yebST* mutants were separated using the optimised spheroplasting and sucrose gradient method. The western blots using anti-TolC and anti-AcrB antibodies confirm the successful separation of the inner and outer membranes. The western blots using anti-PqiB and anti-YebT confirm the presence of PqiB and YebT in the inner membrane fraction, and their absence in the corresponding mutants.

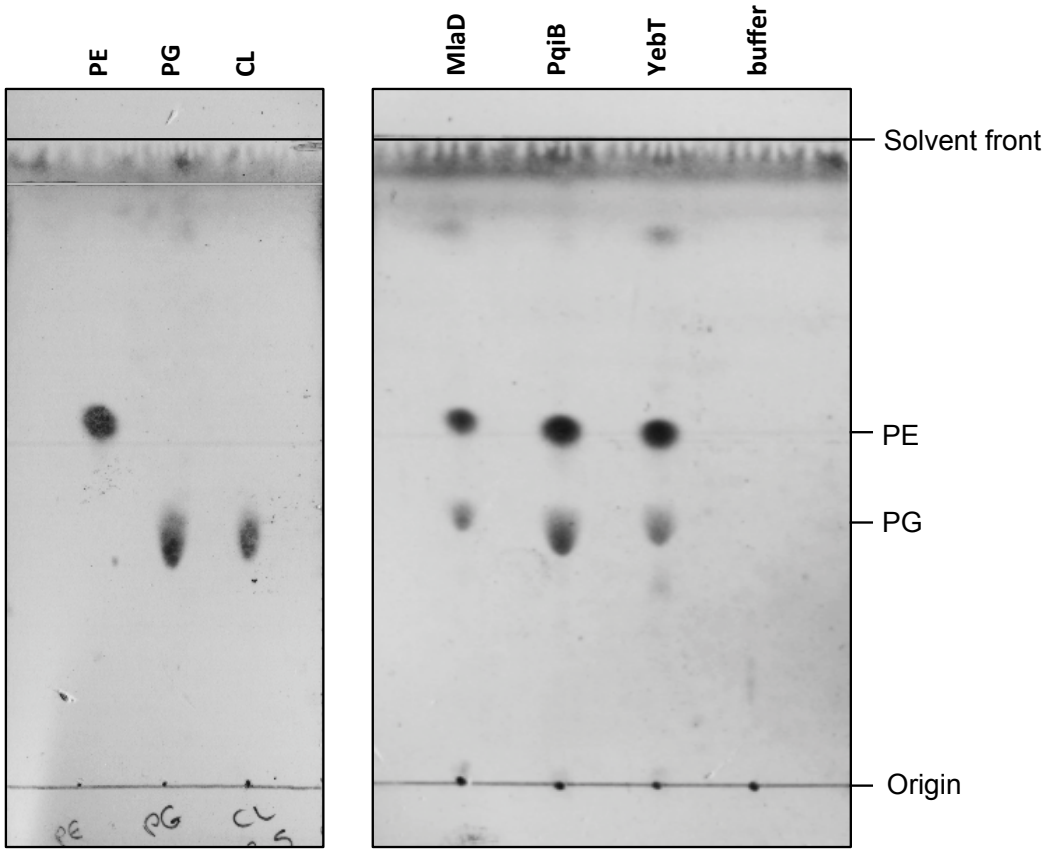


Figure 4.21. Thin layer chromatography of the lipid extracts from purified MlaD, PqiB and YebT (without the transmembrane domains). The lipids were separated in 65:25:4 chloroform:methanol:water alongside PE, PG and CL standards. The lipid species were confirmed by mass spectrometry.

were calculated from the stained TLC plates to predict the locations of the lipid species on the unstained plates. A mass spectrometry was used to determine the masses of all of the major lipid species by electrospray ionisation, directly from the TLC plates. The mass to charge ratios of the major species in the upper spots were 688, 702, 716 and 742, and in the lower spots were 691, 719, 733 and 747. By comparison to the standards, the upper major lipid species in the MlaD, PqiB and YebT lipid extracts were identified as PE and the lower lipid species were identified as PG. This experiment revealed that the soluble domains of MlaD, PqiB and YebT all co-purify with PE and PG.

4.2.24. Two-direction thin layer chromatography to identify changes in membrane phospholipid composition in various MCE mutants

As MlaD is part of a phospholipid trafficking pathway in *E. coli*, the membranes of various MCE mutants were analysed to determine differences in membranes composition relative to the parent strain. The inner and outer membranes of the parent strain and the Δ*mlaD,* Δ*pqiAB*, Δ*yebST,* Δ*pqiAB* Δ*yebST* and the triple deletion strain were prepared using the optimised sucrose gradient protocol. Lipids were extracted from the separated inner and outer membranes using a Bligh-Dyer extraction method (Bligh and Dyer, 1959). The loading of the samples onto TLC plates was normalised by the amount of phospholipid, as determined by an ammonium ferrothiocyanate colourimetric assay (see 2.9.4). The samples were separated by 2 directional TLC. The solvent system for the first direction was 65:25:4 chloroform:methanol:water, which separates out most phospholipids, but not PG and CL. The solvent system for the second direction was 80:12:15:4 chloroform:methanol:acetic acid:water that separates out phospholipid species further, including PG and CL.

The 2D TLC plates were stained with phosphomolybdic acid, revealing the inner and outer membrane profiles of all 6 strains. The relative locations of PE, PG, CL and acyl-PG were known from a previous study, allowing their identification on the TLC plates (Figure 4.22, Dalebroux et al., 2015). The profiles of the inner and outer membranes were very

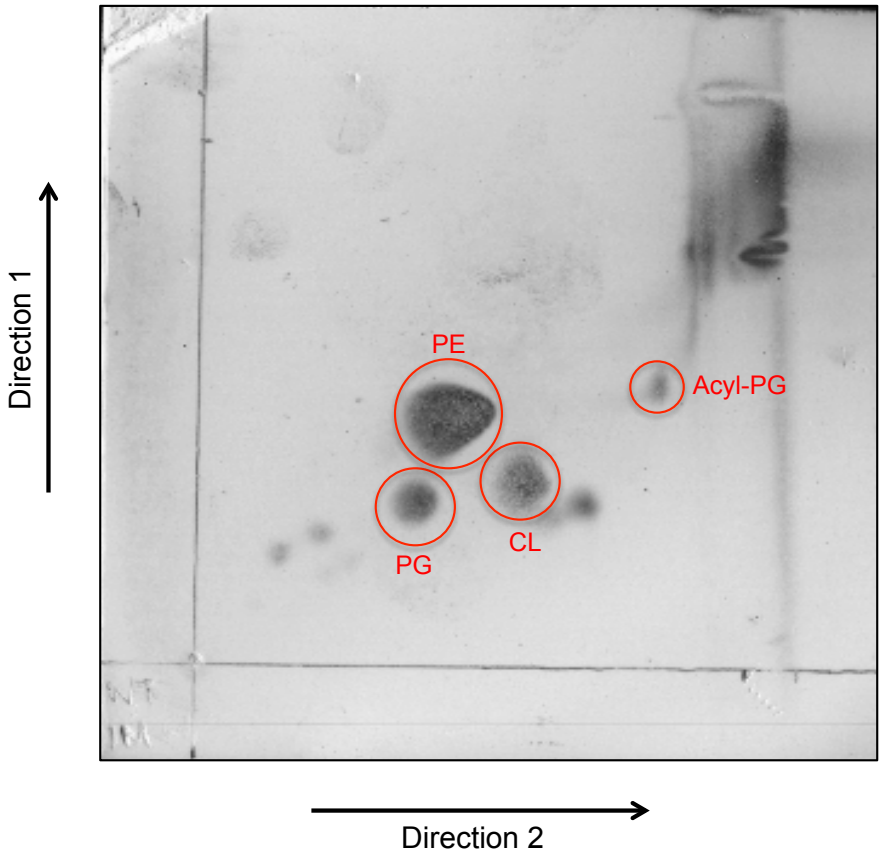


Figure 4.22. Thin layer chromatography of the inner membrane lipid extract of BW25113 with known phospholipid species identified. The lipids were separated in 65:25:4 chloroform:methanol:water (direction 1) and 80:12:15:4 chloroform:methanol: acetic acid:water (direction 2).

similar for all 6 strains (Figure 4.23 and 4.24). Some minor changes were observed in the inner membranes of the Δ*yebST*, Δ*pqiAB* Δ*yebST* and triple deletion strains, but in unknown lipid species (Figure 4.231i-iii). However, there were no major changes in the membrane phospholipid compositions of the various mutants in *E. coli*.

4.3. Discussion

In this chapter, a phenotype associated with defects in all three MCE operons was discovered: sensitivity to the zwitterionic detergent LSB. The similar sensitivities of Δ*pqiAB* and Δ*mlaD* mutants on LSB and additional attenuation of a double mutant suggest that these genes fulfill partially overlapping but nevertheless distinct roles in protection against this detergent. The additive phenotype of the Δ*pqiAB* Δ*yebST* mutant was greater than the additive phenotype of the Δ*mlaD* Δ*yebST* mutant, suggesting the functions of *yebST* and *pqiAB* are more related than the functions of *yebST* and *mlaD*. The triple mutant was unable to grow on 1% (w/v) LSB, demonstrating an increase in sensitivity with increasing loss of MCE domain containing proteins. LSB has been shown to inhibit the carnitine/acyl-carnitine transporter in mitochondria, as a substrate analog (Indiveri et al., 1998). A potential explanation for the observed phenotype is that LSB is also inhibiting a lipid transport pathway in *E. coli*, perhaps one that has some overlap in function with the MCE pathways. Alternatively, MCE domain containing proteins might be involved in trafficking LSB away from its target. Given the related phenotypes for Δ*mlaD*, Δ*pqiAB* and Δ*yebST* mutants it is entirely possible that the roles of the three proteins are similar, albeit with mechanistic differences.

Sensitivity to detergents often signifies an outer membrane defect (Zhou 2012). Therefore, sensitivity to LSB could indicate a loss of outer membrane integrity. Detergents besides LSB were tested to determine the specificity of the phenotype. While no mutants showed sensitivity to SDS alone, the *mlaD* mutants were sensitive to SDS and EDTA, which coincides with previous findings for the Mla pathway (Malinverni and Silhavy, 2009). Therefore, *mlaD* is required for resistance to both LSB and SDS, both which contain a 12-

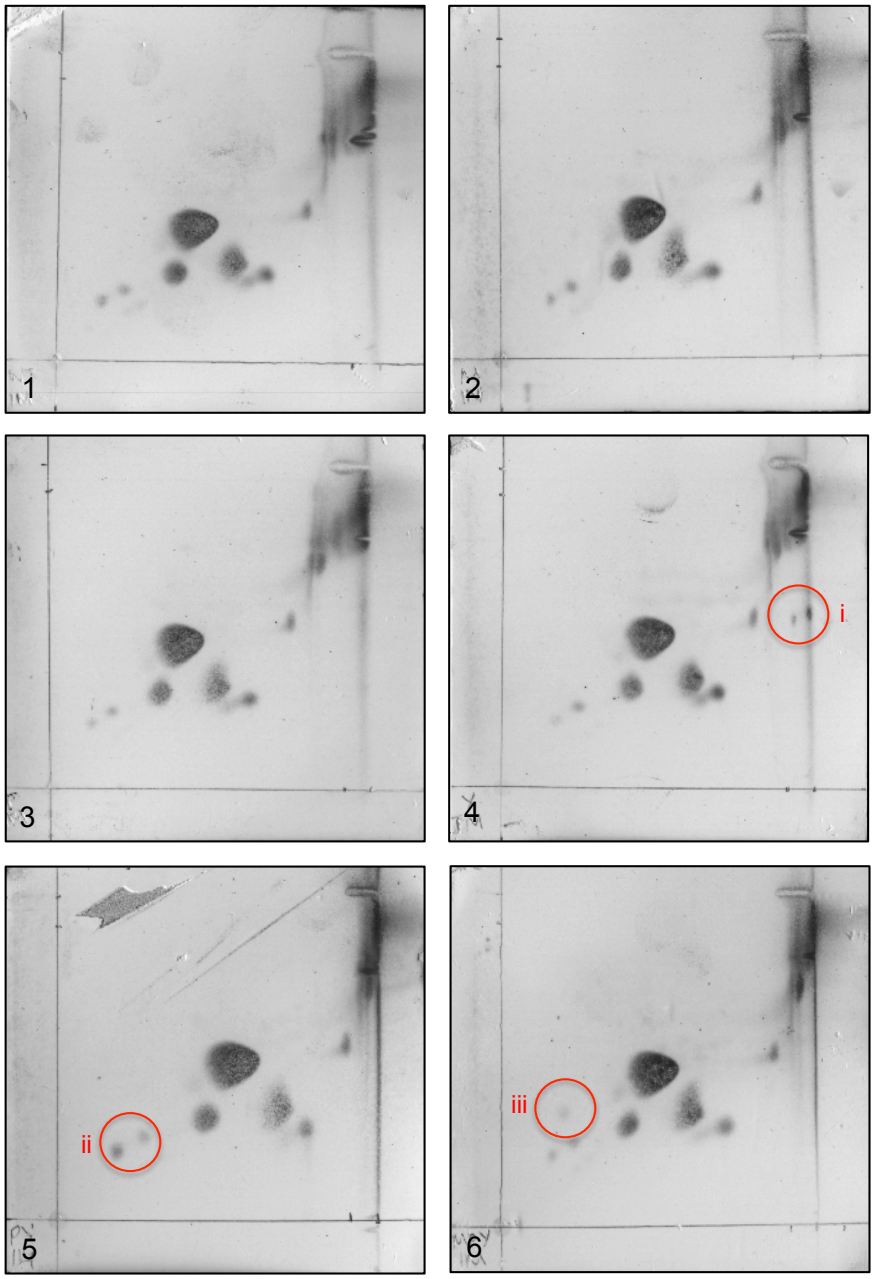


Figure 4.23. Thin layer chromatography of the inner membrane lipid extracts of BW25113 (1), the Δ*mlaD* mutant (2), the Δ*pqiAB* mutant (3), the Δ*yebST* mutant (4), the Δ*pqiAB* Δ*yebST* mutant (5) and the Δ*mlaD* Δ*pqiAB* Δ*yebST* mutant (6). The red circles marked i, ii and iii mark potential changes between membrane profiles. The lipids were separated in 65:25:4 chloroform:methanol:water (direction 1) and 80:12:15:4 chloroform:methanol: acetic acid:water (direction 2).

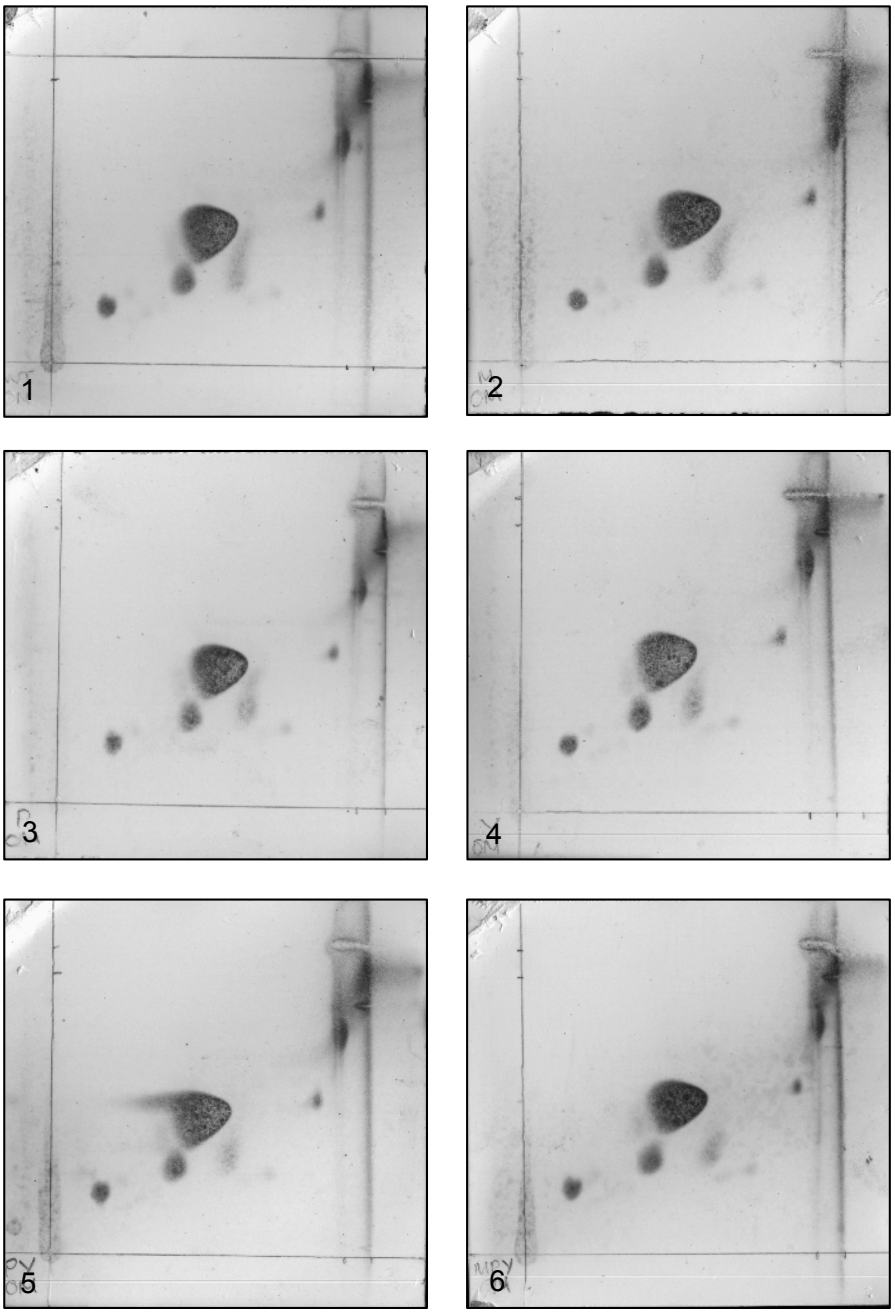


Figure 4.24. Thin layer chromatography of the outer membrane lipid extracts of BW25113 (1), the Δ*mlaD* mutant (2), the Δ*pqiAB* mutant (3), the Δ*yebST* mutant (4), the Δ*pqiAB* Δ*yebST* mutant (5) and the Δ*mlaD* Δ*pqiAB* Δ*yebST* mutant (6). The lipids were separated in 65:25:4 chloroform:methanol:water (direction 1) and 80:12:15:4 chloroform:methanol: acetic acid:water (direction 2).

carbon chain. Screens on other sulfobetaines revealed that a Δ*pqiAB* mutant was equally as sensitive on SB3-10 as on LSB, but with less of an additive phenotype for the Δ*pqiAB* Δ*yebST* mutant. The *pqiAB* gene pair might specifically be required for protection against sulfobetaines (or zwitterionic detergents in general) while *yebST* is only required in the absence of *pqiAB*. The specificity of the phenotypes for each deletion strain could be related to the charge and chain length of the molecules but more detergent screens are required to investigate this hypothesis. Nevertheless, the variation of results on different detergents shows that the functions of *mlaD*, *pqiAB* and *yebST* are not redundant.

Vancomycin is a hydrophobic antibiotic used to treat infections of Gram-positive bacteria. It works by inhibiting the synthesis of peptidoglycan (Hammes and Neuhaus, 1974). Gram-negative bacteria are inherently more resistant to vancomycin due to an additional barrier, the outer membrane, shielding the peptidoglycan. Defects in LPS biogenesis often cause a large increase in sensitivity to hydrophobic antibiotics (Vaara, 1993; Vuorio and Vaara, 1992). Therefore, like detergents, vancomycin sensitivity can be used as a marker for loss of outer membrane integrity. Unexpectedly, some of the deletions strains in this study were much more resistant to vancomycin than the parent strain. This phenotype was observed for all of the *mlaD* mutants, and the Δ*pqiAB* Δ*yebST* mutant. Increased resistance to vancomycin has been linked to accumulation of phosphatidic acid (PA) on the outer leaflet of the outer membrane (Sutterlin et al., 2014). As PA has a large net negative charge, Sutterlin et al. (2014) proposed that vancomycin forms associations with PA in the outer leaflet and is therefore kept away from its target. Mutants defective in the Mla pathway in *E. coli* accumulate phospholipids on the cell surface, which could explain the vancomycin resistance phenotype observed here (Malinverni and Silhavy, 2009). Although it is usually sensitivity to vancomycin that is seen as a marker for loss of membrane integrity, this discussion suggests vancomycin resistance could also be used as a marker for loss of membrane homeostasis.

To test whether the phenotype observed for the *mlaD* mutants were applicable to the whole *mla* operon, *mlaC* mutants were also screened for increased sensitivity to LSB.

The phenotypes of the *mlaC* mutants were exactly those of the *mlaD* mutants on LSB, suggesting that all the individual components of the *mlaFEDCB* operon might be required for resistance to LSB. MlaC is the periplasmic chaperone of the Mla pathway, and there is no homolog in the *pqiAB* and *yebST* operons. If MlaC is the sole chaperone for all MCE domain containing proteins in *E. coli*, an *mlaC* gene deletion would disrupt all 3 MCE pathways and the phenotypes of the *mlaC::aph* and Δ*mlaD* Δ*pqiAB* Δ*yebST* mutants on LSB would be similar. As the phenotype of the Δ*mlaC* mutant was similar to phenotype of the Δ*mlaD* mutant and not the Δ*mlaD* Δ*pqiAB* Δ*yebST* triple mutant, it is unlikely that MlaC is chaperoning lipids for all 3 MCE pathways (or at least it is not the only chaperone).

Similar to the Δ*pqiAB* mutant, the individual gene deletion strains (in *pqiA::aph*, *pqiB::aph* and *ymbA::aph*) were all sensitive to LSB, suggesting that the three genes are functionally connected. The LSB phenotype of the Δ*pqiAB* mutant was also manipulated to identify important components of PqiB. If either the C-terminal alpha helix or MCE domain was removed from PqiB, the protein could not protect against LSB, suggesting these components of PqiB are important for function. Furthermore, a "PqiAB-like" YebST (with the first 3 YebT MCE domains and an added PqiB C terminal alpha helix), could not complement the Δ*pqiAB* mutant, despite their structural similarity. This result indicates that that are more intricate differences between PqiB and YebT than the number of MCE domains and alpha helices, and that these differences are important for function. In the future, the presence of each truncated PqiB and YebT protein will be confirmed by western blot to confirm that these proteins are expressed and stable. This investigation could be taken further to identify other important components of PqiB for function, and to determine whether it is differences between PqiA and YebS, PqiB and YebT, or both that are responsible for these results.

The location of a protein is an important factor in understanding its function. Therefore, one aim of this chapter was to determine the subcellular locations of PqiB and YebT in *E. coli*. Bioinformatic predictions revealed that both proteins localise to the membrane, which was then confirmed experimentally. After the optimisation of the

separation of inner and outer membranes in *E. coli*, PqiB and YebT were identified as inner membrane proteins. This is the first study to identify the location of natively expressed multi-MCE domain proteins in Proteobacteria. This result differs from the previous work on the *Vibrio parahaemolyticus* YebT homolog, MAM-7, which was reported to be surface localised when over-expressed (Krachler et al., 2011). However it is in agreement with all other published literature on the location of MCE domain containing proteins in Gram-negative bacteria, including MlaD, and the chloroplast protein, Tgd2 (Awai et al., 2006; Hong et al., 1998; Monaco et al., 2006). This result is consistent with the fact that MCE domain containing proteins generally have no predicted signal peptide cleavage sites (Petersen et al., 2011). Bioinformatic predictions have led to the conclusion that PqiB and YebT, like MlaD, face the periplasm in *E. coli*. In the future this prediction will be confirmed by immunofluorescence using the anti-PqiB and anti-YebT antibodies generated in this study.

Another aim of this chapter was determine whether MlaD, PqiB and YebT bind lipids, specifically phospholipids. To explore this, lipid extracts of purified MlaD, PqiB and YebT were analysed by TLC and mass spectrometry to reveal the presence of PE and PG. These results are supported by the recently published study on the Mla pathway, which identified PE and PG bound to purified MlaD (Thong et al., 2016). Furthermore, MAM-7 and the protein from *Arabidopsis thaliana*, Tgd2, have both been found to bind phosphatidic acid (Awai et al., 2006; Krachler et al., 2011). It is plausible that these proteins have a general affinity towards phospholipids (or lipids in general) and that the results observed for MlaD, PqiB and YebT are a consequence of the *E. coli* lipid composition. PE and PG are the most abundant phospholipids, making up 80% and 15% of the *E. coli* cell membrane, respectively (Ingram, 1977). In the future, fluorescence-based experiments can be used to determine the lipid binding affinities of MCE domain containing proteins and the kinetics of a particular protein-lipid interaction. This technique could also be used for competition assays to determine whether the proteins have preferences towards particular phospholipids.

With the prior knowledge of the Mla pathway and discovery that all three MCE domain containing proteins bind phospholipids, it was appropriate to determine whether various MCE mutants have altered membrane phospholipid compositions and thus determine whether these proteins have a role in trafficking membrane phospholipids. Two-dimensional TLC of lipid extracts revealed minor changes for the inner membrane samples. Although these might be genuine, they might also be artefacts introduced through sample preparation, loading or staining. Either way, no major changes were observed in the phospholipid compositions of the inner and outer membranes, suggesting that the MCE domain containing proteins do not have a major role in the maintenance of the membrane phospholipid composition. Although the Mla pathway is involved in the removal of unwanted phospholipids from the cell surface, only minor phospholipid accumulation was observed in its absence (Malinverni and Silhavy, 2009). The techniques used here were probably not sensitive enough to detect such minor changes, so a minor role in the phospholipid composition of the cell membranes cannot be ruled out. In the future, mass spectrometry will be used as a more quantitative and sensitive measure for the determination of membrane composition and will also allow chain length determination. Furthermore, no work has been done to check for an overall increase or decrease in phospholipid content of the cell membranes, by comparing the ratios of phospholipid to membrane proteins and/or LPS in the outer membrane. Future work will also focus on repeating the phenotype of phospholipid accumulation on the surface of *mlaD* mutants and determining whether this phenomenon occurs in other MCE deletion strains.

Overall the results in this chapter represent a step towards understanding the function of the three MCE domain containing proteins in *E. coli*, and their relationship to one another. The phenotypes identified in this chapter suggest that there is some overlap in function between the three pathways (particularly between PqiAB and YebST), but that they do not overlap completely. All three MCE domain containing proteins in *E. coli* are located in the inner membrane and have a general affinity towards phospholipids. Although the results did not ultimately determine one way or another whether the proteins are involved

in phospholipid trafficking, they rule out a major role in the trafficking phospholipids to or from membranes. Nevertheless, based on the existing knowledge of the Mla pathway, specific detergent sensitivities and vancomycin resistance it seems feasible that there is a minor function in the maintenance of outer membrane phospholipid content. It is also possible that MCE domain containing proteins have a role in trafficking phospholipids across membranes, perhaps for uptake from the external environment. A role in uptake has been demonstrated for MCE pathways in other bacteria, as shown by the inability of mutants to grow on lipids as a carbon source (Endo et al., 2007; Mohn et al., 2008; Pandey and Sassetti, 2008). Unfortunately the *E. coli* strain used in this study could not grow on phospholipids and therefore this theory could not easily be tested. In the next chapter, the focus is moved towards determining the first crystal structure of an MCE domain containing protein to provide more information on the lipid binding properties and general function of MCE domains.

CHAPTER 5

PURIFICATION AND CRYSTALLOGRAPHY OF PQIB

5.1. Introduction

To date no structure has been determined for any MCE domain containing protein. X-ray crystallography has been a key technique in the field of structural biology since the determination of the structure of myoglobin in the 1950s (Kendrew 1958). This technique involves screening the protein of interest in a variety of conditions with the aim of growing well-ordered crystals. A well-formed crystal can diffract X-ray waves that are detected and visualised as a pattern of spots, from which information about the structure of the protein can be mathematically determined. The diffraction pattern provides information about the direction from which a wave came (the spot position) and the amplification of a wave (the spot intensity). Information is lost about the phase of the waves, resulting in what is termed the "phase problem". Without this information, the crystal structure cannot be fully solved. For some proteins, solving the "phase problem" is made easier by utilising information from existing structures of similar proteins. Where similar structures are unavailable, multiple isomorphous replacement can be used. This method requires datasets from two crystals: one native (picked directly from the conditions the crystal formed in) and one soaked in heavy atoms. The heavy atoms attach to two or more positions in the protein and datasets from both crystals are compared to reveal the heavy atom positions (Smyth and Martin, 2000). From the data, the phase of the heavy atoms can be determined and the phases of the native crystal can then be solved mathematically (Smyth and Martin, 2000). If the crystal(s) diffract to a high enough resolution, then the data can provide information on protein structure.

As discussed previously, the basic secondary structure for an MCE domain containing protein is predicted to be an N-terminal alpha helix in the TM domain and 6 conserved beta sheets in the MCE domain. In more complex MCE domain containing proteins there are large C-terminal alpha helices and/or multiple MCE domains. Both of these characteristics are found in PqiB, in which there are 3 MCE domains and a C-terminal alpha helix of approximately 120 amino acids in length (Buchan et al., 2013). A crystal structure of PqiB would therefore provide information on all of the common features of

MCE domain containing proteins. The aim of this chapter was to determine the structure of PqiB using crystallography.

5.2. Results

5.2.1. Purification of full-length PqiB for crystallography

To obtain the first crystal structure of an MCE domain containing protein, full-length PqiB was expressed and purified for crystallography. The *pqiB* gene sequence was amplified by PCR using pqiB_NdeI_F and pqiB_XhoI_R and ligated into pET22b(+) using NdeI and XhoI restriction sites, and transformed into *E. coli* BL21 DE3. To test for protein production, 10 ml cultures were induced at an OD_{600} of 0.6 with 1 mM IPTG and left overnight at 18°C. The low induction temperature was used to slow down production and improve folding and incorporation of PqiB into the membrane. As this trial experiment resulted in the accumulation of sufficient PqiB the cultures were scaled up to six 1 l cultures. After overnight induction, the membranes were isolated from lysed cells and re-suspended in 0.5% of the detergent *n*-dodecyl-β-D-maltoside (DDM) for overnight solubilisation. The protein was purified as described previously (see 4.2.18) and 0.1% DDM was used throughout to maintain protein solubility. The flow-through, washes and protein elution fractions were analysed by SDS-PAGE (Figure 5.1). Substantial amounts of PqiB were eluted using imidazole, which also contained visible contamination of other proteins or PqiB breakdown products. Therefore, further purification was required to prevent crystallisation of contaminants.

5.2.2. Gel filtration chromatography of purified full-length PqiB

The partially purified PqiB sample was further purified by size exclusion chromatography to remove the unwanted contaminants. This involved passing the sample through a 120 ml Superdex 200 10/300 GL column gel filtration column using the ÄKTA Pure chromatography system, where the rate of travel was proportional to the size of the protein. The eluate from the Superdex column was collected in 2 ml fractions and the amount of

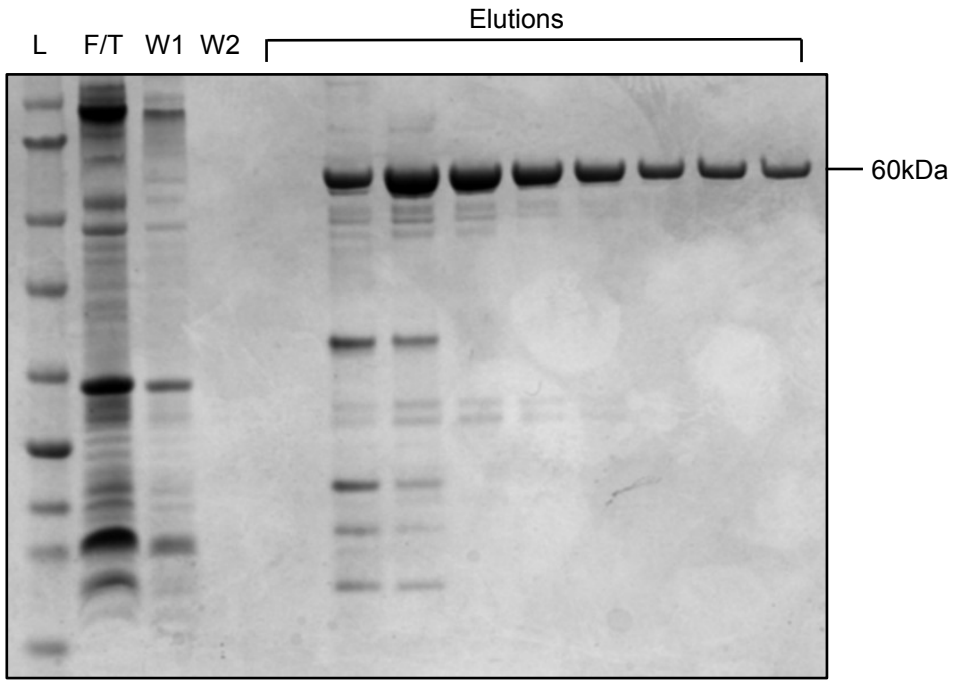


Figure 5.1. SDS-PAGE analysis of the purification of full length PqiB for crystallography. The full-length his-tagged protein was solubilised in DDM and purified using a nickel column. L=molecular weight ladder, F/T=flow through, W1=wash 1 and W2=wash 2.

protein in each fraction was detected by measuring its absorbance at 280 nm, which revealed that PqiB had eluted between 42 and 58 ml (Figure 5.2A). Analysis of alternate fractions by SDS-PAGE revealed much purer PqiB protein than prior to gel filtration (Figure 5.2B). The fractions were pooled and concentrated to ~20 mg/ml for crystallography.

5.2.3. Crystallography of full-length PqiB

As PqiB has an TM region and a large soluble domain, the purified protein was screened in conditions designed for both membrane proteins and soluble proteins. These included the JCSG+, ProPlex, Morpheus HT-96, Midas HT-96, PACT premier and PEGRx screens, each consisting of 96 conditions. The crystallography trials were set up as 600 nl sitting drops in 96-well crystallisation plates with a 1:1 protein to condition ratio. Protein precipitation was observed immediately in most conditions, suggesting either that the protein concentration was too high, or that the protein solubility was too low. Therefore, the same screens were set up with a decreased initial protein concentration of 10 mg/ml, which resulted in slightly less precipitation. After 1 month, no crystals were observed. The focus of the crystallography was moved to the soluble region of PqiB with the aim of improving protein solubility and crystal formation.

5.2.4. Purification and gel filtration of the soluble component of PqiB for crystallography

Expression PqiB without the TM domain was achieved using *E. coli* BL21 DE3 transformed with pET22b-PqiBΔTM. One litre of cells were grown as described previously (see 4.2.17), but with overnight induction at 18°C. The expressed protein was purified using a nickel column and the flow-through, washes and elution fractions were analysed by SDS-PAGE. The yield of PqiB had improved substantially compared to the procedure used previously for the production of antibodies (Figure 5.3). The elution fractions were pooled and further purified by gel filtration chromatography. The 280 nm absorbance trace revealed a double peak, indicating the presence of different protein species or conformations (Figure 5.4A).

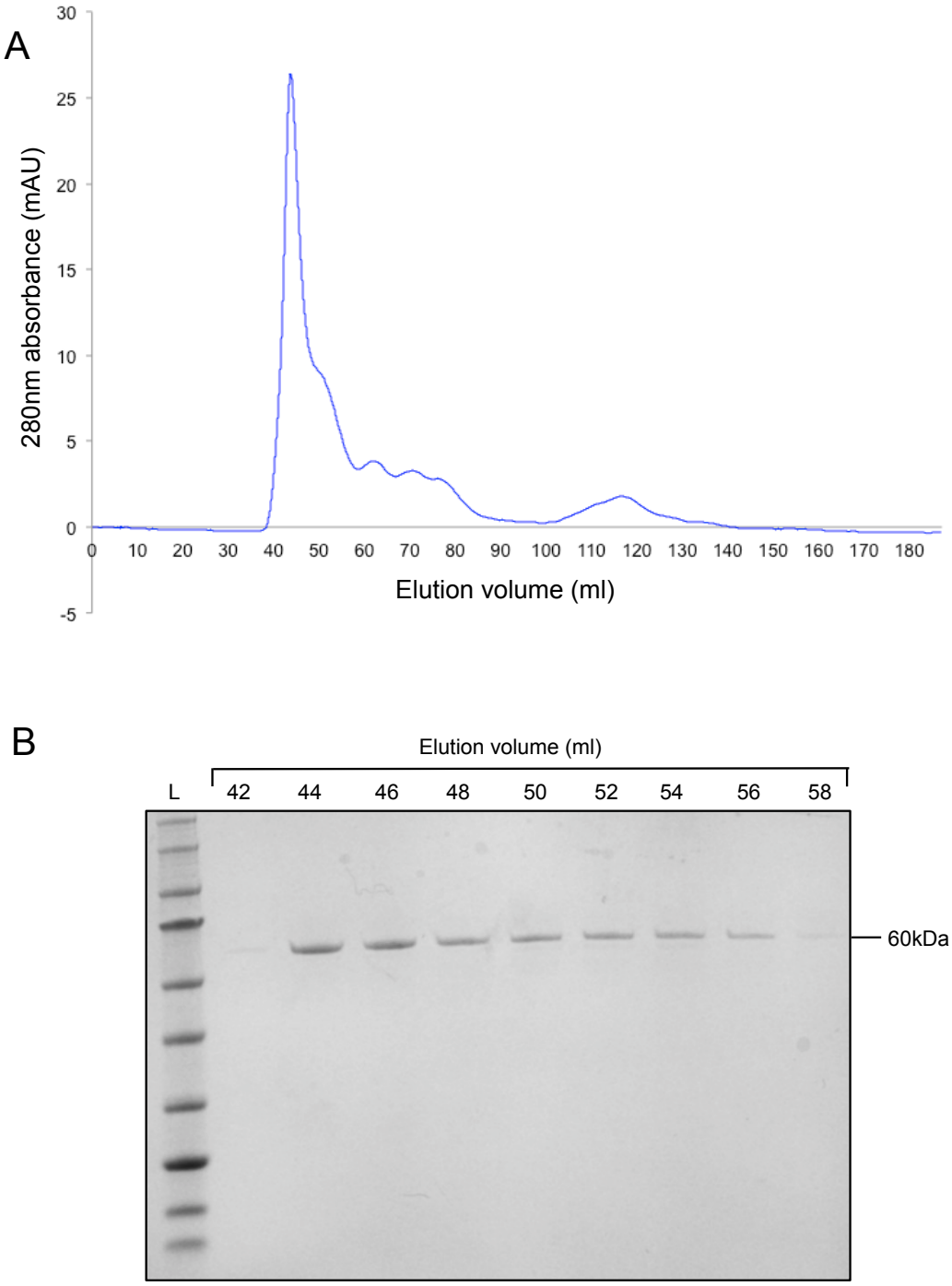


Figure 5.2. Gel filtration of full length PqiB for crystallography using the Superdex 200 10/300 GL column **A.** The 280 nm absorbance peak revealing the elution volume of PqiB. **B.** SDS-PAGE analysis of the alternate PqiB elution fractions. L=molecular weight ladder.

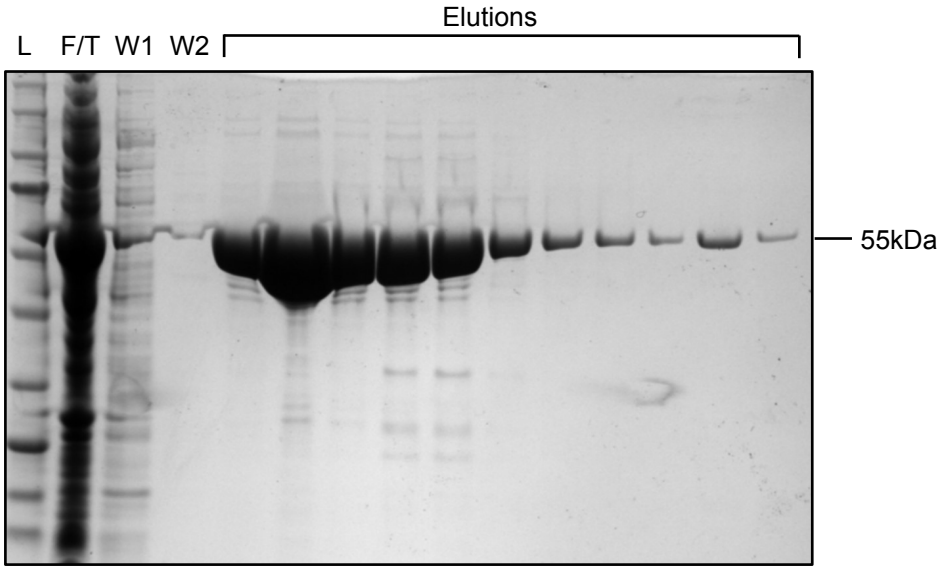


Figure 5.3. SDS-PAGE analysis of the optimised purification of PqiB without the TM region for crystallography. The proteins were his-tagged and purified using a nickel column. L=molecular weight ladder, F/T=flow through, W1=wash 1 and W2=wash 2.

Fractions spanning the entire peak were analysed by SDS-PAGE, which revealed very pure PqiB protein in all fractions (Figure 5.4B). The early eluting PqiB was potentially aggregated and was therefore excluded from further experimentation. The later eluting PqiB was pooled and concentrated to ~20 mg/ml for crystallography.

5.2.5. Initial crystallography of PqiB without the transmembrane region

Purified PqiB without the transmembrane region was placed in the same crystal screens as used previously for the full-length protein. There was less initial protein precipitation than in the previous screens, suggesting that removal of the TM domain had substantially improved protein solubility. After 2.5 weeks the first PqiB crystals were observed in the PEGRx screen in 0.1 M MES, pH 6 and 28% (w/v) polyethylene glycol (PEG) 4000 (Figure 5.5i). The crystals were picked and stored in the condition it formed, or soaked in bromine for phase determination. The crystals were stored in liquid nitrogen and transported to Diamond Light Source for X-ray analysis. The best crystal diffracted to a resolution of 12 Å, which confirmed the crystals were protein but could not provide any structural information.

In an attempt to improve crystal growth the drop size was scaled up to 4 μl with a range of pH (5.5 to 6.5) and PEG 4000 concentrations (20 to 32% (w/v)). After 2.5 weeks new crystals had formed within the range of pH 5.5 to 6.0 and 24 to 28% (w/v) PEG 4000 (Figure 5.5ii). These crystals only diffracted to a resolution of 10 Å and therefore could not be used to obtain structural information.

5.2.6. Microseed matrix screening to improve crystal growth

Different conditions are often required for crystal nucleation and crystal growth, which decreases the chances of crystal formation in a given condition. Microseed matrix screening overcomes this problem by breaking up crystals and using them as a nucleation point in new screens (Ireton and Stoddard, 2004). In an attempt to improve crystal growth, the poorly diffracting PqiB crystals from the previous section were used for microseed matrix screening. The crystals were removed from their original drops and were broken by

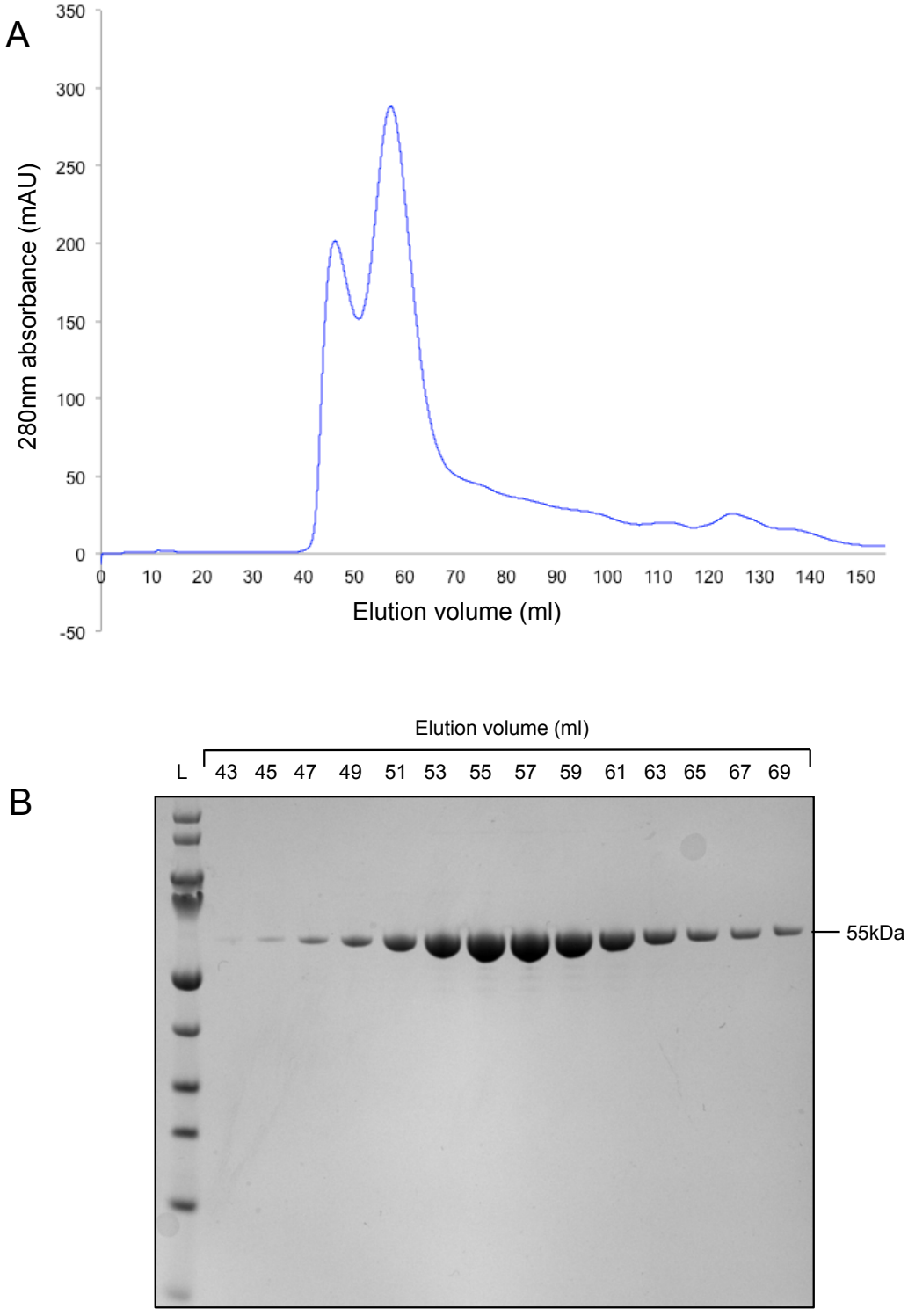


Figure 5.4. Gel filtration of PqiB without the TM region for crystallography using the Superdex 200 10/300 GL column. **A.** The 280 nm absorbance peak revealing the elution volume of PqiB. **B.** SDS-PAGE analysis of the PqiB elutions. L=molecular weight ladder.

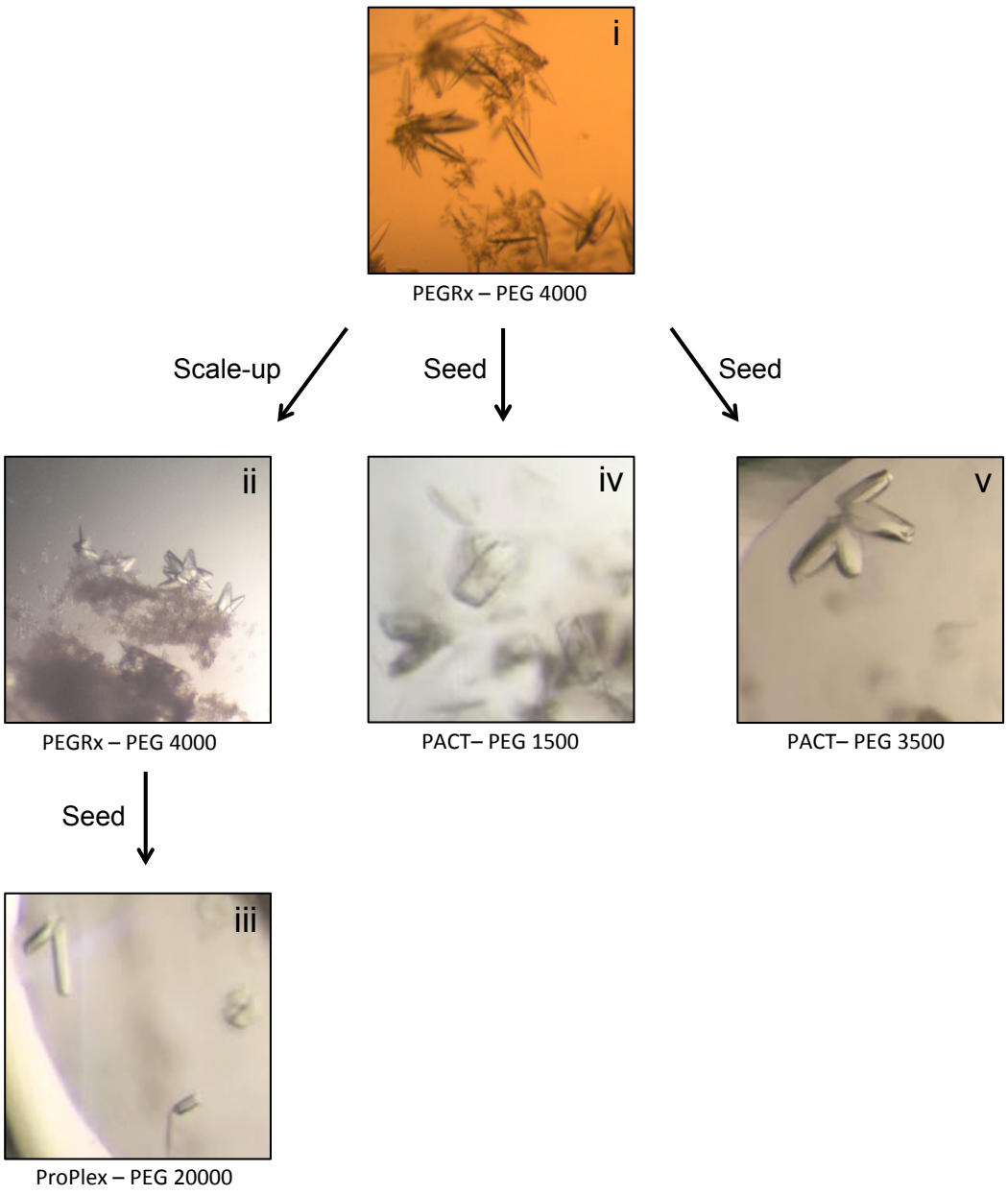


Figure 5.5. The workflow of the crystallography of PqiB without the TM domain. Initially, protein at a concentration of 20 mg/ml was placed in a variety of crystallography trials in 200 nl drops. The first crystals were observed in the PEGRx screen in PEG 4000 based condition. Crystal growth was optimised by scale-up of the condition to a 2: 2 µl drop (Scale-up) or microseed matrix screening into a new condition (Seed).

vortexing. The crystal stock was used for seeding new 600 nl screens with a 1:2:3 ratio of stock : condition : protein. The crystals that diffracted to 10 Å were used for seeding into the ProPlex, PACT premier and Midas HT-96 screens. After 1 week new crystals had formed in the ProPlex screen in 0.1 M Na HEPES, pH 7 and 15% (w/v) PEG 20000 (Figure 5.5iii). However, these crystals did not diffract.

The crystals that diffracted to 12 Å were used for seeding into the PACT screen. After 1 week crystals had formed in 0.1 M MIB buffer (sodium malonate dibasic monohydrate, imidazole, boric acid) pH 8 and 25% (w/v) PEG 1500 (Figure 5.5iv). These crystals diffracted to a resolution of 12 Å. After another 5 days, crystals had formed in 0.02 M sodium/potassium phosphate and 20% (w/v) PEG 3350 (Figure 5.5v). These crystals diffracted to a much-improved resolution of 4 Å, but this resolution was achieved only for the bromine-soaked crystal and not for those obtained from the native condition, and therefore the structure could not be solved.

Overall, the crystallography process identified that PqiB forms crystals in conditions with low Mr PEG, and that the best crystals formed in PEG 3350 following microseed matrix seeding. This information can be used to further optimise PqiB crystallography.

5.2.7. Determining the oligomeric state of PqiB using gel filtration

Gel filtration size exclusion was used to determine the oligomeric state of purified PqiB. The partition coefficients (K_{av}) of proteins with known molecular weights (Table 5.1) were obtained by gel filtration on a Superdex 200 PG 16/60 column. The partition coefficient was calculated as follows: $K_{av} = (V_e - V_o)/(V_t - V_o)$, where V_e is the elution volume of the protein, V_o is the void volume of the column and V_t is the total column volume. The total column volume for the Superdex 200 PG 16/60 column is 120 ml. The void volume of the column was determined using blue dextran, a dye with a molecular weight of 2000 kDa that is too large to enter the gel matrix. Blue dextran eluted at 45.19 ml. The partition coefficients and log molecular weights of the protein standards (Table 5.1) were plotted to

Standard	Mr	Log Mr	K_{av}
Ferritin	474000	5.68	0.169
Aldolase	156800	5.2	0.311
ConAlbumin	76000	4.88	0.426
Albumin	66500	4.82	0.303
Ovalbumin	45000	4.65	0.49
Carbonic anhydrase	29000	4.46	0.581
Cytochrome C	12000	4.08	0.677
Aprotinin	6511	3.81	0.787

Table 5.1. The standards used for generation of a standard curve for gel filtration. The partition coefficients wer calculated for each protein based on the following equation: $K_{av} = (V_e - V_o)/(V_t - V_o)$, where V_e is the elution volume of the protein, V_o is the void volume of the column and V_t is the total column volume.

form a standard curve from which unknown molecular weights could be calculated (Figure 5.6).

The peak elution volume for full length PqiB was 44 ml. This was in the void volume of the column and therefore the molecular weight could not be determined. This also suggested that the purified full length PqiB was aggregated. The peak elution volume for PqiB without the TM domains was 58 ml. This equated to a molecular weight of 375400 Da, as calculated from the standard curve (Figure 5.6). This molecular weight was 6.72 times the molecular weight of PqiB without the TM domain monomer (55863 Da), suggesting that the purified PqiB was either a hexamer or heptamer.

5.3. Discussion

The aim of this chapter was to determine the structure of PqiB. While this was not achieved, crystals that diffracted to 4 Å were obtained. PqiB production and crystallisation were optimised and it was revealed that PqiB forms reasonable crystals in low molecular weight PEG - based conditions. This goes a significant way towards identifying appropriate conditions for crystallisation of MCE domain containing proteins, laying the foundations for the future determination of crystal structures. Other techniques could also be adopted such as NMR (on a single MCE domain or shorter protein), small-angle X-ray scattering and cryo-electron microscopy.

During the crystallography process, data obtained from gel filtration was used to determine the oligomeric state of PqiB. Full-length PqiB eluted in the void volume, suggesting aggregation had occurred. In hindsight, this could be responsible for the lack of crystals obtained for full length PqiB. The molecular weight of the truncated PqiB, without the TM domain, was 6.72 times that of a equivalent PqiB monomer, suggesting that the protein forms either a hexamer or heptamer. Although this could be an artefact of protein overproduction, other MCE domain containing proteins have been reported to form homoligomers. In *A. thaliana* the complex involved in trafficking phosphatidic acid is reported to contain 8 to 12 copies of the MCE domain containing protein Tgd2 (Roston et

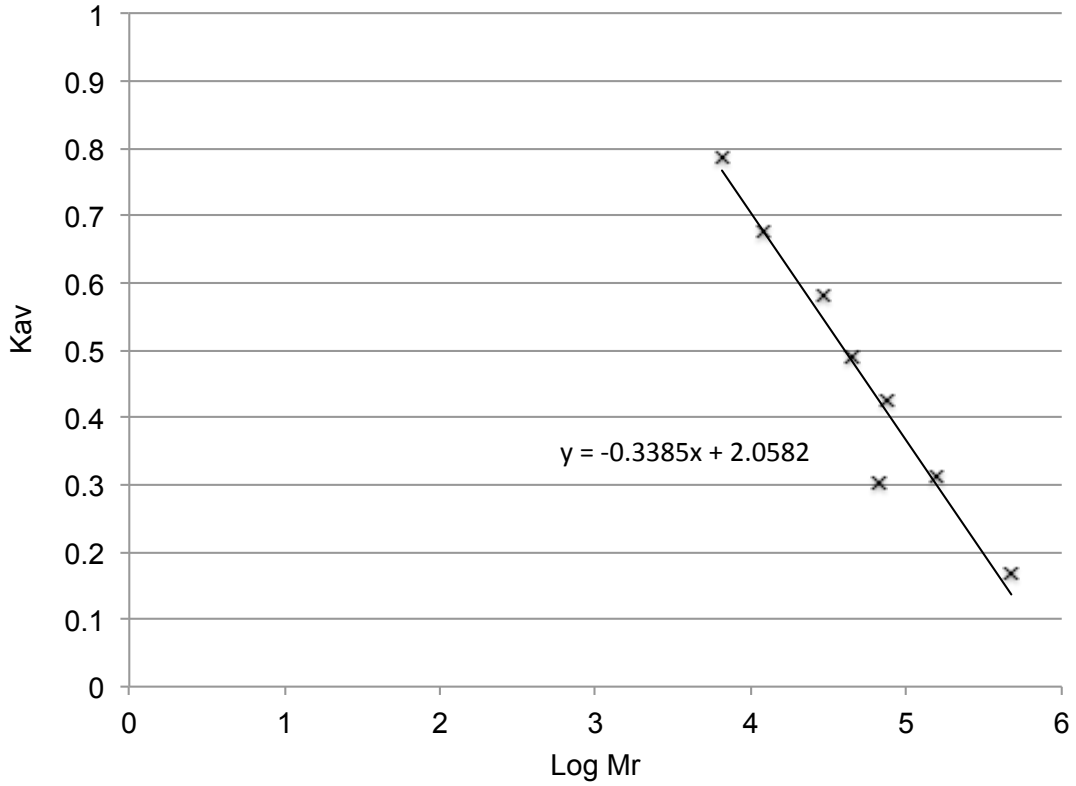


Figure 5.6. A standard curve obtained from gel filtration using a Superdex 200 PG 16/60 column. The log molecular weight (log Mr) of each protein was plotted against the partition coefficient (K_{av}). $K_{av} = (V_e - V_o)/(V_t - V_o)$, where V_e is the elution volume of the protein, V_o is the void volume of the column and V_t is the total column volume.

al., 2012). Furthermore, MlaD from *E. coli* has recently been published as forming a hexamer (Thong et al., 2016). With this in mind, it is likely that the results observed for PqiB are genuine and that it probably forms a hexamer. How or why MCE domain containing proteins form homooligomers will become clear with future functional and structural studies. In the next chapter, the focus is moved to determining the importance of MCE domain containing proteins during pathogenesis using the *Salmonella* murine infection model.

CHAPTER 6

THE ROLE OF MAMMALIAN CELL ENTRY DOMAINS IN *SALMONELLA* TYPHIMURIUM PATHOGENESIS

6.1. Introduction

The *Salmonella* Typhimurium murine infection model is commonly used to investigate the role of proteins in pathogenesis. Whilst it is primarily a pathogen of the gastrointestinal tract, systemic infection results in the spread of bacteria to reticuloendothelial system organs such as the liver and spleen. The importance of MCE domains during murine infection has already been demonstrated in *Mycobacterium tuberculosis,* where *mce* mutants are usually found to be attenuated (Marjanovic et al., 2010; Senaratne et al., 2008). The infection phenotypes of mutants lacking MCE domain containing proteins in Gram-negative bacteria are less well characterised. Infection studies in Gram-negative bacteria have only focussed on homologues of *yebT*. For example, Mahmoud *et al.* (2016) showed that a *Shigella sonnei* strain lacking multivalent adhesion molecule SS01327 was attenuated in *Galleria mellonella*. Furthermore, mutants of *yebT* homologues in a range of Enterobacteriaceae have been shown to exhibit decreased cell attachment and invasion of mammalian cells (Krachler et al., 2011; Mahmoud et al., 2016).

Studies on this topic generally focus on insects or nematodes (Krachler et al., 2011; Mahmoud et al., 2016): the role of MCE domains during infection of mammals has never been studied for a Gram-negative pathogen. Furthermore, nothing is currently known about the importance of MlaD and PqiB during infection. In this chapter, experiments are described in which *Salmonella* Typhimurium was used to determine whether MCE domain containing proteins in Proteobacteria have a role in mammalian pathogenesis. The first aim was to determine, through *in vitro* experiments, whether MlaD, PqiAB and YebST play a role in initial infection. The second aim was to investigate a potential role for MCE domain containing proteins in other stages of infection, through *in vivo* experiments. Given that *S.* Typhimurium is able to grow using phospholipids as a sole carbon source (Antunes et al., 2011), the third and final aim was to determine whether MCE domain containing proteins are important for uptake/growth on phospholipids.

6.2. Results

6.2.1. Construction of pqiAB *deletions in* Salmonella *Typhimurium SL1344*

The *pqiAB* genes span positions 1110249 to 1113147 on the *S.* Typhimurium SL1344 genome. Deletions were constructed using gene doctoring, a method that involves λ Red recombination from a plasmid (Lee et al., 2009). First, the kanamycin resistance cassette was generated by PCR using the plasmid pDoc-K as a template. For generation of the *pqiAB* fragment for cloning, pqiA_GD_F and pqiB_GD_R primers were designed from the SL1344 genome with HindIII (forward primer) and BamHI (reverse primer) restriction sites. The PCR produced a kanamycin resistance cassette flanked by regions of homology upstream and downstream of *pqiAB*. The fragment was ligated into pDoc-C using HindIII and BamHI restriction sites, and the constructs were transformed and checked as described previously. The resulting construct consisted of the kanamycin cassette flanked by two I-SceI digestion sites in pDoc-C, which also contains the gene for sucrose sensitivity, *sacB*.

S. Typhimurium SL1344 was electroporated with both the pDoc-C construct and pACBSCE, a vector harbouring the λ Red recombination machinery, I-SceI digestion sites and the *I-sceI* gene. One transformant was grown in 1 ml of LB with arabinose to induce expression of the recombination machinery and I-SceI endonuclease. Cleavage of the I-sceI sites on pDoc-C allows linearisation of the kanamycin resistance cassette fragment for recombination onto the chromosome. After 5 hours of growth in arabinose, the culture was plated onto LA with kanamycin and 5% sucrose. The sucrose was used to differentiate between true recombinant colonies, which were kanamycin and sucrose resistant, and those that simply carry the pDoc-C construct, which were kanamycin resistant but sucrose sensitive due to the *sacB* gene. Kanamycin and sucrose resistant colonies were checked by colony PCR for the deletion of *pqiAB* using pqiA_STm_check_F and pqiB_STm_check_R primers (Figure 6.1A), which annealed to the flanking regions of *pqiAB*. The PCR included the *S.* Typhimurium SL1344 parent as a control, and a successful deletion was shown by a decrease in PCR product from ~3500 bp to ~2500 bp (Figure 6.1B, lanes 2 and 3). The confirmed *pqiAB::aph* deletion was transduced into a clean background of SL1344

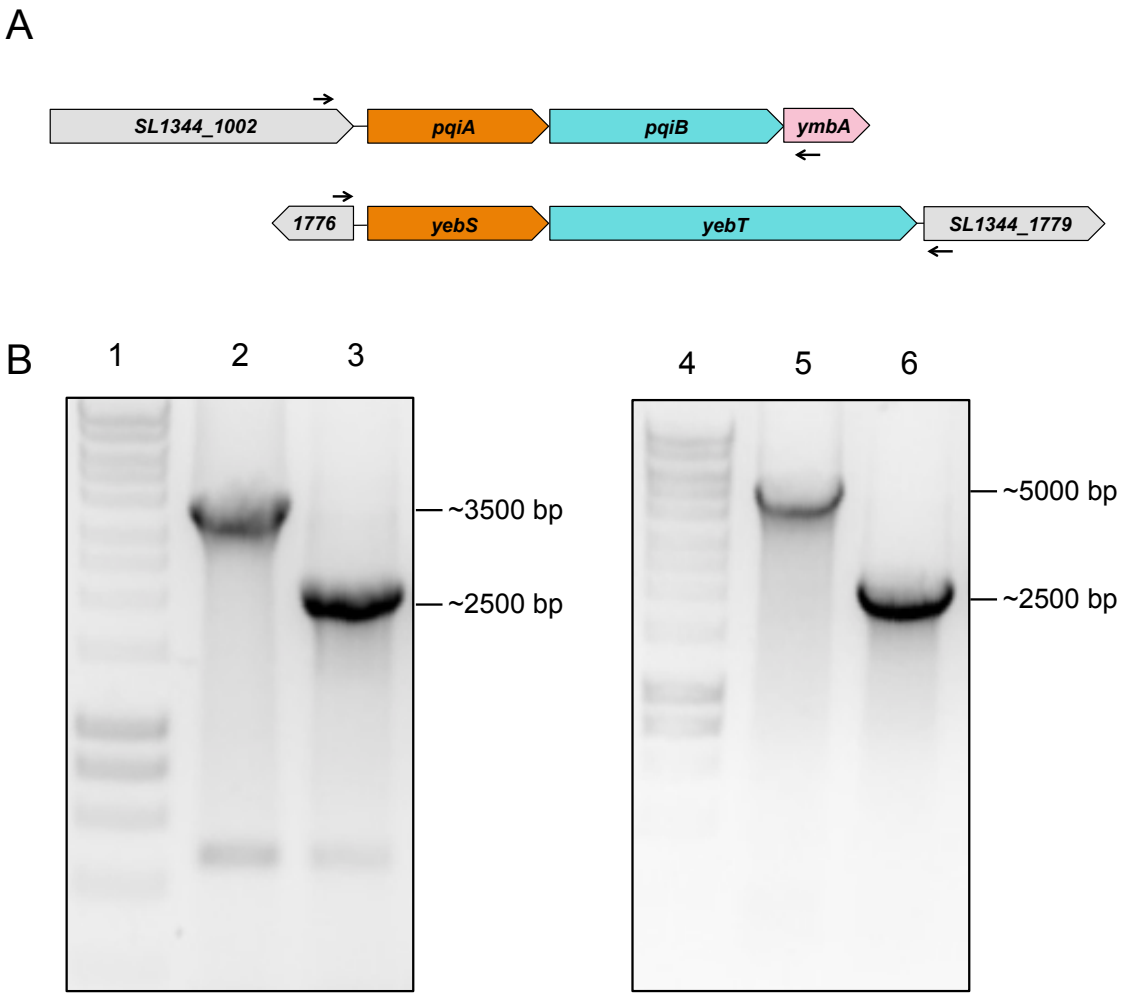


Figure 6.1. Confirmation of the disruption of *pqiAB* and *yebST* in *S.* Typhimurium SL1344. **A.** The positions of the *pqiAB* and *yebST* check primers and **B.** Gel electrophoresis of PCR products confirming the disruption of *pqiAB* and *yebST* by a kanamycin cassette using *pqiAB* and *yebST* check primers: lane 1: molecular weight marker, lane 2: SL1344, lane 3: SL1344 *pqiAB::aph,* lane 4: molecular weight marker, lane 5: SL1344, lane 6: SL1344 *yebST::aph.*

using P22 bacteriophage. Successful transduction was confirmed by colony PCR using the same check primers.

6.2.2. Construction of yebST *deletions in* Salmonella *Typhimurium SL1344*

The *yebST* genes span positions 1902180 to 1906065 on the SL1344 chromosome. Deletions were made in SL1344 using the same protocol described above for construction of the *pqiAB::aph* mutant. A *yebST* fragment was generated by PCR using the primer pair yebS_GD_F and yebT_GD_R and was ligated into pDoc-C using HindIII and BamHI restriction sites. Deletions in *yebST* were constructed as described for construction of the *pqiAB::aph* mutant, and potential mutants were screened using the primer pair yebS_STm_check_F and yebT_STm_check_R. A decrease from ~5000 bp to ~2500 bp showed successful deletion of *yebST* (Figure 6.1B, lanes 5 and 6). The deletion was transferred into a clean SL1344 background using P22 transduction to construct the kanamycin resistant mutant *yebST::aph*.

6.2.3. Preparation of lipopolysaccharide to check for the presence of O-antigen following P22 transduction

The P22 bacteriophage enters *S.* Typhimurium by the attachment and cleavage of O-antigen (Andres et al., 2010). To check that the transduced strains still had intact O-antigen, the lipopolysaccharide (LPS) was extracted from the SL1344 parent and from the *pqiAB::aph* and *yebST::aph* mutants from before and after P22 transduction. The LPS was analysed by SDS-PAGE and silver staining. A characteristic O-antigen ladder was observed for the parent strain and the strains from prior to P22 transduction, representing the different O-antigen chain lengths (Figure 6.2A). However, no O antigen was present in the LPS extracted from the mutants from after P22 transduction, suggesting that it had been removed during the transduction (Figure 6.2B). As O-antigen is essential for pathogenesis, these strains could not be used for further experiments.

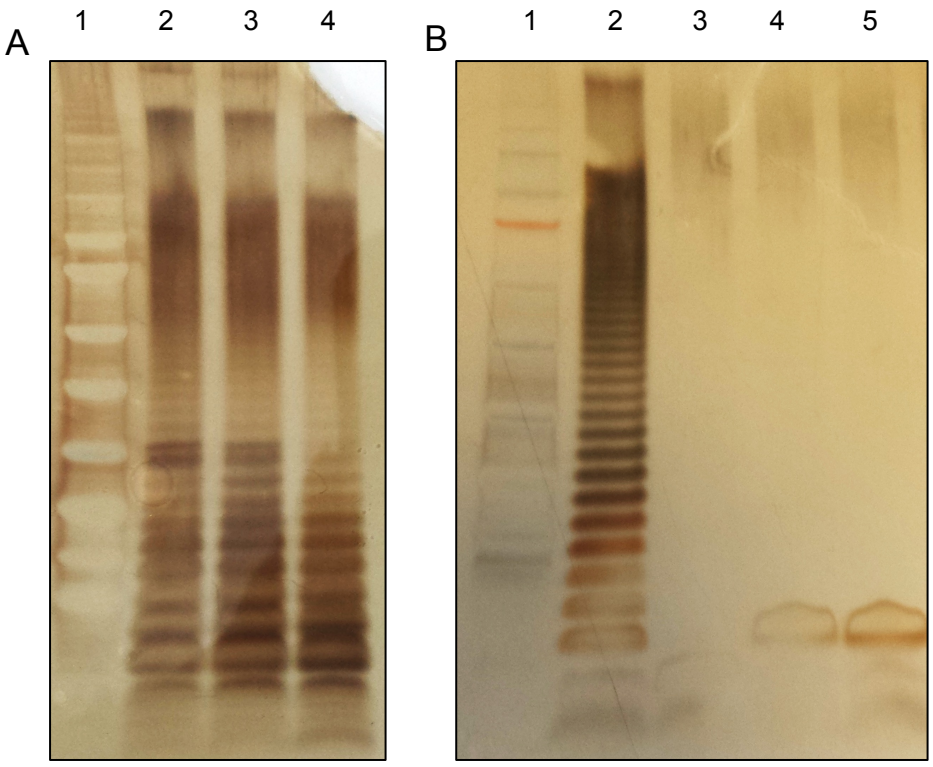


Figure 6.2. SDS-PAGE of the LPS profiles of SLI1344, *pqiAB::aph* and *yebST::aph* mutants before and after P22 transduction. **A.** Silver stained profiles of the mutants from before P22 transduction in comparison to the parent strain, lane 1: molecular weight ladder, lane 2: SL1344, lane 3: SL1344 *pqiAB::aph* and lane 4: SL1344 *yebST::aph*. **B.** Silver stained profiles of the mutants from after P22 transduction in comparison to the parent strain, lane 1: molecular weight ladder, lane 2: SL1344, lane 3: the *pqiAB::aph* mutant and lane 4: the *yebST::aph* mutant.

6.2.4. Genome sequencing of the pqiAB *and* yebST *deletion strains from prior to P22 transduction*

The genomes of *pqiAB::aph* and *yebST::aph* strains from before P22 transduction (with intact O-antigen) were sequenced alongside the genome of the SL1344 parent. When compared to the parent, the strains contained no additional mutations other than the desired gene deletions. The *pqiAB::aph* and *yebST::aph* mutants could therefore be used for further experimentation. To remove the kanamycin resistance, the *pqiAB::aph* and *yebST::aph* mutants were electroporated with plasmid pCP20, as described previously (see 4.2.1). The kanamycin sensitive Δ*pqiAB* and Δ*yebST* mutants were further checked by colony PCR to confirm removal of the kanamycin cassette.

6.2.5. Generation of a pqiAB yebST *double mutant*

To generate a Δ*pqiAB* Δ*yebST* mutant, the vectors used for the generation of the *yebST::aph* mutant were transformed into the Δ*pqiAB* deletion strain. The *yebST::aph* disruption was constructed as described previously (see 6.2.1 and 6.2.2) and colonies were checked using the same primers. Potential double mutants were first screened by colony PCR and then genome sequenced to check for second site mutations on the chromosome. No additional mutations were found. The Δ*pqiAB yebST::aph* mutant was transformed with pCP20 to remove the kanamycin cassette and generate the kanamycin sensitive Δ*pqiAB* Δ*yebST* mutant.

6.2.6. Generation of mlaD *deletion strains in* S. *Typhimurium SL1344 using an alternative method*

The *mlaD* genes span positions 3498363 to 3498998 on the *S.* Typhimurium SLI1344 chromosome. An *mlaD::aph* mutant was already available in a *S.* Typhimurium 14028s library (Porwollik et al., 2014). As the P22 transduction protocol resulted in loss of O-antigen, an alternative method was used to transfer this mutation to the *S.* Typhimurium SLI1344 strain. The primers mlaD_STm_DW_F and mlaD_STm_DW_R were designed

approximately 250 bp upstream and downstream from the *mlaD* gene deletion, flanking a region that is identical in the SL1344 and 14028s parent strains. The primers were used to amplify the *mlaD::aph* mutation from the library with large regions of homology upstream and downstream. Mutations in *mlaD* were generated in the parent strain and the Δ*pqiAB* Δ*yebST* mutant by the Datsenko and Wanner method (see 2.4.1). The *mlaD* mutants were checked by colony PCR using mlaD_STm_check_F and mlaD_STm check_R primers (Figure 6.3A). An increase in PCR product from ~1250 bp to ~2500 bp indicated that the *mlaD* gene had been replaced with a kanamycin resistance cassette (Figure 6.3B, lanes 2, 3 and 4). The *mlaD::aph* and the triple deletion strains were further checked by genome sequencing, revealing no unwanted mutations.

6.2.7. Serum bactericidal assay to test for the ability of the six strains to survive in healthy human serum

A serum bactericidal assay (SBA) assesses the ability of a strain to resist complement-mediated killing. To check for changes in resistance to serum, the SL1344 parent strain and all 5 deletion strains were exposed to serum obtained from healthy humans. *S.* Typhimurium SL1344 without O-antigen was used alongside as a known sensitive strain (positive control). Overnight cultures of all 7 strains were adjusted to an OD_{600} of approximately 0.1 and were mixed 1:10 with serum. Samples were taken before serum exposure (time 0) and at 45, 90 and 180 minutes and plated onto LA plates for counting of viable cells. After 45 minutes of exposure, there was substantial decrease in the number of CFUs for the positive control strain (Figure 6.4). After 180 minutes, a slight reduction in viability was observed for the parent strain and all 5 mutants, but overall they were equally resistant to the serum. These results indicate that MlaD, PqiAB and YebST are not required for *S.* Typhimurium resistance to human serum.

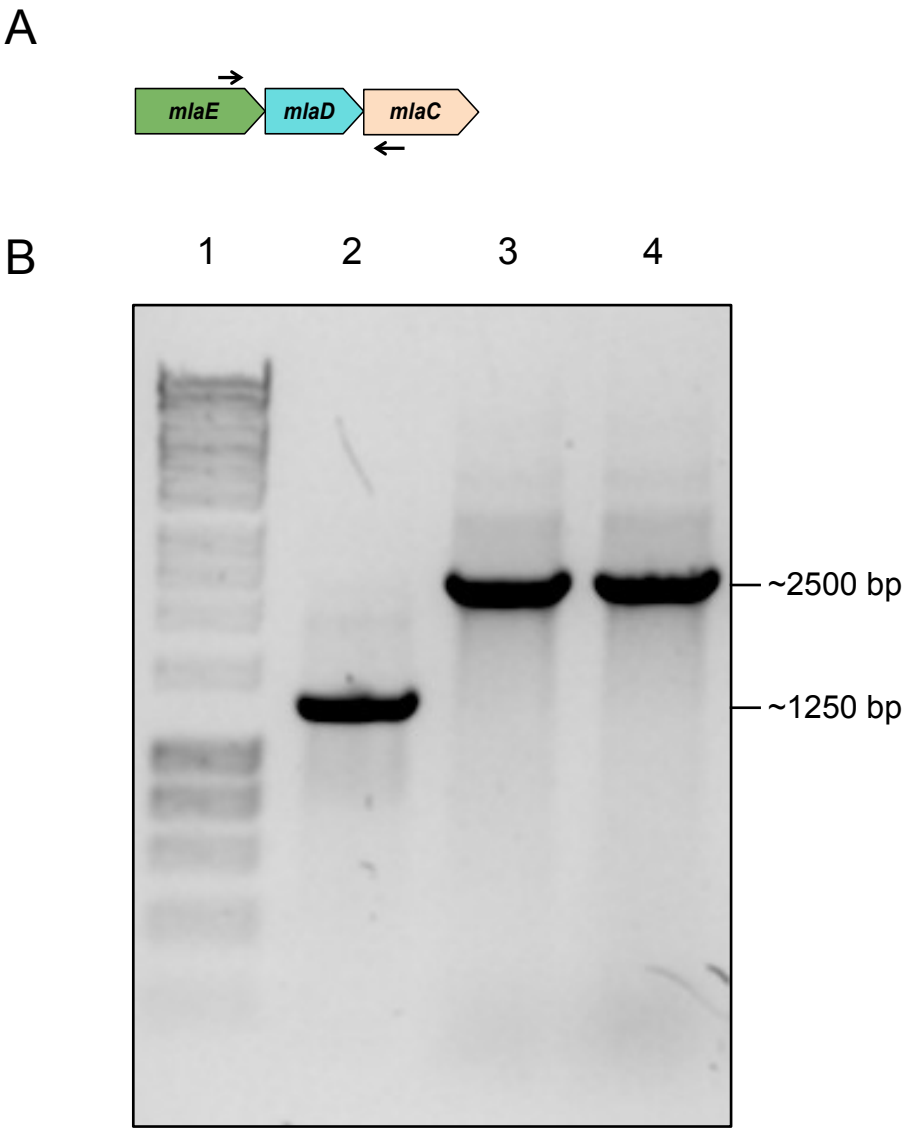


Figure 6.3. Confirmation of the disruptions in *mlaD* to construct an *mlaD::aph* mutant and triple Δ*pqiAB* Δ*yebST mlaD::aph* mutant in *S.* Typhimurium. **A.** The positions of the *mlaD* check primers and **B.** Gel electrophoresis of PCR products confirming the disruptions: lane 1: molecular weight marker, lane 2: SL1344 parent strain, lane 3: *mlaD::aph* mutant and lane 4: triple mutant**.**

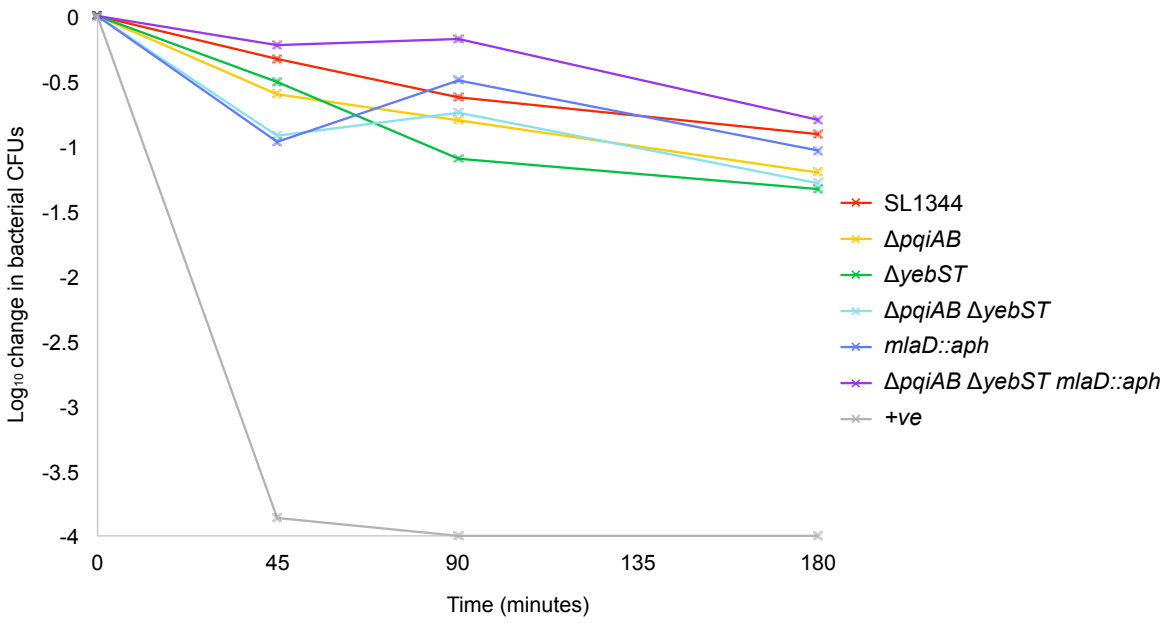


Figure 6.4. Bacterial survival of the *S.* Typhmurium strains in the presence of healthy serum. A serum bactericidal assay was used to check the ability to survive serum obtained from healthy humans. A strain without O-antigen was used alongside as a positive control.

6.2.8. Adhesion and invasion of HeLa cells by SL1344 and the mutant strains

The ability of the *Salmonella* mutants to adhere to and invade HeLa cells was tested to determine whether the MCE domain containing proteins are required for mammalian cell entry. For each of the 6 strains, 10^6 bacterial cells were added to 10^5 HeLa cells. The strains were left for either 30 minutes, to measure adherence, or for an additional 2 hours with gentamycin treatment, to measure invasion. At each time point the HeLa cells were lysed and plated onto LA for counting bacterial CFUs. The experiment repeated 3 times with biological replicates and mean number of CFUs was calculated for each strain.

For adherence, there were no obvious differences in the number of CFUs (Figure 6.5). A Kruskal-Wallis test was used to test the entire adherence dataset and confirmed that there were no significant differences (p=0.08). For invasion, the mean numbers of CFUs for the Δ*pqiAB*, Δ*pqiAB* Δ*yebST* and *mlaD::aph* mutants were noticeably lower than the parent strain (Figure 6.5). However, the decreased mean was only observed when there was a greater amount of variance. A Kruskal-Wallis test revealed that there were no significant differences in the invasion dataset (p=0.06). These results indicate that MlaD, PqiAB and YebST are not required for the adherence and early invasion of HeLa cells.

6.2.9. Survival of the strains in bone marrow derived macrophages

The survival of *S.* Typhimurium inside macrophages is important for the establishment of murine infection. To test whether MlaD, PqiAB and YebST are important for survival in macrophages, bone marrow derived macrophages were infected with the 6 strains, as described for the infection of HeLa cells. After 2, 5 and 24 hours, the cells were lysed and plated on LA. There were no large differences in the number of CFUs at any time point (Figure 6.6). This was confirmed by the Kruskal-Wallis test that produced p values of 0.5, 0.44 and 0.21 for the 2, 5 and 24 hour datasets, respectively. Therefore, it was concluded that MlaD, PqiAB and YebST are not required for the entry or survival of *S.* Typhimurium in bone marrow derived macrophages.

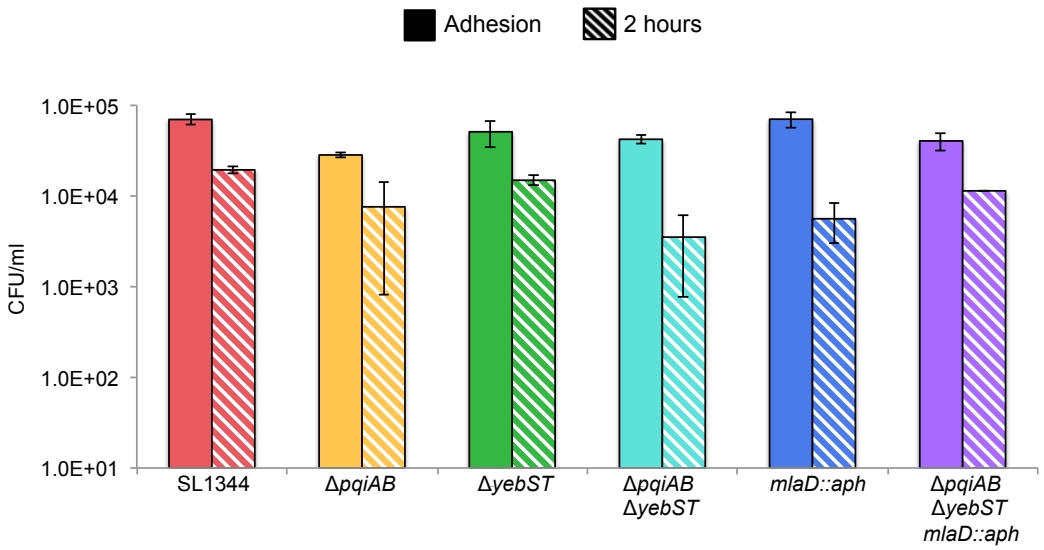


Figure 6.5. Adhesion and invasion of HeLa cells with the parent strain and five deletion strains. For adhesion, the bacteria were left to adhere to HeLa cells for 30 minutes and the HeLa cells were lysed and plated onto LB agar plates. For the 2 hour time point, the bacteria were given an additional 2 hours in the presence of gentamycin to test for invasion only. The bars are a mean of 3 biological replicates and the error bars represent the standard error of the mean.

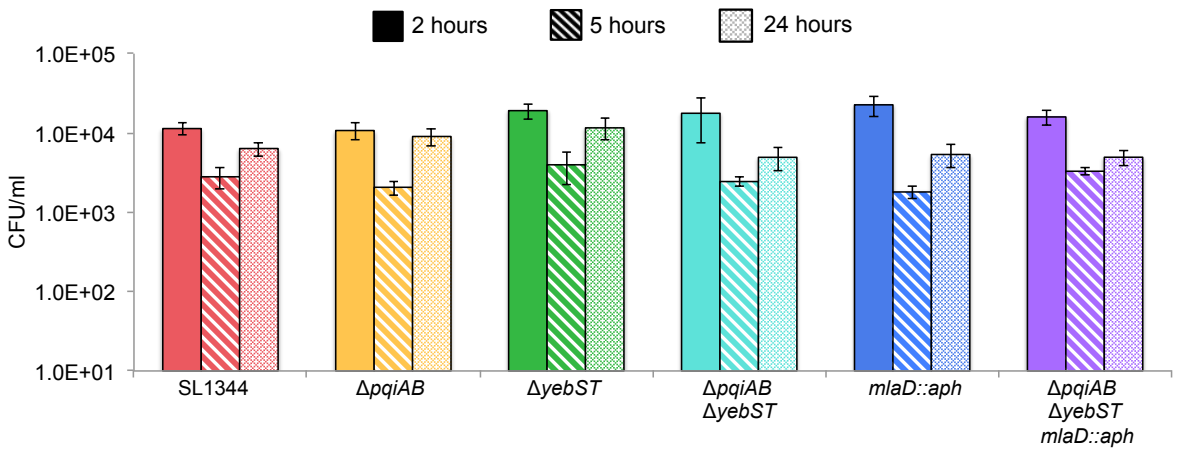


Figure 6.6. Survival of the parent strain and five deletion strains in bone marrow derived macrophages after 2, 5 and 24 hours of invasion. The bars are a mean of 3 biological replicates and the error bars represent the standard error of the mean.

6.2.10. Intraperitoneal infection of mice with the six Salmonella *strains to determine the importance of MCE domains during murine infection*

S. Typhimurium is known to cause systemic infection of mice, particularly in organs of the reticuloendothelial system such as the liver and spleen. A pilot study was conducted in mice to check for differences in systemic infection between the strains. Two C57/BL6 mice were infected via intraperitoneal injection with 1000 bacterial cells for each strain. After 3 days the mice were culled and the livers and spleens were mashed and plated onto LB plates. The pilot study revealed decreased numbers of CFUs/organ for some of the mutants, and therefore the same experiment was repeated with 4 mice per strain. The data from both the pilot and scale-up experiments were combined and the mean numbers of CFUs/organ were calculated for each strain. No viable bacteria were observed in a mouse infected with the Δ*yebST* mutant or in another infected with the triple mutant. This was possibly due to unsuccessful injections and these data were therefore excluded from further analysis.

In both the liver and spleen there appeared to be attenuation of the Δ*pqiAB* Δ*yebST*, *mlaD::aph* and the triple mutants, but no differences were observed for the Δ*pqiAB* and Δ*yebST* single mutants (Figure 6.7A and B). The p values from a Kruskal-Wallis test were 2.4e-05 and 1.0e-04 for the liver and spleen datasets, respectively, confirming that there were significant differences present. Pairwise comparisons between datasets for each strain were done using a Mann-Whitney U test, and the Bonferroni correction was applied to all p values to correct for multiple testing (Figure 6.8). There were no significant differences between the datasets for the parent, Δ*pqiAB* and Δ*yebST* deletion strains in either the liver or the spleen. The numbers of CFUs for the Δ*pqiAB* Δ*yebST, mlaD::aph* and triple deletion strains were significantly different from the parent strain, with p values of less than 0.01 in both organs. In the liver, the numbers of CFUs for these 3 attenuated strains were also significantly different from one another, revealing that the mutants become increasingly attenuated in the following order: Δ*pqiAB* Δ*yebST*, *mlaD::aph* and Δ*pqiAB* Δ*yebST mlaD::aph*. Although the same trend was observed in the spleen, the differences were

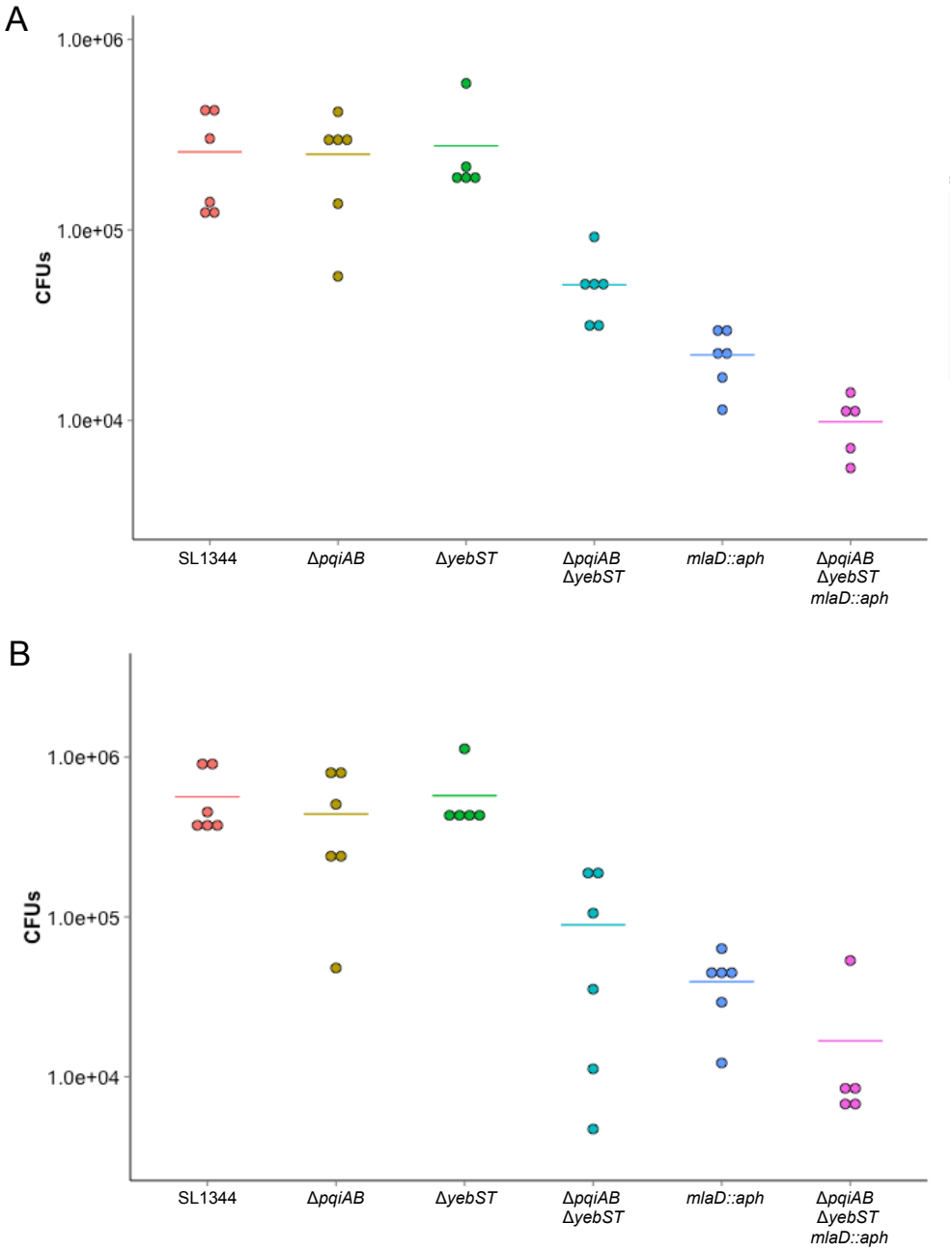


Figure 6.7. Survival of the parent strain and five deletion strains in C57/BL6 mice. The mice were infected with the strains for 3 days via IP injection. The CFUs/organ were counted in the liver (**A**) and spleen (**B**). Each point represents the number of CFUs/organ from a single mouse, and the line indicates the mean number of CFUs.

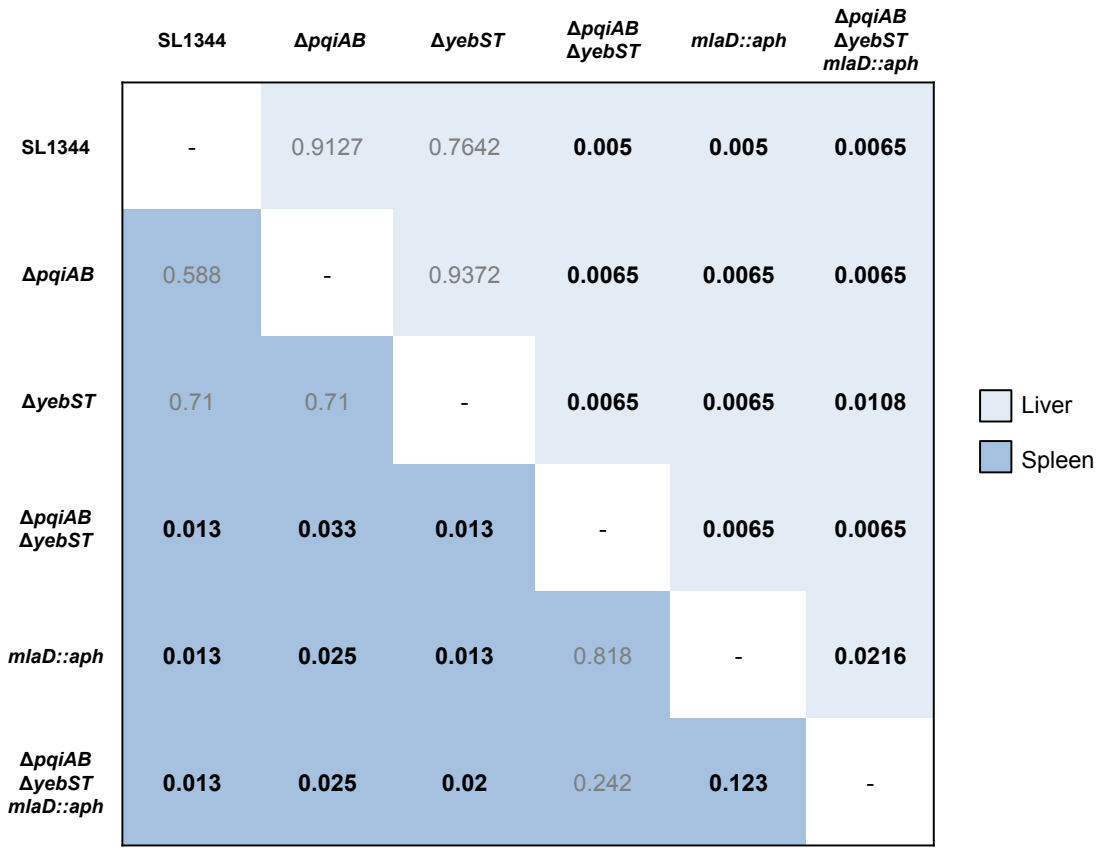


Figure 6.8. Pairwise comparisons of all datasets obtained from the mouse infection experiments using a Mann-Whitney U test. The Bonferroni correction was applied to correct for multiple comparisons. The numbers in bold represent significant p values ($p<0.05$).

not significant. Overall these results reveal an additive phenotype for the Δ*pqiAB* Δ*yebST* and triple deletion strains. Furthermore, they suggest that all MCE domain containing proteins, particularly MlaD, have some role in the systemic infection of *S*. Typhimurium.

6.2.11. Growth of the parent strain and triple mutant on lyso-phosphatidylglycerol as a sole carbon source

A previous study has demonstrated that *S*. Typhimurium metabolises phospholipids in bile, and that it can use some lyso-phospholipids as a sole source of carbon (Antunes et al., 2011). With the clear link of MlaD, PqiB and YebT to phospholipids, it was appropriate to test the growth of the *S*. Typhimurium parent strain and triple mutant with lyso-phospholipids as the only carbon source in M9 minimal medium. To first check the growth profiles, the strains were inoculated into M9 medium supplemented with 200 mM histidine (as *S*. Typhimurium SL1344 is a histidine auxotroph) with or without 10 mg/ml glucose as a carbon source. Neither of the strains grew in M9 and histidine only, confirming that the strain could not grow without a specific carbon source (Figure 6.9A). The growth profiles with glucose were the same for both strains, revealing that the mutant has no altered ability to grow in minimal media.

To test for changes in the ability to grow on lyso-phospholipids, the strains were grown in M9 supplemented with 200 mM histidine and 1 mg/ml lyso-PG. The growth profiles revealed a growth defect for the triple deletion strain, which was present in all 3 biological replicates (Figure 6.9B). The loss of the ability to grow on lyso-PG indicates that MlaD, PqiAB and/or YebST function in the utilisation of lyso-PG as a carbon source.

6.3. Discussion

The first aim of this chapter was to determine whether MlaD, PqiAB and YebST were important for infection *in vitro*. Following the construction of the *mlaD::aph*, Δ*pqiAB*, Δ*yebST,* Δ*pqiAB* Δ*yebST* and Δ*pqiAB* Δ*yebST mlaD::aph* deletion strains, the ability of

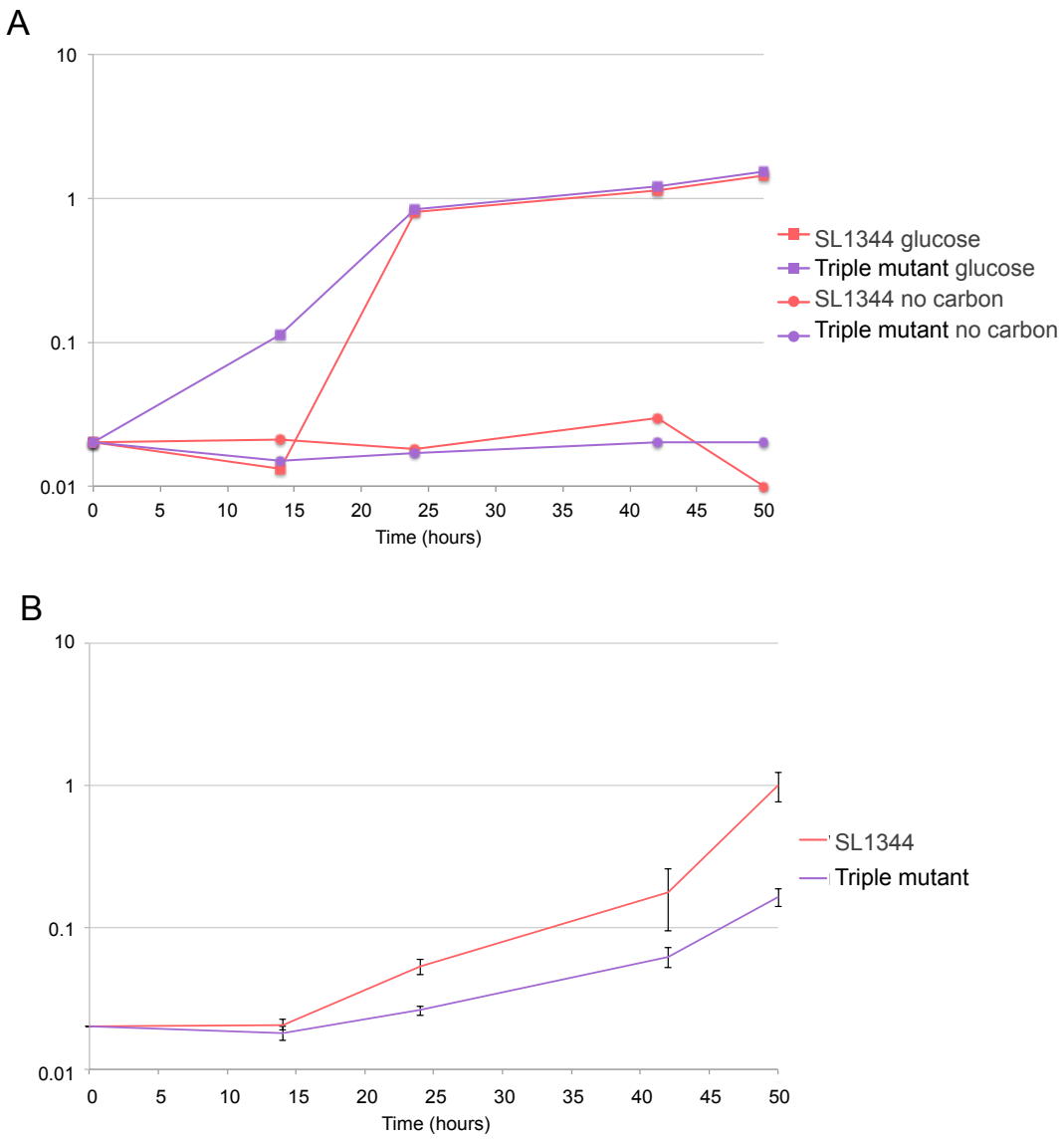


Figure 6.9. Growth of the parent and triple mutant on lyso-PG as a carbon source. The strains were grown in M9 minimal media supplemented with 200 mM histidine. The strains were grown with either no carbon source (**A**-circles), 10 mg/ml glucose (**A**-squares) or 1 mg/ml lyso-PG (**B**). For **B**, each point represents the mean OD_{600} from 3 biological replicates and the error bars represent the standard error of the mean.

the strains to survive challenge by serum from healthy humans was tested. The results show that, as with the parent strain, all 5 mutants were resistant to serum, and are therefore able to withstand complement-mediated killing.

In vitro tissue culture experiments also indicated that MlaD, PqiAB and YebST are not required for the attachment or early invasion of HeLa cells. These results are in contrast to similar experiments with *Shigella sonnei* (SS01327) and *Vibrio parahaemolyticus* (MAM-7), which demonstrated a role for MAM-7 in the attachment and invasion of HeLa cells (Krachler et al., 2011; Mahmoud et al., 2016). It is possible that these results demonstrate differences between the role of MCE domains between bacterial species. Krachler *et al.* (2011) also reported that the MAM-7 deletion strain was less cytotoxic for mammalian cells. Further experiments will be necessary to determine whether the same is true for *S.* Typhimurium mutants.

Whilst HeLa cells are ideal for cell entry experiments, they do not fairly represent the harsh intracellular environment in which *S.* Typhimurium has to persist to establish infection. An important stage of *S.* Typhimurium infection is the ability to survive in macrophages by withstanding factors designed to kill invading microorganisms. The results from this chapter indicate that MCE domain containing proteins are not required for the survival of *S.* Typhimurium in macrophages derived from bone marrow. Together with the results from the HeLa experiments, these results indicate that MCE domain containing proteins are not necessary for *S.* Typhimurium mammalian cell entry or survival.

The second aim of this chapter was to determine whether MCE domain containing proteins are important for *S.* Typhimurium infection of mice. The results revealed that the Δ*pqiAB* Δ*yebST*, *mlaD::aph* and Δ*pqiAB* Δ*yebST mlaD::aph* mutants were all attenuated in mice, but the Δ*pqiAB* and Δ*yebST* mutants were not. The trend in the data suggests that the *mlaD* deletion strain is more attenuated than the Δ*pqiAB* Δ*yebST* deletion strain, and that the triple mutant is the most attenuated. This trend was proven to be significant in the liver, but not in the spleen. However the spread of the data for the spleen was large and it is therefore likely that this trend would be significant with more replicates. The data indicate

that MlaD is the most important MCE domain containing protein for establishing systemic infection (at least in the liver), and is more important than both PqiAB and YebST combined. Furthermore, additive phenotypes for the Δ*pqiAB* Δ*yebST* and triple deletion strains were identified, similar to those observed in *E. coli*. Overall, these results reveal that MlaD, PqiAB and YebST all contribute to some extent to *S.* Typhimurium systemic infection.

In the previous chapter, a role for MCE domain containing proteins in phospholipid uptake was considered, but experiments with *E. coli* K-12 BW25113 were unsuccessful due to a lack of growth on phospholipids. With the knowledge that *S.* Typhimurium is able to grow on lyso-PLs (Antunes et al., 2011), the third aim of this chapter was to determine whether the proteins are required for this process. Decreased growth was observed for the triple mutant on lyso-PG, suggesting that the proteins are necessary for either uptake or metabolism of phospholipids. Future experiments will test growth of the individual mutants to determine whether all three MCE domain-containing proteins are involved in this process. Furthermore, other phospholipids will be tested to determine whether the phenotype is general or specific to lyso-PG.

Overall, the experiments described in this chapter revealed for the first time that MCE domain containing proteins in Gram-negative bacteria are important for mammalian infection. With the knowledge that MCE domain containing proteins are unlikely to be required for *S.* Typhimurium mammalian cell entry and/or survival, other possible scenarios for this attenuation phenotype have to be considered. Whilst there are many harsh environments for *S.* Typhimurium to withstand, the ability to survive in bile is particularly important for systemic infection. It is known that phospholipids in bile can act as a carbon source for *S.* Typhimurium, and therefore the ability to utilise phospholipids might be an important part in establishing systemic infection (Antunes et al., 2011). It is therefore possible that the attenuation phenotype observed in mice is linked to the role of MCE domain containing proteins in phospholipid metabolism in bile, a hypothesis supported by the decreased ability of the triple mutant to grow on lyso-PG.

CHAPTER 7

FINAL DISCUSSION

7.1. The evolution of the MCE domain-containing proteins as components of transporters

Prior to this study, the link between MCE domains and transport had already been established. MCE domain containing proteins have previously been identified as homologous to the substrate binding components of ABC transporters (Kumar et al., 2005) and an in depth bioinformatic study in 2007 identified a strong link between MCE domains and ABC transporters in Actinobacteria (Casali and Riley, 2007).

In this study, a similar bioinformatic study was conducted but with more focus on the different types of MCE domain containing proteins, particularly in Proteobacteria. These analyses revealed that MCE domains have remained conserved throughout bacterial evolution and that proteins with greater than two MCE domains, such as PqiB and YebT, are specific to and have evolved within Proteobacteria. Furthermore, it was shown that single MCE domain containing proteins are nearly always part of an ABC transporter, as already indicated (Casali and Riley, 2007), but that the multi-domain proteins are part of a novel type of transporter with PqiA domain containing proteins. This work complements the existing literature that reports a role of MCE domains in transport, and for the first time highlights the distinction between single and multi-MCE domain containing proteins.

The co-evolution of multi-MCE domain containing proteins and PqiA domain containing proteins in Proteobacteria suggests that these proteins are functionally linked. Whilst the mechanisms of ABC transporter domains are fairly well understood, the function of a PqiA domain is completely unknown. Other than the predicted homology to an antiporter domain (Hunter et al., 2009), there is no evidence that PqiA domains are involved in transport. Therefore, it is important for future experiments to focus on the function of PqiA domains as transporters.

The presence of the multi-domain operons in only Proteobacteria suggests that they have evolved a role that is not required in other bacterial phyla. As Proteobacteria are a large and diverse group of organisms, and the multi-domain proteins are found in all of the major classes, it is difficult to identify why they would be specific to this phylum. Therefore,

a more extensive analysis is required to look more closely at the species in Proteobacteria that harbour the multi-domain protein types, and their lifestyles.

7.2. The relationship between the three pathways in *E. coli* and their link to phospholipids

The Mla pathway of *E. coli* was first identified in 2009 (Malinverni and Silhavy, 2009), as a retrograde phospholipid trafficking pathway and subsequent studies have looked at the more specific characteristics of the pathway (Chong et al., 2015; Thong et al., 2016). In this thesis the two other MCE operons in *E. coli* were highlighted: *pqiAB-ymbA* and *yebST,* both of which encode multi-MCE domain containing proteins and were previously unstudied. The major aim of this study was to gain a general understanding of the multi-domain proteins in *E. coli*, PqiB and YebT, by identifying phenotypes and determining sub-cellular locations. This aim was met and further explored to understand the relationships between PqiB, YebT and the single domain protein, MlaD. It was shown that PqiB and YebT localise to the same place in the cell as MlaD: the inner membrane. Sulfobetaine sensitivity and vancomycin resistance phenotypes revealed that the functions of the three proteins are distinct but overlapping, with the multi-domain pathways overlapping the most. Furthermore, the use of lipidomics revealed that, despite the reported role of the Mla pathway (Malinverni and Silhavy, 2009), none of the MCE pathways play a major role in the phospholipid composition of the cell membrane. However, MlaD, PqiB and YebT proteins were all found to bind phospholipids, suggesting that these proteins most likely impact the cell membrane in some capacity. These novel findings explore, for the first time, the functional relationships between the three MCE pathways in *E. coli*.

Whilst this is the first study to look at all three MCE domain containing pathways and identify that there are differences between them– it is still not known how these pathways differ. It is possible that they have different substrate specificities, are induced under different conditions or are simply functionally redundant – although the phenotypes observed here make the latter unlikely. The varying sensitivities of the mutants to

sulfobetaines with different chain lengths suggest that the different pathways might vary by their chain length specificities. The lack of sensitivity to SDS alone suggests that the phenotypes might be specific to zwitterionic detergents, or specifically sulfobetaines. Future experiments will focus on identifying the range of detergents that these phenotypes are observed for.

Arguably the easiest way to determine chain length specificities is to scrutinise the phospholipid species bound to MlaD, PqiB and YebT. Further investigation of the protein-lipid extracts by mass spectrometry will identify if there are differences, and this could be a key experiment for determining the differences between the three pathways in *E. coli*. Similarly, determining by mass spectrometry the acyl chain lengths in membrane phospholipids derived from the various mutants will help determine whether the proteins are involved in more subtle alterations of membrane phospholipids.

Future work should also focus on the mechanism of transport within the MCE pathways, and what are the precise roles of each component. This has recently been a focus of research on the Mla pathway, where MlaD has been shown to form a hexamer as part of a stable ABC transporter complex with MlaB, E and F, located in the inner membrane (Thong et al., 2016). It has been proposed that the periplasmic chaperone, MlaC, does not strongly interact with either the inner membrane complex or outer membrane complex (MlaA-OmpC), and that the ability of MlaC to transfer phospholipids to/from these complexes is purely due to its affinity to phospholipids (Thong et al., 2016). Whilst such detailed studies do not exist for the PqiAB-YmbA pathway, the results presented here suggest that PqiB, like MlaD, might also form a hexamer in the inner membrane. Furthermore, a recent study has shown that PqiB in the inner membrane and YmbA in the outer membrane interact, suggesting that PqiB spans the periplasm (Nakayama and Zhang-Akiyama, 2016). It is likely that the same is true for YebT as it is a larger protein than PqiB, but this has not been confirmed. The ability of PqiB and YebT to span the periplasm would explain why the *pqiAB-ymbA* and *yebST* operons do not encode predicted periplasmic chaperones.

Currently, it is unknown whether the PqiAB-YmbA and YebST systems are also associated with OMPs, like the Mla pathway (Chong et al., 2015). Further experimentation is required to identify any other components of the multi-domain pathways. Simple protein pull-down experiments could identify other components through their interactions. Alternatively the phenotypes discovered in this thesis could be manipulated to identify additional components. The most thorough way of doing this is through the use of a high-density transposon library, which contains transposon insertions throughout the genome. Such a library could be grown on lauryl sulfobetaine and sequenced to identify all the mutants in *E. coli* that are inhibited by the detergent. Any candidates can be further tested, for example by protein-protein interactions, to determine if they are linked to the pathway of interest.

7.3. The future of MCE domain structural biology

There is currently no published structure for an MCE domain containing protein. To address this, one aim of this study was to determine the first crystal structure of the domain, with particular focus on PqiB. Whilst this was not achieved in this study, significant steps were taken toward optimising conditions for crystal growth and improving the quality of crystals. However, during the writing of this thesis, it became apparent that another group had successfully obtained the structures of MlaD, PqiB and YebT by a combination of crystallography and cryo electron microscopy. Collaboration has been established with these researchers, and all future structural biology will take place in their institution.

Now that structures exist for MlaD, PqiB and YebT, the key residues in the MCE HMM (see Figure 1.12) can be highlighted in the structures to identify why they are so conserved. Furthermore, these key residues can be altered to determine how important they are for the structure and this can be linked with genetics and lipid binding assays to determine their function.

7.4. MCE domains are not required for *Salmonella* to enter mammalian cells

Since the first MCE domain was identified in 1993, the role of MCE domain containing proteins in infection has been a main focus of the literature. As the name suggests, the original function was thought to directly involve mammalian cell entry (Arruda et al., 1993). However, in this study it was revealed that MCE domain containing proteins are not required for *Salmonella* entry of HeLa cells or survival in macrophages, suggesting that they are not directly involved in mammalian cell entry.

Whilst these results might appear counter-intuitive, few studies have actually concluded that mutants display decreased cell entry. The Mce1A protein found in *M. tuberculosis* was the original protein shown to aid *E. coli* entry and survival into macrophages (Arruda et al., 1993). Beads coated with Mce1A also have the ability to enter HeLa cells (Arruda et al., 1993; Chitale et al., 2001; El-Shazly et al., 2007; Saini et al., 2008). However *mce1* mutants do not show decreased ability to enter and survive in mammalian cells (Marjanovic et al., 2010; Shimono et al., 2003). This indicates that whilst MCE domains might be important for cell entry and survival in particular laboratory conditions, this might not be their primary role *in vivo*. This hypothesis is supported by the fact that MCE domains are present in a wide range of non-pathogenic bacterial and eukaryotic species (Figures 3.2 and 3.3).

7.5. MCE domain containing proteins are required for systemic infection of *Salmonella* Typhimurium and for phospholipid uptake

Several studies in Actinobacteria have identified that MCE domain containing proteins are important for mammalian infection *in vivo* (Marjanovic et al., 2010; Senaratne et al., 2008). However, prior to this work no parallel studies existed for Gram-negative bacteria. In agreement with the literature, the results presented here suggest that MCE domain containing proteins in *S.* Typhimurium are important for systemic infection in mice, although their precise role remains unclear.

In Actinobacteria, the same proteins that are required for infection are also required for uptake of various lipids (Cantrell et al., 2013; Klepp et al., 2012; Mohn et al., 2008; Pandey and Sassetti, 2008), suggesting that the attenuation phenotype observed could be linked to lipid uptake. This work indicates that the same is true for *Salmonella*, as a strain lacking all three MCE-domain containing proteins is less able to grow on phospholipids. This suggests that at least one of the MCE domain containing proteins is important for phospholipid uptake. As phospholipids are an important nutrient source in bile (Antunes et al., 2011), this could explain why these strains are attenuated in mice. Interestingly, the same mutants in *S.* Typhimurium that were attenuated in mice (the double *pqiAB yebST* mutant, *mlaD* mutant and triple mutant) were also resistant to vancomycin in *E. coli*. As increased vancomycin resistance might be linked to phospholipid accumulation on the cell surface (Sutterlin et al., 2014), it is possible that the attenuation phenotype observed for *Salmonella* is a direct consequence of phospholipid accumulation on the cell surface.

Further experimentation is required to determine if the observed attenuation phenotype is due to the decreased ability of the mutants to utilise phospholipids in bile. Simple growth curves will identify whether the MCE domain containing proteins are required for growth in bile. Furthermore, longer-term infection studies can investigate the importance of MCE domain containing proteins in the faecal shedding of *Salmonella*, which is largely sourced from the release of infected bile into the intestinal lumen (Monack et al., 2004).

7.6. A hypothesised model for the Mla, Pqi and Yeb pathways

The bioinformatics and lipid binding results presented here reveal that there is a clear link to phospholipid trafficking, which is supported by the literature on the Mla pathway (Malinverni and Silhavy, 2009). The *Salmonella* phospholipid growth curves and lipid uptake experiments in the literature indicate that the pathways are specifically involved in uptake, which is supported by the fact that substrate-binding proteins (to which MCE

domains containing proteins are homologous) have only been identified in importers (Higgins, 2001).

If all three pathways are involved in uptake from the external environment, this would suggest that description of the role of the Mla pathway is incomplete, and that this import pathway functions in uptake of phospholipids from outside the cell, not just from the outer membrane (Malinverni and Silhavy, 2009). Therefore, it is possible that the three pathway can function in maintaining the outer membrane but also in providing phospholipids as a carbon source when other nutrients are scarce, such as in bile. Currently, the final destination of the phospholipids trafficked by the Mla system remains unknown. It has not been determined whether the phospholipids are inserted into the inner membrane or whether they enter the cytoplasm. The requirement of the MCE domain containing proteins for growth on phospholipids suggests that at least one of the MCE pathways result in phospholipids reaching the cytoplasm for degradation.

With this in mind, a model was hypothesised for all three MCE pathways in Proteobacteria (Figure 7.1). This model proposes that all three pathways take phospholipids from the cell surface and external environment. It is possible that these two processes are not distinct, and that the phospholipids are first inserted into the outer leaflet and then removed for uptake. The model for all three MCE pathways starts with OMPs, which for the PqiB and YebT systems are currently unknown. In the Mla pathway, MlaA obtains phospholipids from OmpC, which are subsequently passed to MlaC and then onto the inner membrane MlaBDEF complex, which includes an MlaD hexamer (Thong et al., 2016) (Figure 7.1A). The proposed mechanism for the Pqi pathways is similar, where phospholipids are passed from the unknown OMP to PqiB, via YmbA (Figure 7.1B). In this model, PqiB forms a hexamer that spans the periplasm, through which the phospholipids pass to reach PqiA in the inner membrane. The proposed model for the YebST pathway is the same, except the outer membrane lipoprotein is not present and YebT can directly interact with the OMP (Figure 7.1C). In this model, phospholipids trafficked by the MCE pathways end up the cytoplasm for degradation. The model presented here seems

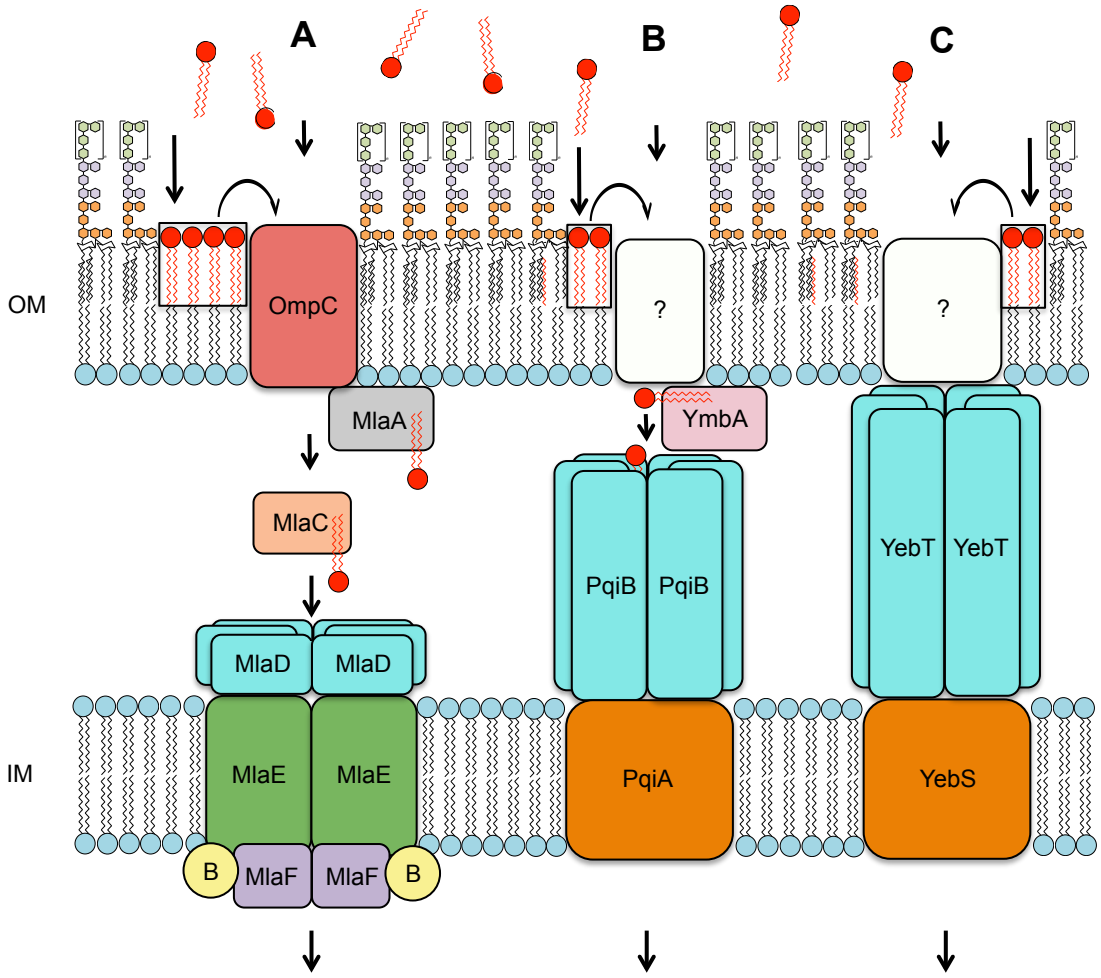


Figure 7.1. A model for the MCE domain containing phospholipid uptake pathways in *E. coli*. Phospholipids are taken from the cell surface or from outside the cell. **A.** The Mla pathway: phospholipids pass through OmpC, to the inner membrane lipoprotein, MlaA. The phospholipids are passed from MlaA to MlaC and through the MlaD channel onto MlaFEB inner membrane complex. The phospholipids are transported through the MlaE permease and into the cytoplasm for degradation. **B.** The Pqi pathway: phospholipids pass through an unknown OMP to the lipoprotein YmbA. From YmbA, phospholipids pass to the PqiB channel and through to PqiA in the inner membrane. Here, the phospholipids are transported through PqiA by and unknown mechanism. **C.** The Yeb pathway: the phospholipids pass from an unknown OMP straight through the YebT pore and on to YebS in the inner membrane.

the most likely given the existing literature and the results presented in this thesis. However, aside from the growth on phospholipids, there is no experimental data to suggest the direction of the phospholipid transport. To address this future experiments will involve tracking radiolabelled phospholipids in *E. coli*, as described by Jones and Osborn (1977).

As fatty acids are the major component of phospholipids, links to the fatty acid degradation (Fad) system must also be considered. FadL is required for transport of longer chain fatty acids across the outer membrane (Black, 1991), and FadD for transport across the inner membrane (Black et al., 1992). However, the method by which fatty acids cross the periplasm is currently unknown (Parsons and Rock, 2013). The link of the MCE pathways to lipid uptake indicates that they could be a missing part of the fatty acid degradation system in Proteobacteria. However, experimentation with the various components of the Fad system and with fatty acids is required to test this hypothesis.

Overall, this study has expanded our knowledge of MCE domain containing proteins in Proteobacteria. The results demonstrate that MCE domains are not directly involved in mammalian cell entry, as originally thought (Arruda et al., 1993). Instead the evidence suggests a role in phospholipid transport in Proteobacteria that is important for phospholipid uptake, either as a source of carbon or for maintenance of the cell envelope.

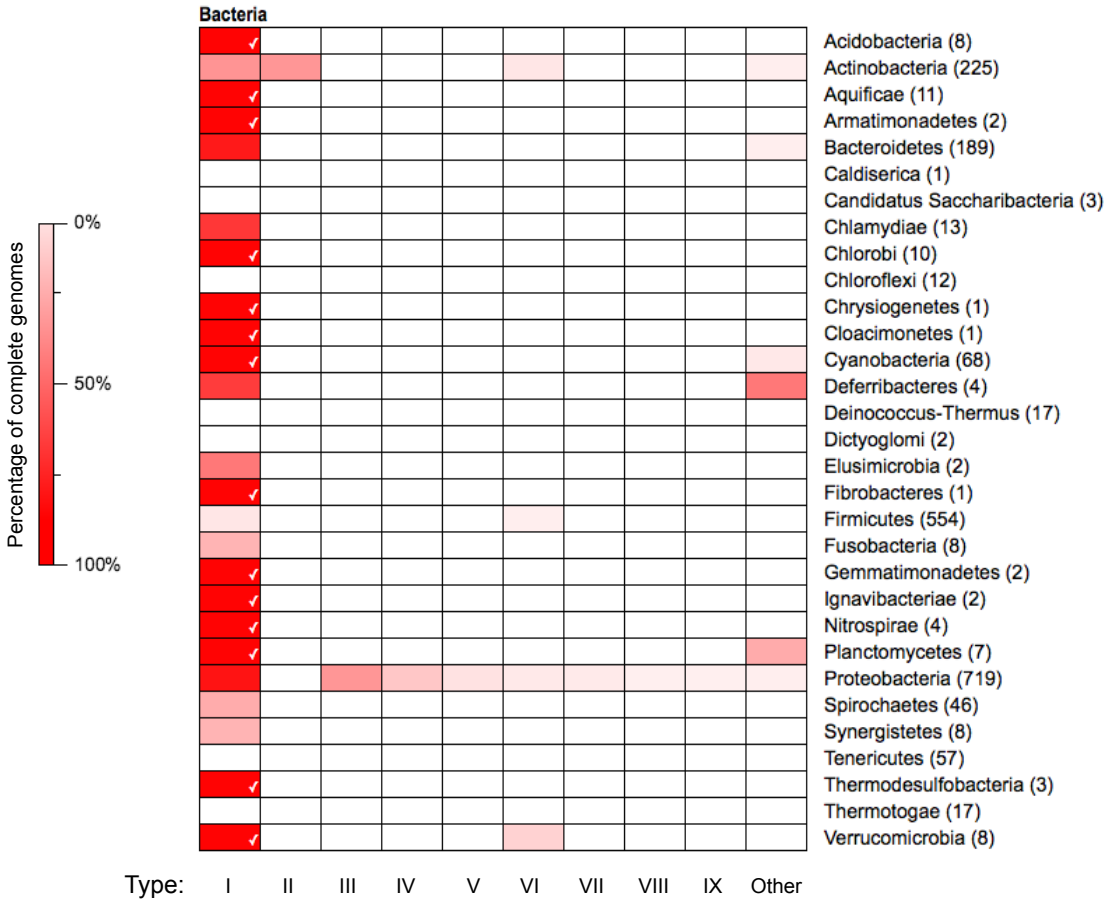
APPENDIX

Appendix i – list of all MCE domain containing protein architectures

Architecture ID	No. proteins	Description
1	11012	MCE
2	2718	MCE, DUF3407
3	2026	MCE, MCE, MCE
4	1042	MCE, MCE, MCE, MCE, MCE, MCE, MCE
5	236	MCE, MCE, MCE, MCE, MCE, MCE
6	125	MCE, MCE
7	35	MCE, DUF330
8	20	MCE, MCE, MCE, MCE
9	15	MCE, MCE, MCE, MCE, MCE
10	8	MCE, MCE, DUF3407
11	7	MCE, MCE, MCE, MCE, MCE, MCE, MCE, MCE
12	4	MCE, DUF3407, MCE
13	4	MCE, MCE, MCE, DUF330
14	2	MCE, AP2
15	2	PqiA, PqiA, MCE, MCE, MCE, DUF330
16	2	PqiA, MCE, MCE
17	2	MCE, PLU-1
18	2	MCE, OmpA
19	1	MCE, MCE, MCE, MCE, Molybdopterin
20	1	PqiA, PqiA, MCE, MCE, MCE, MCE, MCE, MCE, MCE, Nol1_Nop2_Fmu, Methyltransf_31, Methyltransf_18, Methyltransf_26, Nol1_Nop2_Fmu, Nol1_Nop2_Fmu_2
21	1	MCE, MCE, MCE, MCE, MCE, MCE, MCE, MCE, MCE, MCE, MCE, MCE, MCE
22	1	MCE, DUF4455
23	1	OTCace_N, OTCace, MCE
24	1	DUF3407, MCE

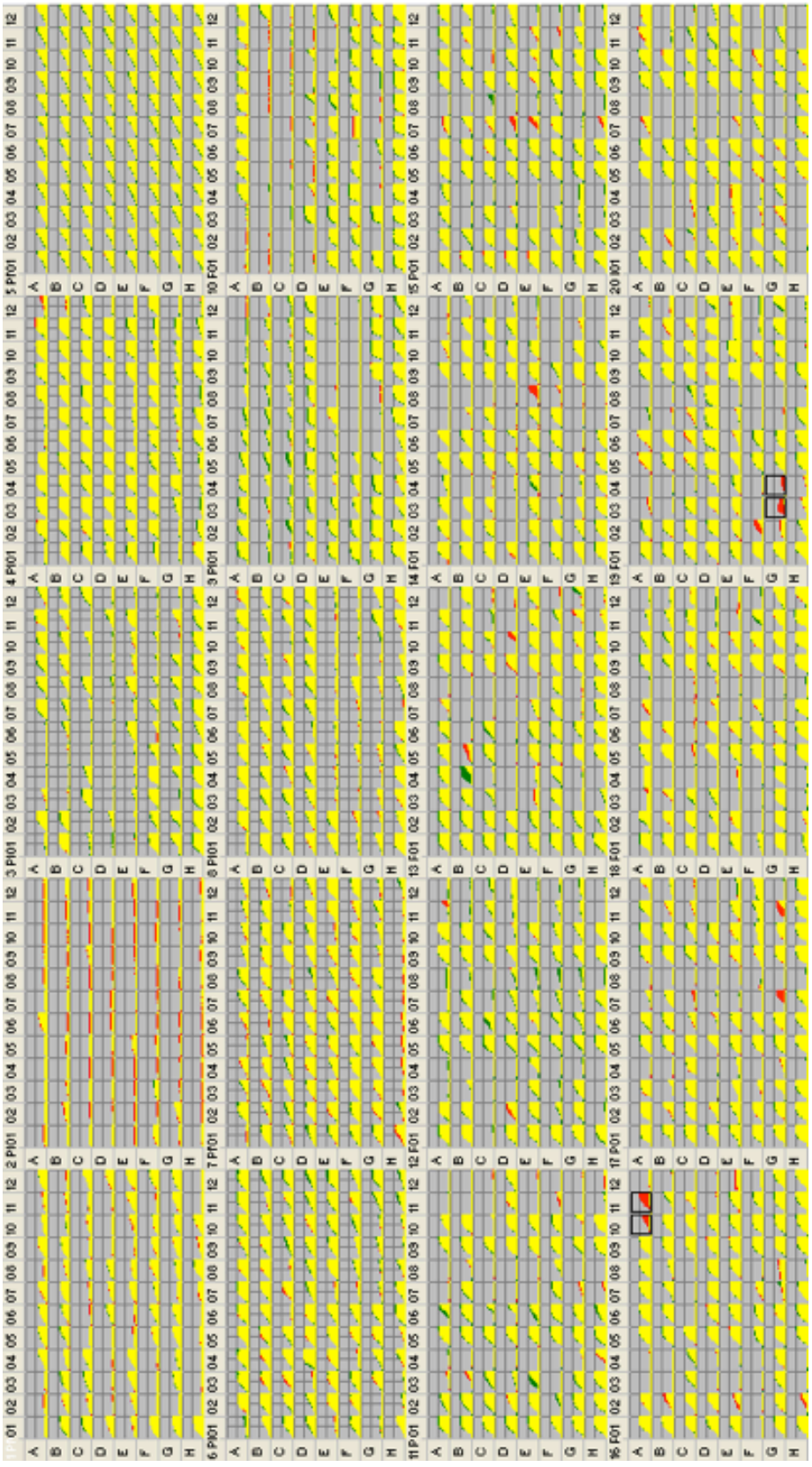
Architecture ID	No. proteins	Description
25	1	MCE, Inositol_P
26	1	Permease, MCE, Tol_Tol_Ttg2, STAS_2, BolA
27	1	MCE, Bacillus_HBL
28	1	MCE, RTX
29	1	MCE, Serine_rich
30	1	MCE, MCE, MCE, MCE, PqiA, MCE, MCE, MCE
31	1	MCE, DUF3407, MCE, DUF3407
32	1	MCE, DUF3407, DUF3407
33	1	PqiA, MCE, MCE, MCE, MCE, MCE, Nol1_Nop2_Fmu, Nol1_Nop2_Fmu_2
34	1	MCE, LXG
35	1	MCE, GCN5L1
36	1	ABC_tran, MCE
37	1	NDUFA12, MCE, DUF2155
38	1	MreC, MCE, MCE

Appendix ii – Distribution of the top nine MCE domain containing architectures across bacteria

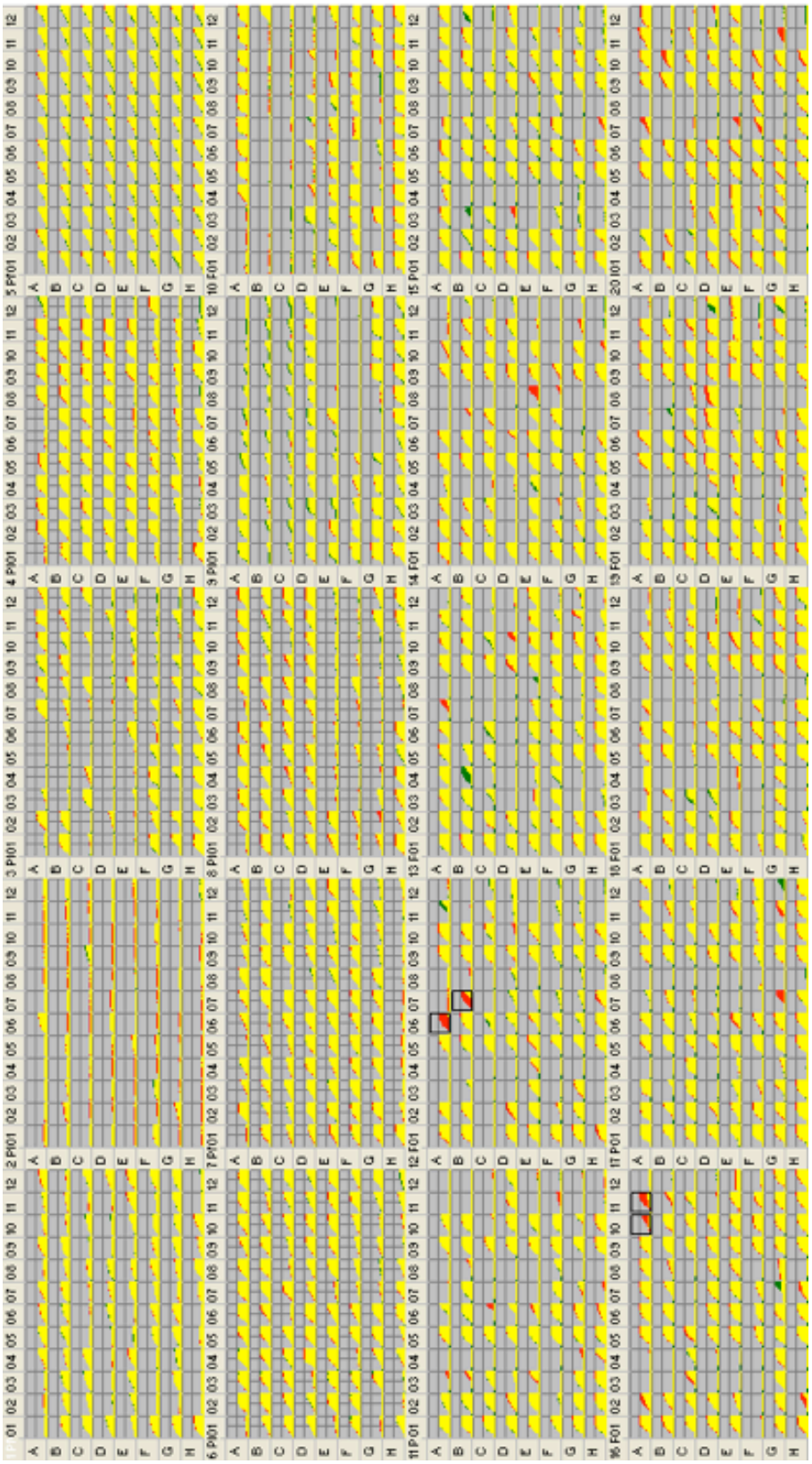


Appendix iii – Biolog data

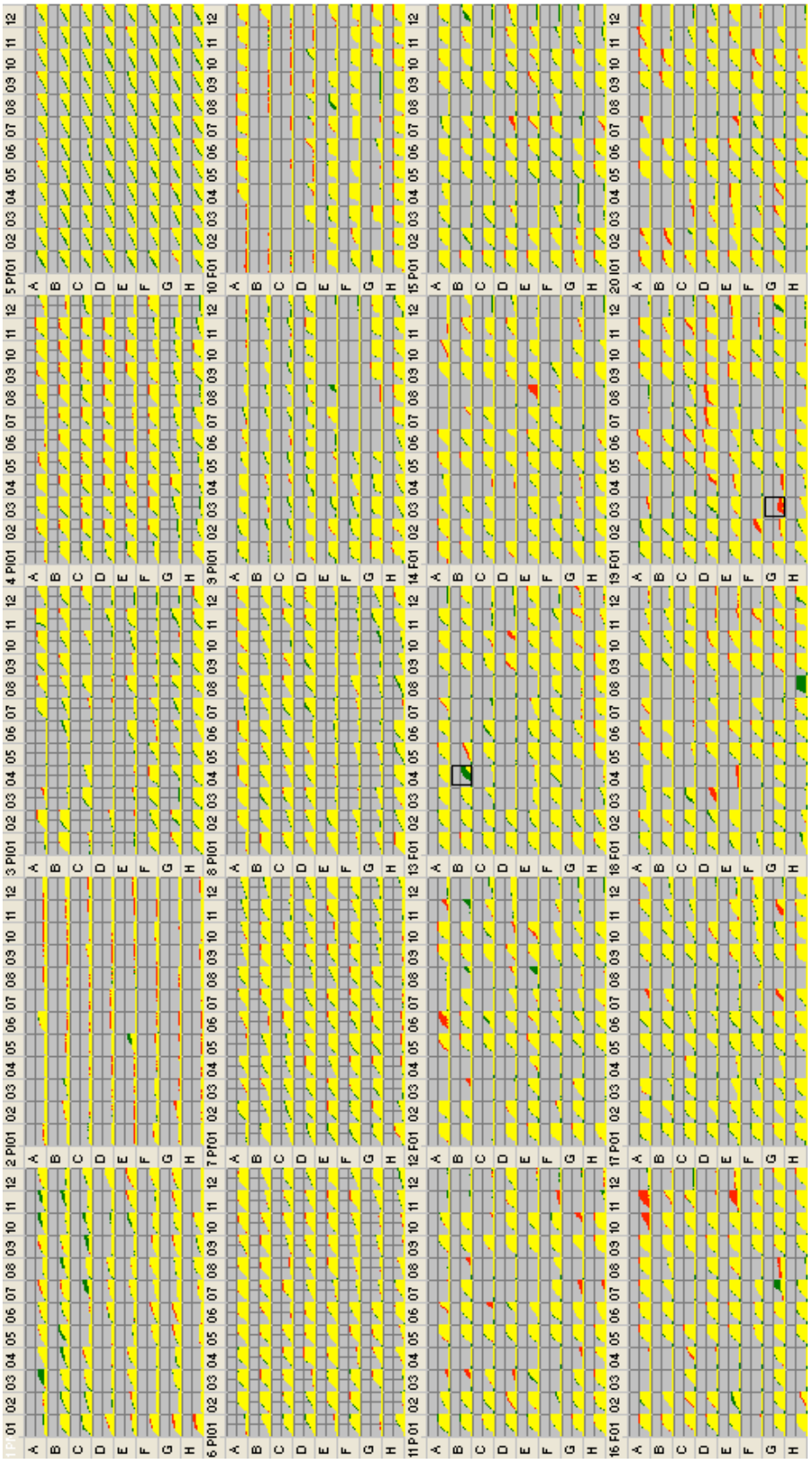
BW25113 (red) vs BW25113 Δ*pqiAB* (green)



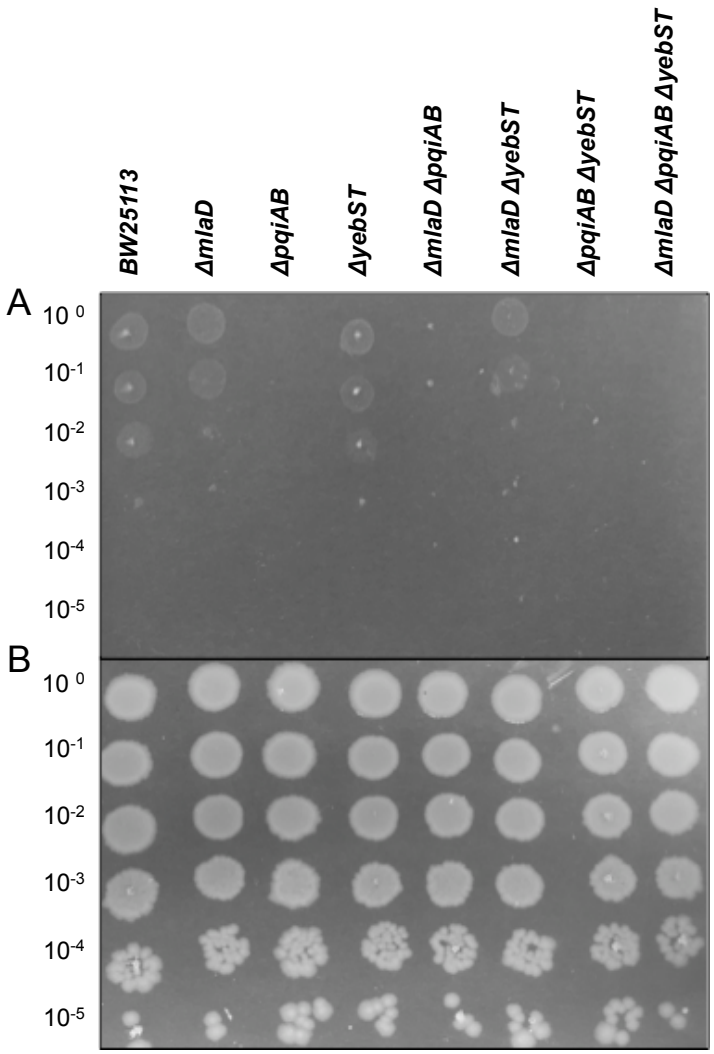
BW25113 (red) vs BW25113 Δ*yebST* (green)



BW25113 (red) vs BW25113 Δ*pqiAB* Δ*yebST* (green)



Appendix iv – Growth of the BW25113 parent and mutant strains on LA supplemented with **A.** 4% SB3-10 and **B.** 4% SB3-14.



BIBLIOGRAPHY

Abeyrathne, P.D., Daniels, C., Poon, K.K., et al. (2005) Functional characterization of WaaL, a ligase associated with linking O-antigen polysaccharide to the core of *Pseudomonas aeruginosa* lipopolysaccharide. **Journal of Bacteriology,** 187 (9): 3002-3012.

Ahmad, S., Akbar, P.K., Wiker, H.G., et al. (1999) Cloning, expression and immunological reactivity of two mammalian cell entry proteins encoded by the *mce1* operon of *Mycobacterium tuberculosis*. **Scandinavian Journal of Immunology,** 50 (5): 510-518.

Ahmad, S., El-Shazly, S., Mustafa, A.S., et al. (2004) Mammalian cell-entry proteins encoded by the mce3 operon of *Mycobacterium tuberculosis* are expressed during natural infection in humans. **Scandinavian Journal of Immunology,** 60 (4): 382-391.

Alvarez-Ordonez, A., Begley, M., Prieto, M., et al. (2011) *Salmonella* spp. survival strategies within the host gastrointestinal tract. **Microbiology,** 157 3268-3281.

Ames, G.F. (1974) Resolution of bacterial proteins by polyacrylamide gel electrophoresis on slabs. Membrane, soluble, and periplasmic fractions. **The Journal of biological chemistry,** 249 (2): 634-644.

Andersson, H. and von Heijne, G. (1993) Sec dependent and sec independent assembly of *E. coli* inner membrane proteins: the topological rules depend on chain length. **The EMBO journal,** 12 (2): 683-691.

Andersson, S.G., Zomorodipour, A., Andersson, J.O., et al. (1998) The genome sequence of Rickettsia prowazekii and the origin of mitochondria. **Nature,** 396 (6707): 133-140.

Andres, D., Baxa, U., Hanke, C., et al. (2010) Carbohydrate binding of *Salmonella* phage P22 tailspike protein and its role during host cell infection. **Biochemical Society transactions,** 38 (5): 1386-1389.

Antunes, L.C., Andersen, S.K., Menendez, A., et al. (2011) Metabolomics reveals phospholipids as important nutrient sources during *Salmonella* growth in bile in vitro and in vivo. **Journal of Bacteriology,** 193 (18): 4719-4725.

Aravind, L. and Koonin, E.V. (2000) The STAS domain - a link between anion transporters and antisigma-factor antagonists. **Current biology,** 10 (2).

Arruda, S., Bomfim, G., Knights, R., et al. (1993) Cloning of an *M. tuberculosis* DNA fragment associated with entry and survival inside cells. **Science,** 261 (5127): 1454-1457.

Audia, J.P., Webb, C.C. and Foster, J.W. (2001) Breaking through the acid barrier: an orchestrated response to proton stress by enteric bacteria. **International journal of medical microbiology,** 291 (2): 97-106.

Awai, K., Xu, C., Tamot, B., et al. (2006) A phosphatidic acid-binding protein of the chloroplast inner envelope membrane involved in lipid trafficking. **Proceedings of the National Academy of Sciences of the United States of America,** 103 (28): 10817-10822.

Baba, T., Ara, T., Hasegawa, M., et al. (2006) Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection. **Molecular systems biology,** 2 (1).

Baglieri, J., Beck, D., Vasisht, N., et al. (2012) Structure of TatA paralog, TatE, suggests a structurally homogeneous form of Tat protein translocase that transports folded proteins of differing diameter. **The Journal of biological chemistry,** 287 (10): 7335-7344.

Bell, R.M. (1975) Mutants of *Escherichia coli* defective in membrane phospholipid synthesis. Properties of wild type and Km defective sn-glycerol-3-phosphate acyltransferase activities. **The Journal of biological chemistry,** 250 (18): 7147-7152.

Benson, T.E., Walsh, C.T. and Hogle, J.M. (1997) X-ray crystal structures of the S229A mutant and wild-type MurB in the presence of the substrate enolpyruvyl-UDP-N-acetylglucosamine at 1.8-A resolution. **Biochemistry,** 36 (4): 806-811.

Benson, T.E., Walsh, C.T. and Hogle, J.M. (1996) The structure of the substrate-free form of MurB, an essential enzyme for the synthesis of bacterial cell walls. **Structure,** 4 (1): 47-54.

Berks, B.C., Palmer, T. and Sargent, F. (2003) The Tat protein translocation pathway and its role in microbial physiology. **Advances in Microbial Physiology,** 47 187-254.

Bhasin, M., Garg, A. and Raghava, G.P. (2005) PSLpred: prediction of subcellular localization of bacterial proteins. **Bioinformatics,** 21 (10): 2522-2524.

Bishop, R.E., Gibbons, H.S., Guina, T., et al. (2000) Transfer of palmitate from phospholipids to lipid A in outer membranes of gram-negative bacteria. **The EMBO journal,** 19 (19): 5071-5080.

Black, P.N. (1991) Primary sequence of the *Escherichia coli* fadL gene encoding an outer membrane protein required for long-chain fatty acid transport. **Journal of Bacteriology,** 173 (2): 435-442.

Black, P.N., DiRusso, C.C., Metzger, A.K., et al. (1992) Cloning, sequencing, and expression of the fadD gene of *Escherichia coli* encoding acyl coenzyme A synthetase. **The Journal of biological chemistry,** 267 (35): 25513-25520.

Bligh, E.G. and Dyer, W.J. (1959) A rapid method of total lipid extraction and purification. **Canadian journal of biochemistry and physiology,** 37 (8): 911-917.

Bonen, L. and Doolittle, W.F. (1975) On the prokaryotic nature of red algal chloroplasts. **Proceedings of the National Academy of Sciences of the United States of America,** 72 (6): 2310-2314.

Bos, M.P. and Tommassen, J. (2004) Biogenesis of the Gram-negative bacterial outer membrane. **Current opinion in microbiology,** 7 (6): 610-616.

Brandtzaeg, P. (1989) Overview of the mucosal immune system. **Current topics in microbiology and immunology,** 146 13-25.

Braun, V. (1973) Molecular organization of the rigid layer and the cell wall of *Escherichia coli*. **The Journal of infectious diseases,** 128 9-16.

Braun, V. and Rehn, K. (1969) Chemical characterization, spatial distribution and function of a lipoprotein (murein-lipoprotein) of the *E. coli* cell wall. The specific effect of trypsin on the membrane structure. **European journal of biochemistry,** 10 (3): 426-438.

Brown, E.D., Vivas, E.I., Walsh, C.T., et al. (1995) MurA (MurZ), the enzyme that catalyzes the first committed step in peptidoglycan biosynthesis, is essential in *Escherichia coli*. **Journal of Bacteriology,** 177 (14): 4194-4197.

Broz, P., Ohlson, M.B. and Monack, D.M. (2012) Innate immune response to *Salmonella* typhimurium, a model enteric pathogen. **Gut microbes,** 3 (2): 62-70.

Brundage, L., Hendrick, J.P., Schiebel, E., et al. (1990) The purified *E. coli* integral membrane protein SecY/E is sufficient for reconstitution of SecA-dependent precursor protein translocation. **Cell,** 62 (4): 649-657.

Buchan, D.W., Minneci, F., Nugent, T.C., et al. (2013) Scalable web services for the PSIPRED Protein Analysis Workbench. **Nucleic acids research,** 41 (Web Server issue): W349-57.

Campbell, J.W. and Cronan, J.E.,Jr (2002) The enigmatic *Escherichia coli fadE* gene is yafH. **Journal of Bacteriology,** 184 (13): 3759-3764.

Campbell, J.W., Morgan-Kiss, R.M. and Cronan, J.E.,Jr (2003) A new *Escherichia coli* metabolic competency: growth on fatty acids by a novel anaerobic beta-oxidation pathway. **Molecular microbiology,** 47 (3): 793-805.

Cantrell, S.A., Leavell, M.D., Marjanovic, O., et al. (2013) Free mycolic acid accumulation in the cell wall of the *mce1* operon mutant strain of Mycobacterium tuberculosis. **Journal of microbiology (Seoul, Korea),** 51 (5): 619-626.

Casali, N. and Riley, L.W. (2007) A phylogenomic analysis of the Actinomycetales *mce* operons. **BMC genomics,** 8 60.

Caspi, R., Billington, R., Ferrer, L., et al. (2016) The MetaCyc database of metabolic pathways and enzymes and the BioCyc collection of pathway/genome databases. **Nucleic acids research,** 44 (D1): D471-80.

Chitale, S., Ehrt, S., Kawamura, I., et al. (2001) Recombinant Mycobacterium tuberculosis protein associated with mammalian cell entry. **Cellular microbiology,** 3 (4): 247-254.

Choi, K.H., Heath, R.J. and Rock, C.O. (2000) beta-ketoacyl-acyl carrier protein synthase III (FabH) is a determining factor in branched-chain fatty acid biosynthesis. **Journal of Bacteriology,** 182 (2): 365-370.

Chong, Z.S., Woo, W.F. and Chng, S.S. (2015) Osmoporin OmpC forms a complex with MlaA to maintain outer membrane lipid asymmetry in *Escherichia coli*. **Molecular microbiology,** 98 (6): 1133-1146.

Clark, L.C., Seipke, R.F., Prieto, P., et al. (2013) Mammalian cell entry genes in *Streptomyces* may provide clues to the evolution of bacterial virulence. **Scientific reports,** 3 1109.

cmdr website (2000) [Online]. Available from: http://cmdr.ubc.ca/bobh/methods/OUTERMEMBRANEPREPARATIONONESTEP.html [Accessed June 2015].

Coburn, B., Grassl, G.A. and Finlay, B.B. (2007) *Salmonella*, the host and disease: a brief review. **Immunology and cell biology,** 85 (2): 112-118.

Cole, S.T., Brosch, R., Parkhill, J., et al. (1998) Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence. **Nature,** 393 (6685): 537-544.

Cowan, S.W., Schirmer, T., Rummel, G., et al. (1992) Crystal structures explain functional properties of two *E. coli* porins. **Nature,** 358 (6389): 727-733.

Cristobal, S., de Gier, J.W., Nielsen, H., et al. (1999) Competition between Sec- and TAT-dependent protein translocation in *Escherichia coli*. **The EMBO journal,** 18 (11): 2982-2990.

Cronan, J.E.,Jr and Waldrop, G.L. (2002) Multi-subunit acetyl-CoA carboxylases. **Progress in lipid research,** 41 (5): 407-435.

Cronan, J.E. and Vagelos, P.R. (1972) Metabolism and function of the membrane phospholipids of *Escherichia coli*. **Biochimica et biophysica acta,** 265 (1): 25-60.

Cunningham, F., Amode, M.R., Barrell, D., et al. (2015) Ensembl 2015. **Nucleic acids research,** 43 (Database issue): D662-9.

Dalebroux, Z.D., Edrozo, M.B., Pfuetzner, R.A., et al. (2015) Delivery of cardiolipins to the *Salmonella* outer membrane is necessary for survival within host tissues and virulence. **Cell host & microbe,** 17 (4): 441-451.

Datsenko, K.A. and Wanner, B.L. (2000) One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. **Proceedings of the National Academy of Sciences of the United States of America,** 97 (12): 6640-6645.

Davis, M.S. and Cronan, J.E.,Jr (2001) Inhibition of *Escherichia coli* acetyl coenzyme A carboxylase by acyl-acyl carrier protein. **Journal of Bacteriology,** 183 (4): 1499-1503.

de Keyzer, J., van der Does, C. and Driessen, A.J. (2003) The bacterial translocase: a dynamic protein channel complex. **Cellular and molecular life sciences,** 60 (10): 2034-2052.

de la Paz Santangelo, M., Klepp, L., Nunez-Garcia, J., et al. (2009) Mce3R, a TetR-type transcriptional repressor, controls the expression of a regulon involved in lipid metabolism in *Mycobacterium tuberculosis*. **Microbiology,** 155 (Pt 7): 2245-2255.

DeChavigny, A., Heacock, P.N. and Dowhan, W. (1991) Sequence and inactivation of the pss gene of *Escherichia coli*. Phosphatidylethanolamine may not be essential for cell viability. **The Journal of biological chemistry,** 266 (8): 5323-5332.

Dekker, N. (2000) Outer-membrane phospholipase A: known structure, unknown biological function. **Molecular microbiology,** 35 (4): 711-717.

Dellaporta, S.L., Xu, A., Sagasser, S., et al. (2006) Mitochondrial genome of *Trichoplax adhaerens* supports placozoa as the basal lower metazoan phylum. **Proceedings of the National Academy of Sciences of the United States of America,** 103 (23): 8751-8756.

Dev, I.K. and Ray, P.H. (1984) Rapid assay and purification of a unique signal peptidase that processes the prolipoprotein from *Escherichia coli* B. **The Journal of biological chemistry,** 259 (17): 11114-11120.

DiRusso, C.C., Heimert, T.L. and Metzger, A.K. (1992) Characterization of FadR, a global transcriptional regulator of fatty acid metabolism in *Escherichia coli*. Interaction with the fadB promoter is prevented by long chain fatty acyl coenzyme A. **The Journal of biological chemistry,** 267 (12): 8685-8691.

Duong, F. and Wickner, W. (1997) Distinct catalytic roles of the SecYE, SecG and SecDFyajC subunits of preprotein translocase holoenzyme. **The EMBO journal,** 16 (10): 2756-2768.

Durand, E., Verger, D., Rego, A.T., et al. (2009) Structural biology of bacterial secretion systems in gram-negative pathogens--potential for new drug targets. **Infectious disorders drug targets,** 9 (5): 518-547.

El-Shazly, S., Ahmad, S., Mustafa, A.S., et al. (2007) Internalization by HeLa cells of latex beads coated with mammalian cell entry (Mce) proteins encoded by the *mce3* operon of *Mycobacterium tuberculosis*. **Journal of medical microbiology,** 56 (Pt 9): 1145-1151.

Endo, R., Ohtsubo, Y., Tsuda, M., et al. (2007) Identification and characterization of genes encoding a putative ABC-type transporter essential for utilization of gamma-hexachlorocyclohexane in *Sphingobium japonicum* UT26. **Journal of Bacteriology,** 189 (10): 3712-3720.

Eppens, E.F., Nouwen, N. and Tommassen, J. (1997) Folding of a bacterial outer membrane protein during passage through the periplasm. **The EMBO journal,** 16 (14): 4295-4301.

Erhmann, M. (2007) **The periplasm.** Washington, D.C.: ASM Press.

Erridge, C., Bennett-Guerrero, E. and Poxton, I.R. (2002) Structure and function of lipopolysaccharides. **Microbes and infection,** 4 (8): 837-851.

Fierer, J. and Guiney, D.G. (2001) Diverse virulence traits underlying different clinical outcomes of *Salmonella* infection. **The Journal of clinical investigation,** 107 (7): 775-780.

Finn, R.D., Clements, J., Arndt, W., et al. (2015) HMMER web server: 2015 update. **Nucleic acids research,** 43 (W1): W30-8.

Finn, R.D., Coggill, P., Eberhardt, R.Y., et al. (2016) The Pfam protein families database: towards a more sustainable future. **Nucleic acids research,** 44 (D1): D279-85.

Forrellad, M.A., McNeil, M., Santangelo Mde, L., et al. (2014) Role of the Mce1 transporter in the lipid homeostasis of Mycobacterium tuberculosis. **Tuberculosis,** 94 (2): 170-177.

Foster, J.W. (2000) Microbial response to acid stress In **Bacterial Stress Responses** Storz, G. and Hengge-Aronis, R. (eds.) Washington, DC: American Society for Microbiology. pp. 99-115.

Fralick, J.A. (1996) Evidence that TolC is required for functioning of the Mar/AcrAB efflux pump of *Escherichia coli*. **Journal of Bacteriology,** 178 (19): 5803-5805.

Francis, C.L., Starnbach, M.N. and Falkow, S. (1992) Morphological and cytoskeletal changes in epithelial cells occur immediately upon interaction with *Salmonella* Typhimurium grown under low-oxygen conditions. **Molecular microbiology,** 6 (21): 3077-3087.

Freinkman, E., Okuda, S., Ruiz, N., et al. (2012) Regulated assembly of the transenvelope protein complex required for lipopolysaccharide export. **Biochemistry,** 51 (24): 4800-4806.

Fujita, Y., Matsuoka, H. and Hirooka, K. (2007) Regulation of fatty acid metabolism in bacteria. **Molecular microbiology,** 66 (4): 829-839.

Galyov, E.E., Wood, M.W., Rosqvist, R., et al. (1997) A secreted effector protein of *Salmonella* Dublin is translocated into eukaryotic cells and mediates inflammation and fluid secretion in infected ileal mucosa. **Molecular microbiology,** 25 (5): 903-912.

Ganz, T. and Weiss, J. (1997) Antimicrobial peptides of phagocytes and epithelia. **Seminars in hematology,** 34 (4): 343-354.

Garcia-del Portillo, F., Foster, J.W. and Finlay, B.B. (1993) Role of acid tolerance response genes in *Salmonella* Typhimurium virulence. **Infection and immunity,** 61 (10): 4489-4492.

Gruenheid, S., Pinner, E., Desjardins, M., et al. (1997) Natural resistance to infection with intracellular pathogens: the Nramp1 protein is recruited to the membrane of the phagosome. **The Journal of experimental medicine,** 185 (4): 717-730.

Gunn, J.S., Marshall, J.M., Baker, S., et al. (2014) *Salmonella* chronic carriage: epidemiology, diagnosis, and gallbladder persistence. **Trends in microbiology,** 22 (11): 648-655.

Guo, L., Lim, K.B., Poduje, C.M., et al. (1998) Lipid A acylation and bacterial resistance against vertebrate antimicrobial peptides. **Cell,** 95 (2): 189-198.

Gupta, S.D., Gan, K., Schmid, M.B., et al. (1993) Characterization of a temperature-sensitive mutant of *Salmonella* typhimurium defective in apolipoprotein N-acyltransferase. **The Journal of biological chemistry,** 268 (22): 16551-16556.

Haile, Y., Bjune, G. and Wiker, H.G. (2002) Expression of the *mceA, esat-6 and hspX* genes in *Mycobacterium tuberculosis* and their responses to aerobic conditions and to restricted oxygen supply. **Microbiology,** 148 (Pt 12): 3881-3886.

Hall, A. (1998) Rho GTPases and the actin cytoskeleton. **Science,** 279 (5350): 509-514.

Hammes, W.P. and Neuhaus, F.C. (1974) On the mechanism of action of vancomycin: inhibition of peptidoglycan synthesis in *Gaffkya homari*. **Antimicrobial Agents and Chemotherapy,** 6 (6): 722-728.

Hancock, R.E. and McPhee, J.B. (2005) *Salmonella*'s sensor for host defense molecules. **Cell,** 122 (3): 320-322.

Hawley, C.A., Watson, C.A., Orth, K., et al. (2013) A MAM7 peptide-based inhibitor of *Staphylococcus aureus* adhesion does not interfere with in vitro host cell function. **PloS one,** 8 (11).

Hawrot, E. and Kennedy, E.P. (1978) Phospholipid composition and membrane function in phosphatidylserine decarboxylase mutants of *Escherichia coli*. **The Journal of biological chemistry,** 253 (22): 8213-8220.

Heath, R.J. and Rock, C.O. (1996) Roles of the FabA and FabZ beta-hydroxyacyl-acyl carrier protein dehydratases in *Escherichia coli* fatty acid biosynthesis. **The Journal of biological chemistry,** 271 (44): 27795-27801.

Heath, R.J. and Rock, C.O. (1995) Enoyl-acyl carrier protein reductase (FabI) plays a determinant role in completing cycles of fatty acid elongation in *Escherichia coli*. **The Journal of biological chemistry,** 270 (44): 26538-26542.

Higgins, C.F. (2001) ABC transporters: physiology, structure and mechanism-an overview. **Research in microbiology,** 152 (3-4): 205-210.

Higgins, C.F., Hiles, I.D., Salmond, G.P., et al. (1986) A family of related ATP-binding subunits coupled to many distinct biological processes in bacteria. **Nature,** 323 (6087): 448-450.

Hobb, R.I., Fields, J.A., Burns, C.M., et al. (2009) Evaluation of procedures for outer membrane isolation from *Campylobacter jejuni*. **Microbiology,** 155 (Pt 3): 979-988.

Hong, M., Gleason, Y., Wyckoff, E.E., et al. (1998) Identification of two *Shigella flexneri* chromosomal loci involved in intercellular spreading. **Infection and immunity,** 66 (10): 4700-4710.

Houben, E.N., Zarivach, R., Oudega, B., et al. (2005) Early encounters of a nascent membrane protein: specificity and timing of contacts inside and outside the ribosome. **The Journal of cell biology,** 170 (1): 27-35.

Huber, D., Jamshad, M., Hanmer, R., et al. (2016) SecA cotranslationally interacts with nascent substrate proteins in vivo. **Journal of Bacteriology.**

Huber, D., Rajagopalan, N., Preissler, S., et al. (2011) SecA interacts with ribosomes in order to facilitate posttranslational translocation in bacteria. **Molecular cell,** 41 (3): 343-353.

Hug, L.A., Baker, B.J., Anantharaman, K., et al. (2016) A new view of the tree of life. **Nature Microbiology.**

Hunter, S., Apweiler, R., Attwood, T.K., et al. (2009) InterPro: the integrative protein signature database. **Nucleic acids research,** 37 (Database issue): D211-5.

Ikeda, M., Wachi, M., Jung, H.K., et al. (1991) The *Escherichia coli mraY* gene encoding UDP-N-acetylmuramoyl-pentapeptide: undecaprenyl-phosphate phospho-N-acetyl muramoyl - pentapeptide transferase. **Journal of Bacteriology,** 173 (3): 1021-1026.

Indiveri, C., Iacobazzi, V., Giangregorio, N., et al. (1998) Bacterial overexpression, purification, and reconstitution of the carnitine/acylcarnitine carrier from rat liver mitochondria. **Biochemical and biophysical research communications,** 249 (3): 589-594.

Ingram, L.O. (1977) Changes in lipid composition of *Escherichia coli* resulting from growth with organic solvents and with food additives. **Applied and Environmental Microbiology,** 33 (5): 1233-1236.

Inouye, M., Shaw, J. and Shen, C. (1972) The assembly of a structural lipoprotein in the envelope of *Escherichia coli*. **The Journal of biological chemistry,** 247 (24): 8154-8159.

Ireton, G.C. and Stoddard, B.L. (2004) Microseed matrix screening to improve crystals of yeast cytosine deaminase. **Acta crystallographica,** 60 601-605.

Ito, Y., Kanamaru, K., Taniguchi, N., et al. (2006) A novel ligand bound ABC transporter, LolCDE, provides insights into the molecular mechanisms underlying membrane detachment of bacterial lipoproteins. **Molecular microbiology,** 62 (4): 1064-1075.

Jackowski, S., Zhang, Y.M., Price, A.C., et al. (2002) A missense mutation in the *fabB* (beta-ketoacyl-acyl carrier protein synthase I) gene confers tiolactomycin resistance to *Escherichia coli*. **Antimicrobial Agents and Chemotherapy,** 46 (5): 1246-1252.

Jackson, B.J., Bohin, J.P. and Kennedy, E.P. (1984) Biosynthesis of membrane-derived oligosaccharides: characterization of *mdoB* mutants defective in phosphoglycerol transferase I activity. **Journal of Bacteriology,** 160 (3): 976-981.

Jenke-Kodama, H., Sandmann, A., Muller, R., et al. (2005) Evolutionary implications of bacterial polyketide synthases. **Molecular biology and evolution,** 22 (10): 2027-2039.

Jones, N.C. and Osborn, M.J. (1977) Translocation of phospholipids between the outer and inner membranes of *Salmonella* typhimurium. **The Journal of biological chemistry,** 252 (20): 7405-7412.

Karow, M. and Georgopoulos, C. (1993) The essential Escherichia coli *msbA* gene, a multicopy suppressor of null mutations in the *htrB* gene, is related to the universally conserved family of ATP-dependent translocators. **Molecular microbiology,** 7 (1): 69-79.

Karp, P.D., Riley, M., Saier, M., et al. (2002) The EcoCyc Database. **Nucleic acids research,** 30 (1): 56-58.

Kaur, J. and Jain, S.K. (2012) Role of antigens and virulence factors of *Salmonella enterica* serovar Typhi in its pathogenesis. **Microbiological research,** 167 (4): 199-210.

Kawasaki, K., Ernst, R.K. and Miller, S.I. (2004) Deacylation and palmitoylation of lipid A by Salmonellae outer membrane enzymes modulate host signaling through Toll-like receptor 4. **Journal of endotoxin research,** 10 (6): 439-444.

Khandekar, S.S., Gentry, D.R., Van Aller, G.S., et al. (2001) Identification, substrate specificity, and inhibition of the *Streptococcus pneumoniae* beta-ketoacyl-acyl carrier protein synthase III (FabH). **The Journal of biological chemistry,** 276 (32): 30024-30030.

Kim, K., Lee, S., Lee, K., et al. (1998) Isolation and characterization of toluene-sensitive mutants from the toluene-resistant bacterium *Pseudomonas putida* GM73. **Journal of Bacteriology,** 180 (14): 3692-3696.

Klepp, L.I., Forrellad, M.A., Osella, A.V., et al. (2012) Impact of the deletion of the six *mce* operons in *Mycobacterium smegmatis*. **Microbes and infection,** 14 (7-8): 590-599.

Knodler, L.A. and Steele-Mortimer, O. (2003) Taking possession: biogenesis of the *Salmonella*-containing vacuole. **Traffic,** 4 (9): 587-599.

Koh, Y.S. and Roe, J.H. (1996) Dual regulation of the paraquat-inducible gene *pqi-5* by SoxS and RpoS in *Escherichia coli*. **Molecular microbiology,** 22 (1): 53-61.

Koh, Y.S. and Roe, J.H. (1995) Isolation of a novel paraquat-inducible (*pqi*) gene regulated by the soxRS locus in *Escherichia coli*. **Journal of Bacteriology,** 177 (10): 2673-2678.

Krachler, A.M., Ham, H. and Orth, K. (2012a) Turnabout is fair play: use of the bacterial Multivalent Adhesion Molecule 7 as an antimicrobial agent. **Virulence,** 3 (1): 68-71.

Krachler, A.M., Ham, H. and Orth, K. (2011) Outer membrane adhesion factor multivalent adhesion molecule 7 initiates host cell binding during infection by Gram-negative pathogens. **Proceedings of the National Academy of Sciences of the United States of America,** 108 (28): 11614-11619.

Krachler, A.M., Mende, K., Murray, C., et al. (2012b) In vitro characterization of multivalent adhesion molecule 7-based inhibition of multidrug-resistant bacteria isolated from wounded military personnel. **Virulence,** 3 (4): 389-399.

Krachler, A.M. and Orth, K. (2011) Functional characterization of the interaction between bacterial adhesin multivalent adhesion molecule 7 (MAM7) protein and its host cell ligands. **The Journal of biological chemistry,** 286 (45): 38939-38947.

Krogh, A., Larsson, B., von Heijne, G., et al. (2001) Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. **Journal of Molecular Biology,** 305 (3): 567-580.

Kumar, A., Bose, M. and Brahmachari, V. (2003) Analysis of expression profile of mammalian cell entry (*mce*) operons of *Mycobacterium tuberculosis*. **Infection and immunity,** 71 (10): 6083-6087.

Kumar, A., Chandolia, A., Chaudhry, U., et al. (2005) Comparison of mammalian cell entry operons of Mycobacteria: in silico analysis and expression profiling. **FEMS immunology and medical microbiology,** 43 (2): 185-195.

Lee, D.J., Bingle, L.E., Heurlier, K., et al. (2009) Gene doctoring: a method for recombineering in laboratory and pathogenic *Escherichia coli* strains. **BMC microbiology,** 9 252.

Li, G.W., Burkhardt, D., Gross, C., et al. (2014) Quantifying absolute protein synthesis rates reveals principles underlying allocation of cellular resources. **Cell,** 157 (3): 624-635.

Lim, J., Stones, D.H., Hawley, C.A., et al. (2014) Multivalent adhesion molecule 7 clusters act as signaling platform for host cellular GTPase activation and facilitate epithelial barrier dysfunction. **PLoS pathogens,** 10 (9).

Lima, P., Sidders, B., Morici, L., et al. (2007) Enhanced mortality despite control of lung infection in mice aerogenically infected with a *Mycobacterium tuberculosis mce1* operon mutant. **Microbes and infection,** 9 (11): 1285-1290.

Louie, K., Chen, Y.C. and Dowhan, W. (1986) Substrate-induced membrane association of phosphatidylserine synthase from *Escherichia coli*. **Journal of Bacteriology,** 165 (3): 805-812.

Lovering, A.L., Safadi, S.S. and Strynadka, N.C. (2012) Structural perspective of peptidoglycan biosynthesis and assembly. **Annual Review of Biochemistry,** 81 451-478.

Lu, Y.H., Guan, Z., Zhao, J., et al. (2011) Three phosphatidylglycerol-phosphate phosphatases in the inner membrane of *Escherichia coli*. **The Journal of biological chemistry,** 286 (7): 5506-5518.

Luirink, J. and Sinning, I. (2004) SRP-mediated protein targeting: structure and function revisited. **Biochimica et biophysica acta,** 1694 (1-3): 17-35.

Luirink, J., von Heijne, G., Houben, E., et al. (2005) Biogenesis of inner membrane proteins in *Escherichia coli*. **Annual Review of Microbiology,** 59 329-355.

Luirink, J., Yu, Z., Wagner, S., et al. (2012) Biogenesis of inner membrane proteins in *Escherichia coli*. **Biochimica et biophysica acta,** 1817 (6): 965-976.

Ma, B., Reynolds, C.M. and Raetz, C.R. (2008) Periplasmic orientation of nascent lipid A in the inner membrane of an Escherichia coli LptA mutant. **Proceedings of the National Academy of Sciences of the United States of America,** 105 (37): 13823-13828.

MacAlister, T.J., Costerton, J.W., Thompson, L., et al. (1972) Distribution of alkaline phosphatase within the periplasmic space of Gram-negative bacteria. **Journal of Bacteriology,** 111 (3): 827-832.

Magnusdottir, A., Johansson, I., Dahlgren, L.G., et al. (2009) Enabling IMAC purification of low abundance recombinant proteins from *E. coli* lysates. **Nature methods,** 6 (7): 477-478.

Mahmoud, R.Y., Stones, D.H., Li, W., et al. (2016) The Multivalent Adhesion Molecule SSO1327 plays a key role in *Shigella sonnei* pathogenesis. **Molecular microbiology,** 99 (4): 658-673.

Malinverni, J.C. and Silhavy, T.J. (2009) An ABC transport system that maintains lipid asymmetry in the Gram-negative outer membrane. **Proceedings of the National Academy of Sciences of the United States of America,** 106 (19): 8009-8014.

Malinverni, J.C., Werner, J., Kim, S., et al. (2006) YfiO stabilizes the YaeT complex and is essential for outer membrane protein assembly in *Escherichia coli*. **Molecular microbiology,** 61 (1): 151-164.

Marjanovic, O., Iavarone, A.T. and Riley, L.W. (2011) Sulfolipid accumulation in *Mycobacterium tuberculosis* disrupted in the *mce2* operon. **Journal of microbiology,** 49 (3): 441-447.

Marjanovic, O., Miyata, T., Goodridge, A., et al. (2010) *Mce2* operon mutant strain of *Mycobacterium tuberculosis* is attenuated in C57BL/6 mice. **Tuberculosis,** 90 (1): 50-56.

Matos, C.F., Di Cola, A. and Robinson, C. (2009) TatD is a central component of a Tat translocon-initiated quality control system for exported FeS proteins in *Escherichia coli*. **EMBO reports,** 10 (5): 474-479.

Matsuyama, S., Tajima, T. and Tokuda, H. (1995) A novel periplasmic carrier protein involved in the sorting and transport of *Escherichia coli* lipoproteins destined for the outer membrane. **The EMBO journal,** 14 (14): 3365-3372.

Matsuyama, S., Yokota, N. and Tokuda, H. (1997) A novel outer membrane lipoprotein, LolB (HemM), involved in the LolA (p20)-dependent localization of lipoproteins to the outer membrane of *Escherichia coli*. **The EMBO journal,** 16 (23): 6947-6955.

McGhie, E.J., Brawn, L.C., Hume, P.J., et al. (2009) Salmonella takes control: effector-driven manipulation of the host. **Current opinion in microbiology,** 12 (1): 117-124.

McGhie, E.J., Hayward, R.D. and Koronakis, V. (2001) Cooperation between actin-binding proteins of invasive *Salmonella*: SipA potentiates SipC nucleation and bundling of actin. **The EMBO journal,** 20 (9): 2131-2139.

Medzhitov, R. (2010) Inflammation 2010: new adventures of an old flame. **Cell,** 140 (6): 771-776.

Mendez-Ortiz, M.M., Hyodo, M., Hayakawa, Y., et al. (2006) Genome-wide transcriptional profile of *Escherichia coli* in response to high levels of the second messenger 3',5'-cyclic diguanylic acid. **The Journal of biological chemistry,** 281 (12): 8090-8099.

Mengin-Lecreulx, D., Texier, L., Rousseau, M., et al. (1991) The murG gene of *Escherichia coli* codes for the UDP-N-acetylglucosamine: N-acetylmuramyl-(pentapeptide) pyrophosphoryl-undecaprenol N-acetylglucosamine transferase involved in the membrane steps of peptidoglycan synthesis. **Journal of Bacteriology,** 173 (15): 4625-4636.

Miller, J.D., Bernstein, H.D. and Walter, P. (1994) Interaction of *E. coli* Ffh/4.5S ribonucleoprotein and FtsY mimics that of mammalian signal recognition particle and its receptor. **Nature,** 367 (6464): 657-659.

Miller, S.I., Kukral, A.M. and Mekalanos, J.J. (1989) A two-component regulatory system (phoP phoQ) controls *Salmonella* Typhimurium virulence. **Proceedings of the National Academy of Sciences of the United States of America,** 86 (13): 5054-5058.

Minagawa, S., Ogasawara, H., Kato, A., et al. (2003) Identification and molecular characterization of the Mg2+ stimulon of *Escherichia coli*. **Journal of Bacteriology,** 185 (13): 3696-3702.

Mitra, D., Saha, B., Das, D., et al. (2005) Correlating sequential homology of Mce1A, Mce2A, Mce3A and Mce4A with their possible functions in mammalian cell entry of Mycobacterium tuberculosis performing homology modeling. **Tuberculosis,** 85 (5-6): 337-345.

Miyadai, H., Tanaka-Masuda, K., Matsuyama, S., et al. (2004) Effects of lipoprotein overproduction on the induction of DegP (HtrA) involved in quality control in the *Escherichia coli* periplasm. **The Journal of biological chemistry,** 279 (38): 39807-39813.

Miyazaki, C., Kuroda, M., Ohta, A., et al. (1985) Genetic manipulation of membrane phospholipid composition in *Escherichia coli*: *pgsA* mutants defective in phosphatidylglycerol synthesis. **Proceedings of the National Academy of Sciences of the United States of America,** 82 (22): 7530-7534.

Mohn, W.W., van der Geize, R., Stewart, G.R., et al. (2008) The actinobacterial *mce4* locus encodes a steroid transporter. **The Journal of biological chemistry,** 283 (51): 35368-35374.

Monack, D.M., Mueller, A. and Falkow, S. (2004) Persistent bacterial infections: the interface of the pathogen and the host immune system. **Nature reviews microbiology,** 2 (9): 747-765.

Monaco, C., Tala, A., Spinosa, M.R., et al. (2006) Identification of a meningococcal L-glutamate ABC transporter operon essential for growth in low-sodium environments. **Infection and immunity,** 74 (3): 1725-1740.

Mori, H. and Ito, K. (2001) The Sec protein-translocation pathway. **Trends in microbiology,** 9 (10): 494-500.

Mullineaux, C.W., Nenninger, A., Ray, N., et al. (2006) Diffusion of green fluorescent protein in three cell environments in *Escherichia coli*. **Journal of Bacteriology,** 188 (10): 3442-3448.

Nakayama, T. and Zhang-Akiyama, Q. (2016) *pqiABC* and *yebST*, putative *mce* operons of *Escherichia coli*, encode transport pathways and contribute to membrane integrity. **Journal of Bacteriology.**

Narita, S. and Tokuda, H. (2009) Biochemical characterization of an ABC transporter LptBFGC complex required for the outer membrane sorting of lipopolysaccharides. **FEBS letters,** 583 (13): 2160-2164.

Nikaido, H. (2003) Molecular basis of bacterial outer membrane permeability revisited. **Microbiology and molecular biology reviews,** 67 (4): 593-656.

Nissen, P., Hansen, J., Ban, N., et al. (2000) The structural basis of ribosome activity in peptide bond synthesis. **Science,** 289 (5481): 920-930.

Norris, F.A., Wilson, M.P., Wallis, T.S., et al. (1998) SopB, a protein required for virulence of *Salmonella* Dublin, is an inositol phosphate phosphatase. **Proceedings of the National Academy of Sciences of the United States of America,** 95 (24): 14057-14059.

Ohl, M.E. and Miller, S.I. (2001) *Salmonella*: a model for bacterial pathogenesis. **Annual Review of Medicine,** 52 259-274.

Okuda, S., Freinkman, E. and Kahne, D. (2012) Cytoplasmic ATP hydrolysis powers transport of lipopolysaccharide across the periplasm in *E. coli*. **Science,** 338 (6111): 1214-1217.

Okuda, S. and Tokuda, H. (2011) Lipoprotein sorting in bacteria. **Annual Review of Microbiology,** 65 239-259.

Osborn, M.J. and Munson, R. (1974) Separation of the inner (cytoplasmic) and outer membranes of Gram-negative bacteria. **Methods in enzymology,** 31 642-653.

Paetzel, M., Karla, A., Strynadka, N.C., et al. (2002) Signal peptidases. **Chemical reviews,** 102 (12): 4549-4580.

Palmer, T. and Berks, B.C. (2012) The twin-arginine translocation (Tat) protein export pathway. **Nature reviews microbiology,** 10 (7): 483-496.

Pandey, A.K. and Sassetti, C.M. (2008) Mycobacterial persistence requires the utilization of host cholesterol. **Proceedings of the National Academy of Sciences of the United States of America,** 105 (11): 4376-4380.

Parsons, J.B. and Rock, C.O. (2013) Bacterial lipids: metabolism and membrane homeostasis. **Progress in lipid research,** 52 (3): 249-276.

Patel, J.C. and Galan, J.E. (2005) Manipulation of the host actin cytoskeleton by *Salmonella*-all in the name of entry. **Current opinion in microbiology,** 8 (1): 10-15.

Pearse, V.B. and Voigt, O. (2007) Field biology of placozoans (*Trichoplax*): distribution, diversity, biotic interactions. **Integrative and comparative biology,** 47 (5): 677-692.

Petersen, T.N., Brunak, S., von Heijne, G., et al. (2011) SignalP 4.0: discriminating signal peptides from transmembrane regions. **Nature methods,** 8 (10): 785-786.

Pie, S., Matsiota-Bernard, P., Truffa-Bachi, P., et al. (1996) Gamma interferon and interleukin-10 gene expression in innately susceptible and resistant mice during the early phase of *Salmonella* Typhimurium infection. **Infection and immunity,** 64 (3): 849-854.

Porwollik, S., Santiviago, C.A., Cheng, P., et al. (2014) Defined single-gene and multi-gene deletion mutant collections in *Salmonella enterica* sv Typhimurium. **PloS one,** 9 (7).

Pramanik, A., Pawar, S., Antonian, E., et al. (1979) Five different enzymatic activities are associated with the multienzyme complex of fatty acid oxidation from *Escherichia coli*. **Journal of Bacteriology,** 137 (1): 469-473.

Prinz, A., Behrens, C., Rapoport, T.A., et al. (2000) Evolutionarily conserved binding of ribosomes to the translocation channel via the large ribosomal RNA. **The EMBO journal,** 19 (8): 1900-1906.

Qiu, X., Choudhry, A.E., Janson, C.A., et al. (2005) Crystal structure and substrate specificity of the beta-ketoacyl-acyl carrier protein synthase III (FabH) from *Staphylococcus aureus*. **Protein science,** 14 (8): 2087-2094.

Raetz, C.R., Guan, Z., Ingram, B.O., et al. (2009) Discovery of new biosynthetic pathways: the lipid A story. **Journal of lipid research,** 50 S103-8.

Raetz, C.R. and Newman, K.F. (1978) Neutral lipid accumulation in the membranes of *Escherichia coli* mutants lacking diglyceride kinase. **The Journal of biological chemistry,** 253 (11): 3882-3887.

Raetz, C.R., Reynolds, C.M., Trent, M.S., et al. (2007) Lipid A modification systems in Gram-negative bacteria. **Annual Review of Biochemistry,** 76 295-329.

Raffatellu, M., Wilson, R.P., Chessa, D., et al. (2005) SipA, SopA, SopB, SopD, and SopE2 contribute to *Salmonella enterica* serotype Typhimurium invasion of epithelial cells. **Infection and immunity,** 73 (1): 146-154.

Rajagopal, S., Eis, N., Bhattacharya, M., et al. (2003) Membrane-derived oligosaccharides (MDOs) are essential for sodium dodecyl sulfate resistance in *Escherichia coli*. **FEMS microbiology letters,** 223 (1): 25-31.

Ramos-Morales, F. (2012) Impact of *Salmonella enterica* Type III Secretion System Effectors on the Eukaryotic Host Cell. **ISRN Cell Biology,** Volume 2012.

Roston, R.L., Gao, J., Murcha, M.W., et al. (2012) TGD1, -2, and -3 proteins involved in lipid trafficking form ATP-binding cassette (ABC) transporter with multiple substrate-binding proteins. **The Journal of biological chemistry,** 287 (25): 21406-21415.

Ruiz, N., Gronenberg, L.S., Kahne, D., et al. (2008) Identification of two inner-membrane proteins required for the transport of lipopolysaccharide to the outer membrane of *Escherichia coli*. **Proceedings of the National Academy of Sciences of the United States of America,** 105 (14): 5537-5542.

Saini, N.K., Sharma, M., Chandolia, A., et al. (2008) Characterization of Mce4A protein of *Mycobacterium tuberculosis*: role in invasion and survival. **BMC microbiology,** 8 200-2180-8-200.

Sanchez-Pulido, L., Devos, D., Genevrois, S., et al. (2003) POTRA: a conserved domain in the FtsQ family and a class of beta-barrel outer membrane proteins. **Trends in biochemical sciences,** 28 (10): 523-526.

Sankaran, K. and Wu, H.C. (1994) Lipid modification of bacterial prolipoprotein. Transfer of diacylglyceryl moiety from phosphatidylglycerol. **The Journal of biological chemistry,** 269 (31): 19701-19706.

Sargent, F., Bogsch, E.G., Stanley, N.R., et al. (1998) Overlapping functions of components of a bacterial Sec-independent protein export pathway. **The EMBO journal,** 17 (13): 3640-3650.

Sayers, E.W., Barrett, T., Benson, D.A., et al. (2009) Database resources of the National Center for Biotechnology Information. **Nucleic acids research,** 37 (Database issue): D5-15.

Senaratne, R.H., Sidders, B., Sequeira, P., et al. (2008) *Mycobacterium tuberculosis* strains disrupted in *mce3* and *mce4* operons are attenuated in mice. **Journal of medical microbiology,** 57 (2): 164-170.

Sham, L.T., Butler, E.K., Lebar, M.D., et al. (2014) Bacterial cell wall. MurJ is the flippase of lipid-linked precursors for peptidoglycan biogenesis. **Science,** 345 (6193): 220-222.

Shimono, N., Morici, L., Casali, N., et al. (2003) Hypervirulent mutant of *Mycobacterium tuberculosis* resulting from disruption of the *mce1* operon. **Proceedings of the National Academy of Sciences of the United States of America,** 100 (26): 15918-15923.

Signorovitch, A.Y., Buss, L.W. and Dellaporta, S.L. (2007) Comparative genomics of large mitochondria in placozoans. **PLoS genetics,** 3 (1): e13.

Silhavy, T.J., Kahne, D. and Walker, S. (2010) The bacterial cell envelope. **Cold Spring Harbor perspectives in biology,** 2 (5).

Silva, J., Leite, D., Fernandes, M., et al. (2011) *Campylobacter* spp. as a foodborne pathogen: A Review. **Frontiers in microbiology,** 2 200.

Singh, P., Katoch, V.M., Mohanty, K.K., et al. (2016) Analysis of expression profile of mce operon genes (*mce1, mce2, mce3* operon) in different Mycobacterium tuberculosis isolates at different growth phases. **The Indian journal of medical research,** 143 (4): 487-494.

Smyth, M.S. and Martin, J.H. (2000) X Ray Crystallography. **Molecular pathology,** 53 (1): 8-14.

Sperandeo, P., Deho, G. and Polissi, A. (2009) The lipopolysaccharide transport system of Gram-negative bacteria. **Biochimica et biophysica acta,** 1791 (7): 594-602.

Spiess, C., Beil, A. and Ehrmann, M. (1999) A temperature-dependent switch from chaperone to protease in a widely conserved heat shock protein. **Cell,** 97 (3): 339-347.

Stanley, N.R., Findlay, K., Berks, B.C., et al. (2001) *Escherichia coli* strains blocked in Tat-dependent protein export exhibit pleiotropic defects in the cell envelope. **Journal of Bacteriology,** 183 (1): 139-144.

Stewart, J.C. (1980) Colorimetric determination of phospholipids with ammonium ferrothiocyanate. **Analytical Biochemistry,** 104 (1): 10-14.

Sutterlin, H.A., Zhang, S. and Silhavy, T.J. (2014) Accumulation of phosphatidic acid increases vancomycin resistance in *Escherichia coli*. **Journal of Bacteriology,** 196 (18): 3214-3220.

Suzuki, M., Hara, H. and Matsumoto, K. (2002) Envelope disorder of *Escherichia coli* cells lacking phosphatidylglycerol. **Journal of Bacteriology,** 184 (19): 5418-5425.

Taboada, B., Ciria, R., Martinez-Guerrero, C.E., et al. (2012) ProOpDB: Prokaryotic Operon DataBase. **Nucleic acids research,** 40 (Database issue): D627-31.

Tan, B.K., Bogdanov, M., Zhao, J., et al. (2012) Discovery of a cardiolipin synthase utilizing phosphatidylethanolamine and phosphatidylglycerol as substrates. **Proceedings of the National Academy of Sciences of the United States of America,** 109 (41): 16504-16509.

Thomason, L.C., Costantino, N. and Court, D.L. (2007) *E. coli* genome manipulation by P1 transduction. **Current protocols in molecular biology,** Chapter 1 Unit 1.17.

Thong, S., Ercan, B., Torta, F., et al. (2016) Defining key roles for auxiliary proteins in an ABC transporter that maintains bacterial outer membrane lipid asymmetry. **eLife,** 5.

Tran, A.X., Trent, M.S. and Whitfield, C. (2008) The LptA protein of *Escherichia coli* is a periplasmic lipid A-binding protein involved in the lipopolysaccharide export pathway. **The Journal of biological chemistry,** 283 (29): 20342-20349.

Trent, M.S., Stead, C.M., Tran, A.X., et al. (2006) Diversity of endotoxin and its impact on pathogenesis. **Journal of endotoxin research,** 12 (4): 205-223.

Tsay, J.T., Oh, W., Larson, T.J., et al. (1992) Isolation and characterization of the beta-ketoacyl-acyl carrier protein synthase III gene (*fabH*) from *Escherichia coli* K-12. **The Journal of biological chemistry,** 267 (10): 6807-6814.

Uchida, Y., Casali, N., White, A., et al. (2007) Accelerated immunopathological response of mice infected with *Mycobacterium tuberculosis* disrupted in the *mce1* operon negative transcriptional regulator. **Cellular microbiology,** 9 (5): 1275-1283.

Ullers, R.S., Luirink, J., Harms, N., et al. (2004) SecB is a bona fide generalized chaperone in *Escherichia coli*. **Proceedings of the National Academy of Sciences of the United States of America,** 101 (20): 7583-7588.

UniProt Consortium (2015) UniProt: a hub for protein information. **Nucleic acids research,** 43 (Database issue): D204-12.

UniProt Consortium (2007) The Universal Protein Resource (UniProt). **Nucleic acids research,** 35 (Database issue): D193-7.

Vaara, M. (1993) Antibiotic-supersusceptible mutants of *Escherichia coli* and *Salmonella* Typhimurium. **Antimicrobial Agents and Chemotherapy,** 37 (11): 2255-2260.

Vance, R.E., Isberg, R.R. and Portnoy, D.A. (2009) Patterns of pathogenesis: discrimination of pathogenic and nonpathogenic microbes by the innate immune system. **Cell host & microbe,** 6 (1): 10-21.

Vidal, S., Tremblay, M.L., Govoni, G., et al. (1995) The Ity/Lsh/Bcg locus: natural resistance to infection with intracellular parasites is abrogated by disruption of the Nramp1 gene. **The Journal of experimental medicine,** 182 (3): 655-666.

Vollmer, W., Blanot, D. and de Pedro, M.A. (2008) Peptidoglycan structure and architecture. **FEMS microbiology reviews,** 32 (2): 149-167.

Vuorio, R. and Vaara, M. (1992) The lipid A biosynthesis mutation *lpxA2* of *Escherichia coli* results in drastic antibiotic supersusceptibility. **Antimicrobial Agents and Chemotherapy,** 36 (4): 826-829.

Waterman, S.R. and Holden, D.W. (2003) Functions and effectors of the *Salmonella* pathogenicity island 2 type III secretion system. **Cellular microbiology,** 5 (8): 501-511.

Waterman, S.R. and Small, P.L. (1998) Acid-sensitive enteric pathogens are protected from killing under extremely acidic conditions of pH 2.5 when they are inoculated onto certain solid food sources. **Applied and Environmental Microbiology,** 64 (10): 3882-3886.

Weber, H., Polen, T., Heuveling, J., et al. (2005) Genome-wide analysis of the general stress response network in *Escherichia coli*: sigmaS-dependent genes, promoters, and sigma factor selectivity. **Journal of Bacteriology,** 187 (5): 1591-1603.

Wu, T., Malinverni, J., Ruiz, N., et al. (2005) Identification of a multicomponent complex required for outer membrane biogenesis in *Escherichia coli*. **Cell,** 121 (2): 235-245.

Wu, T., McCandlish, A.C., Gronenberg, L.S., et al. (2006) Identification of a protein complex that assembles lipopolysaccharide in the outer membrane of *Escherichia coli*. **Proceedings of the National Academy of Sciences of the United States of America,** 103 (31): 11754-11759.

Xie, J., Pierce, J.G., James, R.C., et al. (2011) A redesigned vancomycin engineered for dual D-Ala-D-ala And D-Ala-D-Lac binding exhibits potent antimicrobial activity against vancomycin-resistant bacteria. **Journal of the American Chemical Society,** 133 (35): 13946-13949.

Xu, G., Li, Y., Yang, J., et al. (2008) *Mycobacterium bovis* Mce4E protein may play a role in modulating cytokine expression profile in alveolar macrophage. **The international journal of tuberculosis and lung disease,** 12 (6): 664-669.

Xu, G., Li, Y., Yang, J., et al. (2007) Effect of recombinant Mce4A protein of *Mycobacterium bovis* on expression of TNF-alpha, iNOS, IL-6, and IL-12 in bovine alveolar macrophages. **Molecular and cellular biochemistry,** 302 (1-2): 1-7.

Xue, L.J., Cao, M.M., Luan, J., et al. (2007) Mammalian cell entry protein of *Mycobacterium tuberculosis* induces the proinflammatory response in RAW 264.7 murine macrophage-like cells. **Tuberculosis,** 87 (3): 185-192.

Yokota, N., Kuroda, T., Matsuyama, S., et al. (1999) Characterization of the LolA-LolB system as the general lipoprotein localization mechanism of *Escherichia coli*. **The Journal of biological chemistry,** 274 (43): 30995-30999.

Zhang, Y.M. and Rock, C.O. (2008) Membrane lipid homeostasis in bacteria. **Nature reviews microbiology,** 6 (3): 222-233.

Zhou, Z., White, K.A., Polissi, A., et al. (1998) Function of *Escherichia coli* MsbA, an essential ABC family transporter, in lipid A and phospholipid biosynthesis. **The Journal of biological chemistry,** 273 (20): 12466-12475.