Structural and functional assessment of the *Staphylococcus aureus* multidrug transporter QacC

By

Sylvia A. Sapula

A thesis submitted for the Degree of Doctor of Philosophy

The School of Biological Sciences

Flinders University

July 2015

TABLE OF CONTENTS

ABSTRACT

Antimicrobial efflux is a common bacterial resistance mechanism performed by membrane transport proteins. Some of these transporters can extrude a number of structurally dissimilar toxic compounds including antibiotics and biocides. Amongst the six currently characterised superfamilies and families of membrane transporters, is the small multidrug resistance (SMR) family. Members of this family typically range in size from ~100 to 150 amino acids and are composed of four α-helical transmembrane segments (TMS). Due to their small size their functional unit is thought to be multimeric. Within the SMR family, the *Staphylococcus aureus* QacC multidrug efflux pump was one of the first identified and characterised. Composed of 107 amino acids, QacC extrudes quaternary ammonium compounds (QACs) and dyes such as ethidium.

Extensive analysis of both the structure and function of the QacC multidrug transporter was carried out using Cys-scanning mutagenesis, whereby residues are selectively changed to a Cys residue and then challenged with various sulphydryl reagents to evaluate exposure and accessibility. In this study, 103 QacC residues were consecutively replaced with Cys and their influence on protein topology examined by the accessibility of these residues to fluorescein-5-maleimide, a membrane-impermeable sulphydryl-specific reagent. The relative position of each residue was then mapped as either being within the TMS, when shown to be non-accessible, or part of a connecting loop when accessible. These results were compared to models of QacC generated by computer modelling, which allowed for the formation of a final topological map of QacC. As SMR proteins like QacC are believed to function as dimers, with monomers displaying an antiparallel arrangement within the membrane, whole cell accessibility studies were also carried out to determine whether this monomeric arrangement occurs in QacC.

In addition to structural analysis, the ability of each QacC mutant to extrude the fluorescent substrate ethidium was analysed as was the resistance profile to a select group of compounds. Finally, topological assessment coupled with functional analysis was used to evaluate conserved SMR family motifs found in TMS1 and

TMS4, and to assess the distribution of positively charged residues within QacC which may be contributing the orientation of the monomers with the membrane.

Of the 103 residues examined, 14 were identified to be significant for protein function and/or expression. Analysis of the conserved motifs identified residues involved in substrate binding, such as Glu14 and those possibly involved in dimerisation, such as Ser90 and Gly97. Assessment of topology using whole cells has revealed that QacC orientation of monomers within the membrane takes on an antiparallel arrangement with the same reactive residues found on both side of the membrane. Finally, computer analysis and accessibility studies allowed for the formulation of a topological map of QacC, mapping each residue to the TMS or loop region. The data obtained will shed more light into the topology of transport proteins which will further our understanding of the evolution of membrane transporters. Furthermore, the essential residues identified in this study will help us gain a better understanding into the transport mechanism of multidrug transporters and SMR proteins in particular.

DECLARATION

I certify that this thesis does not incorporate without acknowledgment any material previously submitted for a degree or diploma in any university; and that to the best of my knowledge and belief it does not contain any material previously published or written by another person except where due reference is made in the text.

Sylvia A. Sapula

ACKNOWLEDGEMENTS

First of all I would like to thank my supervisors Professor Melissa Brown and Dr Uwe Stroeher for giving me the opportunity and support to complete this PhD. Your advice and patience is greatly appreciated. I would also like to thank everyone in the Brown Lab, Mohsen, Sarah, Felise, Jenny, Betty and Hayden, for simply being around and for helping me out. I know that I was very lucky to work with people that are all exceptional. Every one of you contributed in keeping me sane and making me laugh. I'd like to say a special thanks to Uwe and Hayden for bringing a continuous flow of Sci-Fi into my life, Felise for being simply Felise, and by that I mean simply awesome, Sarah for being understanding and supportive and Jenny for never ceasing to amaze me (sometimes in strange ways). Finally I would also like to thank Mohsen, for being patient and helping me out whenever I needed it.

In addition to the Brown Lab, I would also like to thank members of the Leterme Lab, in particular Charlotte. Char baby you are amazing and a wonderful friend. I appreciate all our talks and your support throughout this process.

I'd also like to say a big thanks to my parents - Dziękuję za wszystko, jesteście naprawdę wspaniali i bardzo, bardzo was kocham. Bez was bym tego nigdy nie zrobiła (Bez Gracie też nie!!). Finally, to Travis, I would just like to say - ♋ ●□❖♏ ♲□◆ ❍□❒♏ ♦♒♋■ ⬩□❒♎⬩ ♍♋■ ♍□■❖♏♲ - which on my home planet Kepler-186f (found in the Kepler-186 system in the Cygnus constellation), means absolutely nothing – we don't use Wingdings 2. You are the most amazing human that I have ever come across, thanks for supporting me and loving me – without you I would have gone ~~back to my planet~~ insane doing this.

CHAPTER 1 - INTRODUCTION

1.1 The staphylococci

The family of staphylococci are Gram–positive bacteria comprising more than 70 species and subspecies (Prax et al., 2013). These spherical bacteria have a grape like growth characteristic and are generally facultatively anaerobic, catalyse positive and typically non–motile (Lyon and Skurray, 1987). The most prominent and important member of this genus is undoubtedly *Staphylococcus aureus*, an opportunistic pathogen capable of causing a wide variety of infections (Lindsay and Holden, 2004). Although considered to be the most pathogenic of the staphylococcal species, it is also a human commensal, colonising the skin and mucosal membranes, and is carried by a significant percentage of the population (Casey et al., 2007; Shinefield and Ruff, 2009; Wassenaar et al., 2015). As a common organism it has the ability, once the skin mucosal membranes are breached, to cause a range of infections varying in degrees of severity from minor skin infections such as pimples and boils, to the more life threatening such as bacteraemia and endocarditis (Holden et al., 2004; Lyon and Skurray, 1987). One of its success as an opportunistic pathogen stems from its capacity to adapt and develop antimicrobial resistance (Lowy, 2003; Schito, 2006). This adaptability has led to the emergence of multidrug resistant strains which have spread worldwide, in both community and healthcare environments (Deurenberg and Stobberingh, 2008).

1.1.1 Development of multidrug resistance in *Staphylococcus aureus*

Prior to the introduction of penicillin in the 1940s, the mortality rate of individuals infected with *S. aureus* was in the order of 80% (Skinner and Keefer, 1941). However, the hope that came with this discovery was short lived as penicillin resistant *S. aureus* began to appear with a few years of its introduction. By the late 1940's, as a result of its widespread use in hospitals, susceptible strains were outnumbered by penicillin resistant *S. aureus* strains (Shinefield and Ruff, 2009). To overcome the problems associated with penicillin resistance, the antibiotic

methicillin which was the first semisynthetic penicillin, was introduced in the UK in 1961 (Casey et al., 2007). Unfortunately, methicillin resistant *S. aureus* (MRSA) strains were also discovered shortly after its introduction (Jevons et al., 1961). Since its identification, strains of MRSA have spread among hospitals and eventually began to appear in the wider community (de Lencastre et al., 2007; Oliveira et al., 2002). Infections caused by antibiotic resistant *S. aureus* have become a critical worldwide threat, with community acquired MRSA (CA-MRSA) rapidly spreading throughout Australia, USA, Canada as well as Europe (Vandenesch et al., 2003). In addition to its dissemination, CA–MRSA has been reported to be particularly virulent, causing tissue-destructive infections such as fulminant, necrotising pneumonia and necrotising fasciitis, whose incidence was rare before the emergence of CA–MRSA (Chambers and Deleo, 2009). In addition to methicillin, *S. aureus* has also developed resistance against a range of other antimicrobials such as gentamicin, oxacillin, erythromycin and vancomycin, the latter being considered a “last-resort” drug (McDougal et al., 2010).

The high levels of drug resistance seen in *S. aureus* are partly due to its ability to rapidly adapt to the selective pressures exerted by humans. This includes the over and inappropriate use of antibiotics, which has helped to select for the more antibiotic resistant MRSA strains (Lowy, 2003). In addition to this, contributing to its rise in resistance is its ability to adapt and acquire mobile genetic elements such as bacteriophages, transposons, insertion sequences, plasmids and pathogenicity islands by horizontal gene transfer. These elements are defined as DNA fragments that are capable of encoding one or more resistance and/or virulence determinants and as such, contribute to the wide spectrum of *S. aureus* pathogenicity (Alibayov et al., 2014).

1.2 Antibiotic resistance in *S. aureus*

1.2.1 Mechanisms of antibiotic resistance

History has shown that *S. aureus* is able to act against each new antibiotic by acquiring at least one resistance mechanism against the drug. These mechanisms include enzymatic modification or destruction of the drug, alteration of the drug

target by mutation or enzymatic inactivation resulting in reduced affinity for the drug and extruding the drug from the cell via an efflux protein (Pantosti et al., 2007; Poole, 2002). From the aforementioned, drug efflux appears to be one of the most widespread mechanisms of resistance and has been demonstrated in both Gram-negative and Gram-positive bacteria (De Rossi et al., 2002), including medically relevant strains such as MRSA isolates (Saiful et al., 2008). Resistance by this mechanism is the focus of this study.

1.3 Membrane transport systems

Drug efflux as a means of antibiotic resistance is a commonly employed mechanism which allows bacteria such as *S. aureus* to limit the accumulation of antimicrobial compounds within the cell. This is achieved by membrane proteins (MPs) which actively efflux the drug out of the cell. While many resistance mechanisms are target specific, this is not necessarily the case for efflux pumps. There are currently numerous efflux proteins that can extrude a broad range of structurally diverse antimicrobial compounds, and as such are designated multidrug efflux proteins (Poole, 2002) as they show resistance against three or more different classes of antibiotics (Hujer et al., 2006). Bacterial efflux proteins membrane fall into six different protein superfamilies or families (Figure 1.1), namely the ATP–binding cassette (ABC) superfamily, the major facilitator superfamily (MFS), the resistance–nodulation–division (RND) family, the multidrug and toxic compound extrusion (MATE) family, the small multidrug resistance (SMR) family, and the newly discovered proteobacterial chlorhexidine efflux (PCE) protein family (Collu and Cascella, 2013; Hassan et al., 2013; Paulsen et al., 1996b).

Efflux, as a mechanism of resistance was first identified in *Escherichia coli* cells that carried a plasmid–encoded single component MFS protein that exported tetracycline (Ball et al., 1980; McMurry et al., 1980). Since then, the number of chromosomally and plasmid–encoded drug membrane transporters discovered and assessed, has continued to climb, as such, it is now believed that drug efflux pumps contribute significantly to acquired bacterial resistance. Not only do they expel a broad range of antimicrobials but, by lowering the intracellular antimicrobial

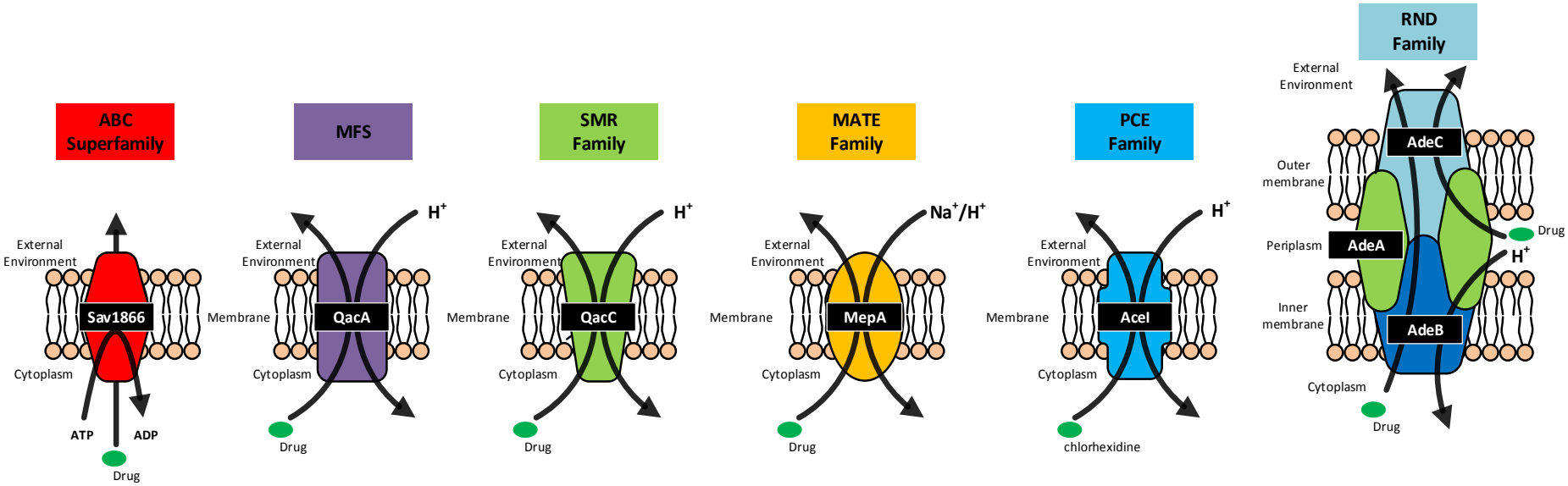

Figure 1.1: Diagrammatic representation of transporters from the six major drug transporter families

The representative transport systems are shown along with their energy-coupling mechanisms (i.e., ATP hydrolysis for the ABC superfamily export system and the use of electrochemical gradients for the remaining systems). The export systems classified within the ABC, MFS, SMR, MATE and PCE families typically transport their substrates across the cytoplasmic membrane. However, some MFS and ABC systems, as well as, most commonly, RND transporters, assemble with periplasmic and outer-membrane-bound proteins to form tripartite systems that are able to expel substrates across both the inner- and outer-membranes.

concentration, they also drive the acquisition of additional resistance mechanisms by prompting mutation accumulation (Sun et al., 2014). They can also work in synergy with other drug membrane transporters, and due to their overlapping substrate range, can act as a backup to cover the loss of functionality of other transporters, (Tal and Schuldiner, 2009) thereby maintaining resistance.

Their ability to expel a broad range of substrates and their cooperation with other mechanisms of resistance (Van Bambeke et al., 2000), has made them important therapeutic targets. Recent studies have shown that 39 out of 50 US prescription drugs have been developed to target membrane transporters (http://www.drugs.com/top200.html) and constitute 30% of the FDA approved drugs (Giorgetti, 2014; Scott et al., 2013).

Found in almost all bacterial species (Sun et al., 2014), membrane transport proteins carry out numerous vital physiological roles. These include the uptake of essential nutrients, including sugars, amino acids and peptides, expulsion of bile and host–defence molecules, as well as the extrusion of toxic compounds (Kumar and Schweizer, 2005; Paulsen et al., 1996b; Piddock, 2006). Their organisation within the prokaryotic cell membrane occurs in one of two ways. Either as single component transporters, where they catalyse the transport of substrates across the cytoplasmic membrane, or they form part of a tripartite system, containing not only the active cytoplasmic membrane transporter but also outer membrane channels and periplasmic fusion proteins. These function together to efflux substrates across both the outer and inner membrane (Li and Nikaido, 2004).

Originally described in bacterial cells, drug membrane transporters are ubiquitous and are also expressed in archaea as well as in eukaryotic cells. Their role within eukaryotic cells does not differ greatly to those mentioned above; however, in eukaryotes, they are also used to efflux neurotransmitters such as acetylcholine and play an important role in conferring resistance to anticancer drugs, discussed further in Section 1.3.1.

The recent advent of next generation sequencing has allowed many bacterial genomes to be sequenced, which has revealed a large number of genes encoding multidrug efflux proteins. Analysis of whole genome sequences of *S. aureus* strains

has revealed a number of genes which encode putative drug transport proteins (Holden et al., 2004; Kuroda et al., 2001), (listed in - http://www.membranetransport.org/), although a large number of these have yet to be experimentally confirmed (reviewed in Costa et al., 2013; Hassan et al., 2007a). These have been identified as belonging to the ABC, MFS, MATE and SMR families. Amongst these, within *S. aureus,* representing the best characterised members within each family are the ABC superfamily Sav1866 multidrug transporter (Section 1.3.1.1), the MFS QacA and QacB drug transporters (Section 1.3.2.3), the MATE family MepA efflux pump (Section 1.3.4.2) and the SMR QacC membrane transporter (Section 1.6.1), discussed below.

This review will examine each family in more detail, with a focus on substrate profile, structure and molecular mechanism. In addition, conserved sequence motifs previously identified within members of each family will also be reviewed. These will be discussed with particular reference to membrane transporters found in *S. aureus*, as mentioned above.

1.3.1 ATP-binding cassette superfamily

The ABC superfamily is characterised by proteins that can use primary energy released from ATP hydrolysis to transport substrates across the cell membrane. Consisting of both uptake and efflux transport systems, ABC transporters represent one the largest superfamilies of proteins known in both prokaryotic and eukaryotic organisms (Rees et al., 2009; Vetrivel and Subramanian, 2014). Proteins within this family have been shown to import or export a wide range of substrates, including, amongst others, amino acids, sugars and lipids (Chang, 2003; Higgins, 1992). This substrate diversity is mirrored by the diversity of the physiological roles that ABC members play in the cell. These include nutrient uptake, elimination of waste products from the cell and export of cellular components such as cell wall polysaccharides (Linton, 2007; Locher et al., 2002). Members of the ABC superfamily are found in all species, with their importance exemplified by the number of proteins certain species express. For example, in *Escherichia coli*, almost

5% of its genomic coding capacity is made up of genes encoding ABC transporters (Linton and Higgins, 1998).

Analysis of ABC membrane transporters has also shown them to be medically important, involved in cystic fibrosis, various eye diseases, liver disease and multidrug resistance in cancer cells, amongst others (Jones et al., 2009). One of the best characterised human multidrug transporters is the P-glycoprotein (P-gp) (Callaghan et al., 2014; Fojo et al., 1987; Martin et al., 2000; Pluchino et al., 2012). It is one of the most important multidrug transporter expressed in humans as it has an unusually broad substrate specificity, capable of recognising hundreds of compounds, most of which are hydrophobic, and has been described to act as a ”hydrophobic vacuum cleaner” (Aller et al., 2009; Raviv et al., 1990). It is responsible for induced drug-resistance in cancer cells as well as for modulating the pharmacokinetics of a number of drugs in human tissues (Collu and Cascella, 2013; Gottesman et al., 2002).

1.3.1.1 Structural organisation of ABC transporter proteins

Several crystal structures of ABC transporters have been solved (Dawson and Locher, 2006, 2007; Hollenstein et al., 2007; Hvorup et al., 2007; Locher et al., 2002; Oldham et al., 2007), and their structural organisation can be exemplified by two structures, the *E. coli* BtuCD vitamin B12 importer (Locher et al., 2002) and the *S. aureus* Sav1866 drug exporter (Dawson and Locher, 2006). The BtuCD transport protein consists of four individual polypeptide chains, containing four domains, including two copies of the BtuC subunit, which make up the transmembrane domain (TMD), and two copies of the BtuD subunits, which constitute the nucleotide–binding domain (NBD) (Locher et al., 2002). The Sav1866 transporter, on the other hand, is comprised of a dimer of two elongated subunits, with each subunit consisting of an amino–terminal TMD and a carboxyl-terminal NBD (Dawson and Locher, 2006). In contrast to the TMD, NBDs are homologous throughout the ABC superfamily, and contain several motifs which are characteristic of this family (Figure 1.2). These include the Walker A and B motifs, which are present in many nucleotide binding proteins, motif C which is specific to

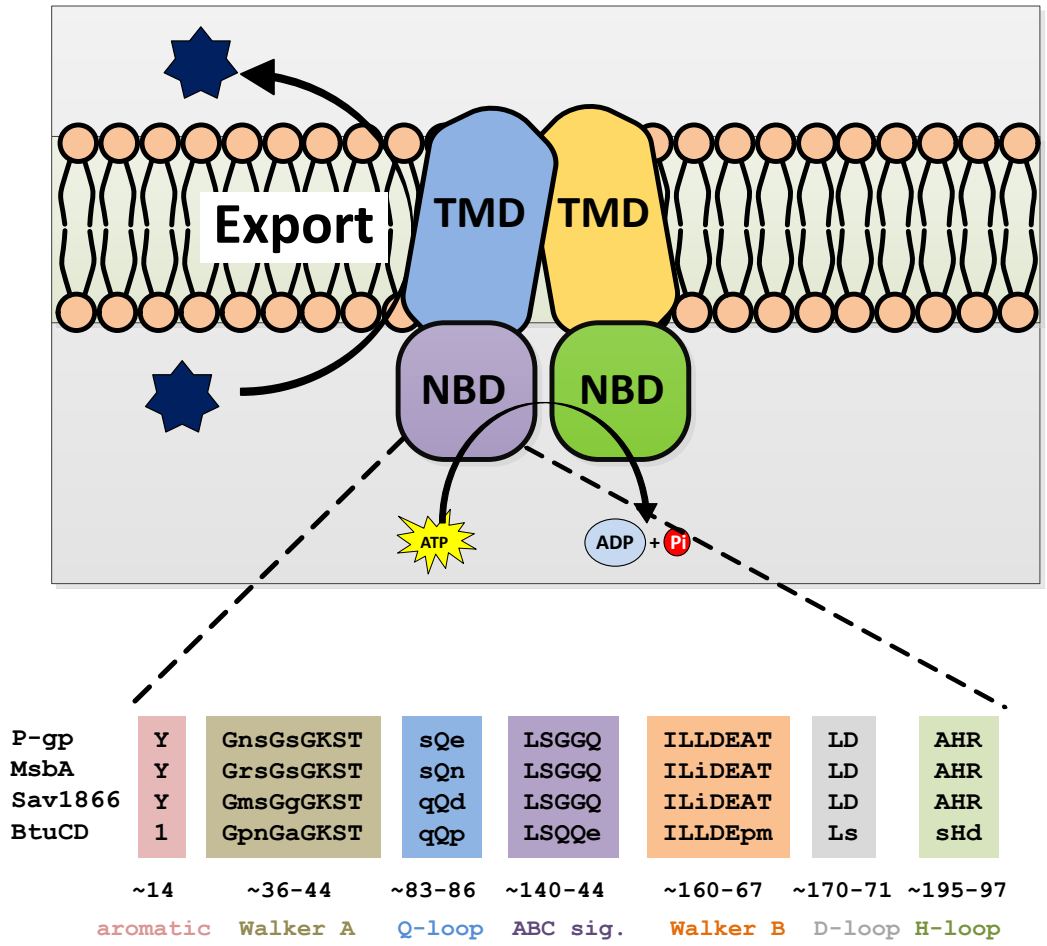

Figure 1.2: The four domains of ABC transporters

Of the four identified ABC transporter domains, two transmembrane domains (TMDs), shown in blue and yellow, are involved in substrate binding. Transport of substrates is energised by ATP binding and hydrolysis carried out by the two nucleotide binding domains (NBDs), shown in purple and green. As shown in the multiple sequence alignment of structurally characterised members of the ABC superfamily, NBDs are homologous throughout the family, and consists of seven highly conserved motifs. Unlike the NBDs, TMDs from different ABC subfamilies have been shown to vary considerably. Figure modified from Linton (2007).

the ABC superfamily (Higgins, 1992) and the stacking aromatic D, H and Q loops (Linton, 2007).

Structural studies have revealed that the TMD, containing the putative substrate binding sites, are composed of multiple α–helical transmembrane segments (TMS) that form the transmembrane channel. The TMS extend, in many cases, into the cytosol where they fold and form a physical interface with the NBDs (Jones and George, 2013). The majority of ABC membrane transporters are predicted to have six TMS per domain that appear to function as homodimers, with each transport complex containing 12 TMS in total. Although exceptions to this rule exist in the form of the *E. coli* maltose transporter MalF, which contains an N–terminal extension with two additional TMS (Froshauer and Beckwith, 1984; Froshauer et al., 1988), most ABC transporters conform to this two–times–six TMS paradigm (Biemans-Oldehinkel et al., 2006).

In addition to the above mentioned domains, importers in Gram–negative bacteria also utilise periplasmic binding proteins (PBP), whose role is to deliver the substrate to the periplasmic side of the transporter (Procko et al., 2009). The exact roles of these proteins is still unknown, however, it is thought that they impose directionality on transport (Higgins, 1992). In contrast to importers, members of the ABC superfamily that are involved in export do not associate with PBP, and obtain their substrates from the cytoplasmic side of the bacterial membrane (Locher et al., 2002).

1.3.1.2 Molecular basis of export

The exact mechanisms by which multidrug ABC transporters export their substrates are still unclear, although analysis of the crystal structures has presented a possible mechanism for coupling ATP hydrolysis to transport (Locher, 2009). Substrate extrusion is thought to begin with the interaction between the substrate and the TMDs, which initiates conformational changes in both the TMDs and the NBDs. This leads to reduced substrate binding affinity which allows for the substrate to be released and the transporter to be reset for further transport (Higgins, 2001). The nature of these conformational changes and how the TMDs

and the NBDs are coupled are still, not clearly understood. More structural, biochemical and biophysical studies need to be carried out to further elucidate the transport mechanism of ABC transporters.

1.3.1.3 The *Staphylococcal aureus* Sav1866 protein

The first ABC membrane transporter structure to be determined at a high resolution of 3.0 Å was the *S. aureus* Sav1866 (Dawson and Locher, 2006). Since then, it has frequently been used as a model for homologous human and bacterial ABC multidrug transporters, however, little is known about the function of this protein. In fact, only after the determination of the Sav1866 structure were studies carried out to determine whether or not Sav1866 could transport multiple drugs. Functional studies have now revealed that Sav1866 can transport structurally unrelated substrates such as tetraphenylphosphonium (TPP) and Hoechst 33342, in intact cells as well as from plasma membrane vesicles and proteoliposomes containing purified and functionally reconstituted protein. It was also shown that this transporter could be inhibited with a known modulator, verapamil (Velamakanni et al., 2008).

1.3.2 Major facilitator superfamily

The MFS is the largest family of secondary transporters, catalysing the transport of a diverse range of substrates (Law et al., 2008; Madej et al., 2013). As found for ABC transporters, proteins belonging to this superfamily are ubiquitously found in the membranes of all living cells (Reddy et al., 2012; Saier et al., 1999), making up ~25% of all known transport MPs in prokaryotes (Madej et al., 2013). This continually expanding superfamily is currently composed of 74 recognised MFS families which transport amino acids, peptides, vitamins and drugs, amongst others. Of these, 17 MFS families are made up of as of yet functionally uncharacterised members and are referred to as unknown major facilitators (Reddy et al., 2012). Amongst this superfamily are three different types of transporters, these are; uniporters, which are capable of transporting only one molecule of substrate at a time, symporters, which can translocate two or more substrates in

the same direction at the same time and lastly antiporters, which are capable of transporting one or more substrates but in opposite directions (Law et al., 2008).

Within the currently recognised families, three subfamilies known as DHA1 (drug:H^+ antiporter), DHA2 and DHA3 have been shown to occur most frequently in bacteria and are specific for drug efflux. Proteins belonging to the DHA1 and DHA2 subfamilies are known to efflux a wide range of structurally dissimilar compounds and have been demonstrated to play a major role in bacterial multidrug resistance (Yan, 2013), while members of the DHA3 subfamily, found only in prokaryotes, are shown to efflux antibiotics such as tetracycline and macrolides (Kumar and Schweizer, 2005). Studies have revealed that the majority of proteins within the DHA1 and DHA3 subfamily have their amino acids arranged into 12 TMS, whilst those belonging to DHA2 are comprised of 14 TMS (Hassan et al., 2006b). These TMS are connected by hydrophilic loops, with both the N and C-termini facing the cytoplasm (Saier, 2003). Members with 12 TMS are thought to have evolved from a single 2 TMS hairpin segment that triplicated, giving rise to a protein with 6 TMS which through a duplication event resulted in the formation of 12 TMS. MFS transporters with 14 TMS are believed to have obtained two centrally localised TMS as a result of an intragenic duplication event of an adjacent hairpin (Reddy et al., 2012).

Like members of the ABC superfamily, sequence analysis of MFS transporters has identified a number of superfamily and family specific sequence motifs (Ginn et al., 2000; Paulsen et al., 1996b; Saier et al., 1999). High conservation of these motifs throughout the superfamily implies that they play a role in structure and/or function. Amongst these motifs, motif C, specific to antiporters, is highly conserved between efflux proteins. Positioned in TMS5 of all antiporters it is believed that this motif may influence or specify the direction of substrate transport (Varela et al., 1995). Analysis of this motif in the MFS *S. aureus* tetracycline resistance efflux protein TetA(K), revealed that a number of conserved residues within this motif are important for tetracycline transport (Ginn et al., 2000), with Gly residues identified as conferring conformational plasticity required for drug efflux.

1.3.2.1 Molecular basis of transport

Unlike transporters belonging to the ABC superfamily, efflux by MFS members and those belonging to the RND, SMR and in some cases the MATE family, is driven by the proton motive force (PMF) (Alekshun and Levy, 2007; Piddock, 2006). Although extensively studied, the exact mechanism by which MFS drug efflux pumps bind and translocate substrates across the cell membrane is still unclear. There are many proposed theories as to the molecular mechanism behind these; however, they differ with reference to the number of binding sites and conformational changes associated with substrate translocation (Lemieux et al., 2004). Although only a small number of MFS proteins have been structurally determined (Yan, 2013), solved structures of such MFS transporters as the *E. coli* multidrug efflux protein EmrD (Yin et al., 2006) and the lactose permease LacY (Abramson et al., 2003b), have found that these MPs contain a hydrophobic interior, consistent with their function as a transporter of lipophilic substrates. Several interior residues have been shown to be bulky and aromatic and are thought to play a role in both substrate translocation as well as substrate specificity. Based on structural studies carried out on the *E. coli* EmrD, it has been speculated that drug transport and proton translocation occur at different locations. The mechanism of drug translocation is postulated to occur through a rocker-switch alternating-access model, which is coupled with H^+ antiport (Yin et al., 2006) (Figure 1.3). According to this model, MFS transporters have two main key alternating conformations, which include an inward and an outward-facing conformation and substrate translocation is assumed to be facilitated by the interconversion between these two conformations (Karpowich and Wang, 2008; Yin et al., 2006).

1.3.2.2 MFS multidrug efflux pumps in *S. aureus*

Analysis of *S. aureus* has revealed this bacterium to possess a wide array of multidrug and drug specific efflux pumps belonging to the MFS family. Chromosomally encoded drug transporters include the NorA pump (Neyfakh et al., 1993; Yoshida et al., 1990), NorB (Truong-Bolduc et al., 2005),

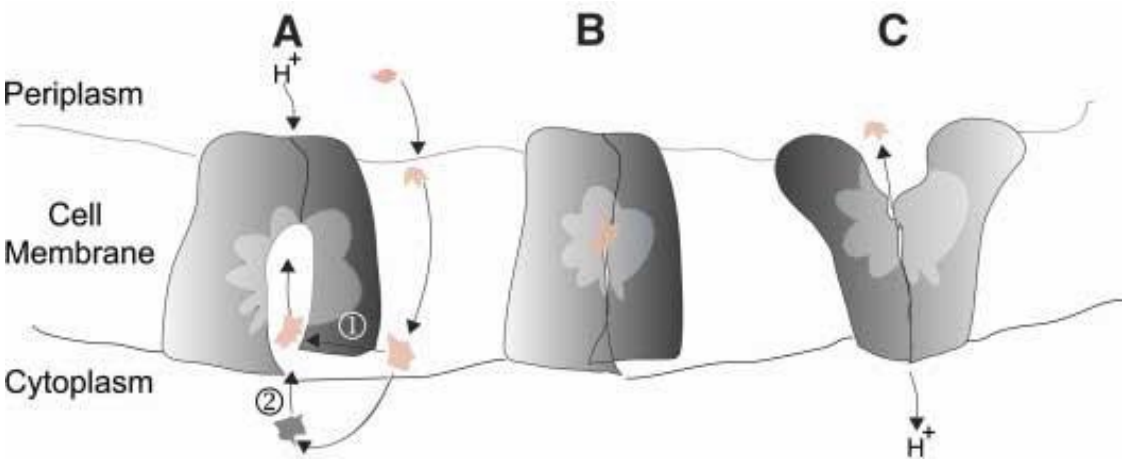

Figure 1.3: Possible transport mechanism of the *E. coli* EmrB multidrug efflux protein

(A) The substrate (shown in pink) can enter the internal cavity of the membrane transporter by two different pathways, path 1 involves going via the inner leaflet of the cell membrane whilst path 2 sees the substrate entering the binding site directly through the cytoplasm. (B) The substrate is then transported through the rocker-switch alternating-access model (Section 1.3.2.1) with the energy required to achieve this is provided by H^+ antiport. (C) Finally the substrate is transported across the lipid bilayer and released into the periplasm. Figure modified from Yin et al., (2006).

NorC (Truong-Bolduc et al., 2006), Tet38 (Truong-Bolduc et al., 2005), MdeA (Huang et al., 2004; Yamada et al., 2006b), SdrM (Yamada et al., 2006a) and the LmrS (Floyd et al., 2010) transporters. Additionally *S. aureus* can carry plasmid encoded transporters, such as QacA (Brown and Skurray, 2001; Mitchell et al., 1998; Tennent et al., 1985) and QacB (Paulsen et al., 1996a). Assessment of these transporters has shown that these MPs are capable of transporting a wide range of substrates, ranging from fluoroquinolones which are transported by nearly all the chromosomally encoded *S. aureus* MFS proteins, to quaternary ammonium compounds (QACs) such as benzalkonium chloride and dyes such as ethidium, which are substrates for all MFS MPs expressed in *S. aureus* (Costa et al., 2013).

1.3.2.3 QacA and QacB multidrug transporters – MFS prototypes

Of the *S. aureus* MFS transporters, the most widely studied multidrug efflux pumps are the chromosomally encoded NorA pump and the plasmid encoded QacA and QacB membrane transporters (Costa et al., 2013). The QacA multidrug efflux pump was the first bacterial multidrug efflux system to be discovered (Tennent et al., 1985) and since has been extensively analysed (Brown and Skurray, 2001; Hassan et al., 2007b; Hassan et al., 2008; Paulsen et al., 1996a; Xu, 2005; Xu et al., 2006). Classified within the DHA2 subfamily of the MFS, QacA has 514 amino acids that have been demonstrated to be organised into 14 TMS (Paulsen et al., 1996a) (Figure 1.4). Assessment of resistance has revealed that QacA transports a broad range of cationic and lipophilic antimicrobials, including monovalent cationic dyes such as ethidium and pyronin Y, QACs such as benzalkonium and cetylpyridinum and bivalent cations such as chlorhexidine and pentamidine (Mitchell et al., 1998; Mitchell et al., 1999). QacA has been identified in clinically relevant *S. aureus* isolates across the world, and has been shown to contribute to resistance to disinfectants widely used in medical environments and in a number of antiseptics such as eye drops (Cervinkova et al., 2013).

Studies of the QacA and its closely related MFS member QacB protein, have revealed that that an acidic residue at position 323 (Asp323) in TMS10, which is an uncharged Ala residue in QacB, affords the QacA membrane transporter the

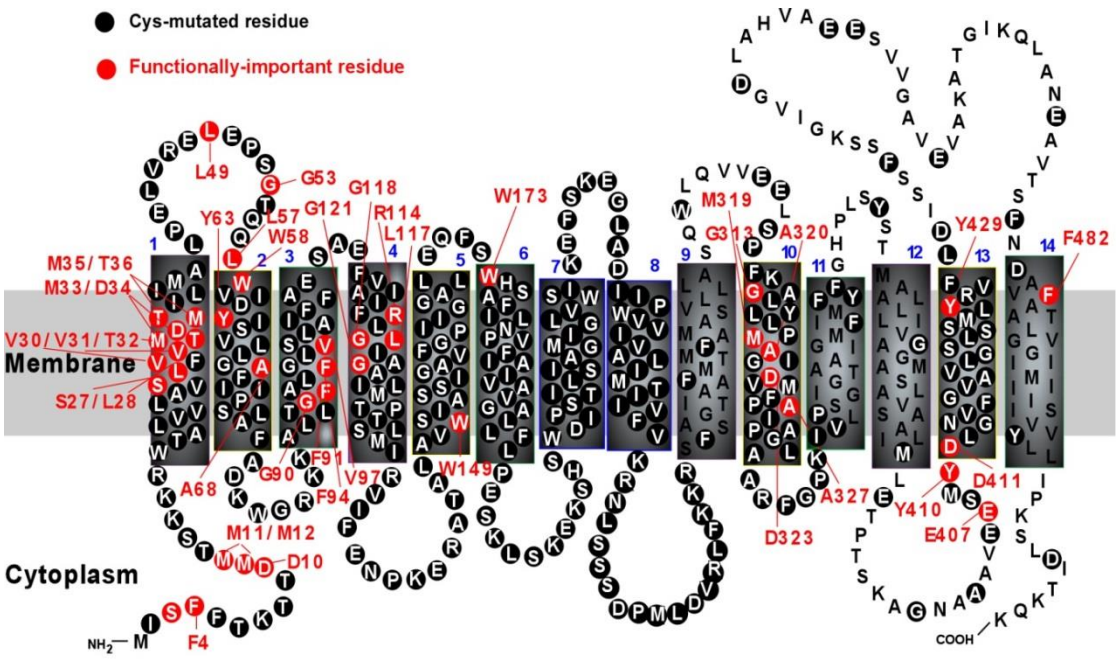

Figure 1.4: Schematic topological representation of the *S. aureus* QacA multidrug transporter

Topology of the QacA transport protein based on hydropathy analysis. The 14 TMS are indicated in dark grey, with residues found to be functionally important highlighted in red. Residues in circles show those which have been analysed by Cys-scanning mutagenesis (Hassan et al., 2006a; Hassan et al., 2007b; Xu et al., 2006). Figure kindly provided by M. H. Brown (2015).

capacity to transport a wider range of substrates that includes bivalent cations, which QacB can only poorly transport (Brown and Skurray, 2001; Paulsen et al., 1998). This implies that this region of TMS10 in QacA is involved in the binding of bivalent cations (Mitchell et al., 1999; Xu et al., 2006). Additionally another residue, Gly377, located in TMS12 may also facilitate the transport of bivalent cationic substrates. Mutagenesis studies substituting the Asp323 with Cys and Gly377 with Glu or Asp revealed that QacA double mutants Asp323Cys_Gly377Glu and Asp323Cys_Gly377Asp retained an overall capacity to confer resistance against such bivalent cations as dequalinium and chlorhexidine (Hassan et al., 2007b).

The differences between QacA and QacB, and their natural variants, have allowed for the assessment of their evolution and their ability to recognise different subsets of substrates. Initially beginning with the discovery of the prototypical QacA protein, carried on the pSK1 plasmid (Tennent et al., 1985) and the QacB protein, carried on pSK23 (Paulsen et al., 1998), these two proteins are an excellent example of fortuitous mutations which can extend the substrate profile of membrane transport proteins. As such, although most likely initially being able to expel only one class of antimicrobials, a single mutation, such as exemplified by the change to a charged residue in QacA at residue position 323, is able to render MPs capable of transporting new classes of compounds.

1.3.2.4 Regulation

Given that both QacA and QacB are encoded on plasmids they can exist as multiple copies within the bacterial cell and when needed can be expressed to levels which can protect the cell. However, high levels of expression of MPs may be deleterious to the cell (Miroux and Walker, 1996), and as such they are tightly regulated. Expression of the *qacA* and *qacB* determinants is regulated by the divergently expressed transcriptional repressor, QacR (Grkovic et al., 1998). A member of the TetR family of repressors, the QacR repressor protein functions as a pair of dimers by binding in the absence of QacA substrates to an inverted repeat (IR1) DNA sequence which overlaps the *qacA/B* transcriptional initiation sites, resulting in repression of *qacA/B* transcription (Grkovic et al., 2001). Assessment of

the substrate recognition profile of QacR has revealed that it binds a number of structurally dissimilar cationic lipophilic antiseptics, disinfectants and cytotoxins including crystal violet, rhodamine and ethidium (Schumacher and Brennan, 2003). These, as mentioned in Section 1.3.2.3, are also QacA substrates. Binding of these substrates induces a conformational change in QacR which renders the protein unable to bind IR1, triggering its release from the *qacA/B* promoter region, thus allowing *qacA/B* transcription to occur (Grkovic et al., 2003; Hassan et al., 2007a; Schumacher et al., 2001).

Much of what is known pertaining to the structure and function of QacR comes from high resolution structures of QacR bound to 10 structurally diverse compounds (Brooks et al., 2007; Murray et al., 2004; Schumacher et al., 2001). These structures revealed that QacR contains two separate but overlapping drug binding sites, which exist in one extended binding pocket. The orientation of these two pockets, referred to as the "ethidium pocket" and "rhodamine 6G binding pocket", shows that the ethidium pocket partially overlaps the rhodamine 6G pocket and lies closer to the proposed drug-binding "portal" entrance than the rhodamine 6G pocket (Schumacher et al., 2001). The volume of the extended binding pocket, when drug free, was found to be under 400 $Å^3$, however this expanded to 1100 $Å^3$ when bound to substrate (Schumacher et al., 2001), exhibiting immense flexibility. The binding pocket itself is rich in aromatic residues and contains acidic residues which play a role in substrate discrimination by affecting the positioning of the drugs within this pocket (Peters et al., 2011; Schumacher et al., 2001). The aromatic residues also play a vital role in the QacR induction mechanism, whereby drug binding triggers a conformational change of QacR by initiating a coil-to-helix transition of residues Thr89 through to Tyr93, resulting in the elongation of the C-terminus of helix 5 by a turn. This results in the expulsion of Tyr92 and Tyr93 from the interior of the protein leading to the formation of the large multidrug-binding pocket (Brooks et al., 2007; Peters et al., 2011; Schumacher et al., 2001). Upon this conformational change the relocation of the DNA-binding domain is instigated leading to the QacR protein no longer being able to bind to the IR1 operator site, resulting in de-repression of *qacA* (Peters et al., 2011). The

presence of these distinct pockets provides an explanation as to why QacR is able to bind both charged and neutral compounds as well as two different drugs simultaneously (Schumacher et al., 2004). However, the exact mechanism by which QacR is able to recognise such a wide variety of compounds is still unknown (Schumacher and Brennan, 2003).

1.3.3 Resistance nodulation-cell division family

In Gram–negative bacteria, the majority of multidrug membrane transporters are organised as a tripartite complex made up of transporter located in the inner membrane, which functions with an outer membrane channel and a periplasmic adaptor protein (Nikaido, 1994, 1996). This organisation ensures the expulsion of compounds across both the inner and outer bacterial membrane. Members of this family expel a broad range of antimicrobial compounds and confer resistance to antibiotics, dyes, detergents and organic solvents (Zgurskaya and Nikaido, 2000). Analysis of the RND family has shown that it consists of at least eight different subfamilies based on phylogeny (Tseng et al., 1999). Amongst these subfamilies are efflux systems which expel drugs, metal ions, polysaccharides, lipooligosccharides and glycolipids, with only one subfamily shown to include multidrug transporters (Paulsen et al., 2001).

Topological assessment has shown that RND transporters are comprised of 12 TMS, and contain two large periplasmic loops consisting of more than 300 amino acids located between TMS1 and 2 and TMS7 and 8. Studies of the *E. coli* AcrB and AcrD multidrug efflux pump have revealed that these loops play a role in substrate recognition and binding (Elkins and Nikaido, 2002). Through crystallographic structures of the AcrB transporter (Das et al., 2007; Murakami et al., 2002; Tornroth-Horsefield et al., 2007), the *Pseudomonas aeruginosa* MexB transporter (Sennhauser et al., 2009) and the membrane fusion protein MexA (Akama et al., 2004), as well as the *E. coli* OM channel TolC (Koronakis et al., 2000) it became possible to speculate on the interaction of these components and evaluate their role in the transport mechanism. Speculation could finally be resolved with the release of the pseudo-atomic structure of the complete AcrAB-TolC multidrug

efflux pump which provided insight into the molecular mechanism of transport and the conformational changes that accompanied it (Du et al., 2014).

1.3.3.1 RND transport mechanism

Structures of the RND proteins revealed that these pumps functioned as trimers. Crystal structures of AcrB demonstrated that each monomer within the trimer exists in three different conformations, with conformational cycling appearing to be responsible for substrate transport (Blair and Piddock, 2009; Murakami et al., 2006; Murakami et al., 2002; Seeger et al., 2008; Seeger et al., 2006). Conformational cycling, also known as the functional rotating hypothesis, was assessed in AcrB using a functional 'linked-trimer' and one inactivated monomer. Results of this study revealed that the AcrB trimer was the functional unit, as the inactivation of one monomer abolished the activity of the whole complex (Takatsuka and Nikaido, 2009).

These studies together with site-directed mutagenesis have identified a number of key AcrB residues in the transmembrane domains. These include three charged residues, Asp408 and Asp470 located in TMS4 and TMS5, and Lys940 (Blair and Piddock, 2009; Goldberg et al., 1999; Guan and Nakae, 2001), and Thr978 located in TMS11 (Takatsuka and Nikaido, 2006). These residues are thought to play a role in the proton relay network, which ultimately leads to drug efflux (Li and Nikaido, 2009). The functional rotating mechanism of transport, mentioned above, sees each of the three monomers in one of the functional states; access (loose), binding (tight) and extrusion (tight). Drug binding, predicted to occur in the periplasm, and extrusion occurs through the conformational changes which are initiated by the protonation of one of the residues in the abovementioned proton relay network within the transmembrane domain (Li and Nikaido, 2009).

The organisation of the RND tripartite system indicates that substrates are pumped out into the external medium, and not into the periplasm (Nikaido, 1998). Amphiphilic substrates are thought to diffuse into the periplasm where they are captured and extruded into the outer leaflet of the cytoplasmic membrane. However, should they continue to diffuse into the cytoplasm they are captured

from the cytosol and extruded out of the cell into the external medium (Alvarez-Ortega et al., 2013; Kim et al., 2011).

RND pumps work synergistically with the outer membrane barrier, making Gram-negative bacteria resistant to most lipophilic antibiotics. Thus, the presence of the outer membrane barrier in combination with RND efflux pumps has rendered Gram-negative bacteria "intrinsically resistant" to a large number of antimicrobial agents (Nikaido and Takatsuka, 2009). Such synergy, as also mentioned in Section 1.3., explains the importance of multidrug efflux pumps and their contribution to resistance, exemplifying the need for greater understanding of these proteins and how they function.

1.3.4 Multidrug and toxic compound extrusion family

Exhibiting a membrane topology which is similar to the MFS, members of the MATE family are ubiquitous to all kingdoms (Brown et al., 1999). Membrane transporters of the MATE family have 400 to 500 amino acids predicted to form 12 TMS. Although no apparent consensus is observed between MATE proteins, they exhibit around 40% similarity (Omote et al., 2006) with multiple sequence alignments revealing slightly conserved regions located in TMS1 and TMS7 and within six loops, two extracellular and four cytoplasmic loops. Given this symmetrical repetition of conserved regions, which is distributed between the amino and carboxyl halves of the MATE pumps, it is hypothesised that these transporters evolved from a common ancestral gene that underwent genetic duplication (Moriyama et al., 2008).

The MATE family includes the *S. aureus* MepA multidrug transporter (Kaatz et al., 2005; McAleese et al., 2005) and the *Vibrio parahaemolyticus* NorM membrane transporter (Morita et al., 1998), amongst others. Phylogenetic analysis has revealed that the MATE family is composed of three large families with 14 small subfamilies (Moriyama et al., 2008). In addition to drug efflux, members of the MATE family have also been shown to be involved in basic mechanisms which maintain homeostasis, by extruding metabolic waste products and xenobiotics in nature (Moriyama et al., 2008; Omote et al., 2006). Functional assays revealed that

these proteins also mediate resistance to cationic dyes such as ethidium, aminoglycosides and fluoroquinolones using energy stored in either the Na^+ or H^+ electrochemical gradient (Li and Nikaido, 2004; Lu et al., 2013b).

1.3.4.1 Molecular basis of transport

Although determined to be clinically significant, the transport mechanism of MATE proteins is poorly understood. However, analysis of solved structures of the *Neisseria gonorrhoeae* NorM transporter (NorM–NG), in apo form and substrate bound (He et al., 2010; Lu et al., 2013b; Tanaka et al., 2013), has revealed a multidrug binding cavity which is composed largely of negatively charged amino acids, with a limited number of hydrophobic residues. Key residues identified in drug binding were the negatively charged Asp41, Asp355 and Asp356 residues, thought to play a role in an electrostatic interaction between substrates that carry a partial positive charge. In addition Ser61, Glu284 and Ser288 were proposed to contribute to drug–charge complementation. Lastly, Phe265 was identified in an aromatic stacking interaction with a substrate (Lu et al., 2013b). Interestingly, at least three negatively charged amino acids of the NorM–NG transporter were determined to be involved in substrate neutralisation, as they were essential for precluding electroneutral or negatively charged compounds from being bound and transported, thereby ensuring specificity (Lu et al., 2013b).

Based on the above structural and functional assessment of the NorM-NG transporter, it was suggested that Na^+ triggers multidrug extrusion by inducing conformation changes within the protein. The binding of Na^+ to a drug bound transporter promotes the movement of TMS7 and TMS8 which causes the drug to disassociate from the drug binding site. The Na^+ bound, drug free transporter then switches to an inward–facing conformation, where it can then bind to another drug. It is believed that drug binding and the subsequent movement of TMS7 and TMS8 weaken Na^+ binding. As a result of this change Na^+ is released into the cytoplasm and the transporter takes on an inward–facing conformation, the then drug–bound protein returns to the outward–facing conformation where the transport cycle is completed (Lu et al., 2013b).

High resolution structural data on the *Vibrio cholera* NorM-VC provided more information regarding the transport mechanism (He et al., 2010). As with NorM–NG, a large internal cavity within the transmembrane domain was identified as being involved in substrate binding. This cavity was shown to be composed of mainly hydrophobic and/or aromatic residues. Important residues thought to be involved in cation binding were also identified in TMS7, TMS8 and TMS10 to TMS12. In particular, Glu255 and Asp371 were revealed to be critical for cation binding, as site–directed mutagenesis of the corresponding residues in the homologous NorM protein of *V. parahaemolyticus* (Glu251 and Asp367) abolished cation binding (Otsuka et al., 2005).

Although Na^+ driven, the mechanism of transport is comparable to the mechanism described for H^+ coupled transporters. In fact, structural analysis and functional assessment of DinF, a H^+-coupled MATE transporter from *Bacillus halodurans,* also identified a membrane embedded substrate binding site. It was suggested that drug–mediated DinF transport resulted from the direct competition between H^+ and the drug, and occurred through conformational inward and outward-facing changes (Lu et al., 2013a), a mechanism that has also been described for proteins belonging to the MFS (Section 1.3.2.1).

1.3.4.2 The *S. aureus* MepA multidrug transporter

The MepA protein was the first and only MATE family multidrug transporter to be discovered in *S. aureus* (Figure 1.5) (Kaatz et al., 2005). Functional analysis of this protein revealed that when overexpressed it had a broad substrate profile that included both monovalent and bivalent biocides such as ethidium and chlorhexidine, respectively, as well as fluoroquinolone antibiotics such as norfloxacin and ciprofloxacin (Kaatz et al., 2005). Functional and structural analysis revealed that MepA is regulated by the MepR repressor which is encoded immediately upstream of the *mepA*. A MarR family member, MepR is substrate responsive and binds to the promoter regions of both *mepR* and *mepA* (Kaatz et al., 2006; Kaatz et al., 2005; Schindler et al., 2015), a mechanism of repression which is comparable to the QacA /QacR system discussed in Section 1.3.2.3.

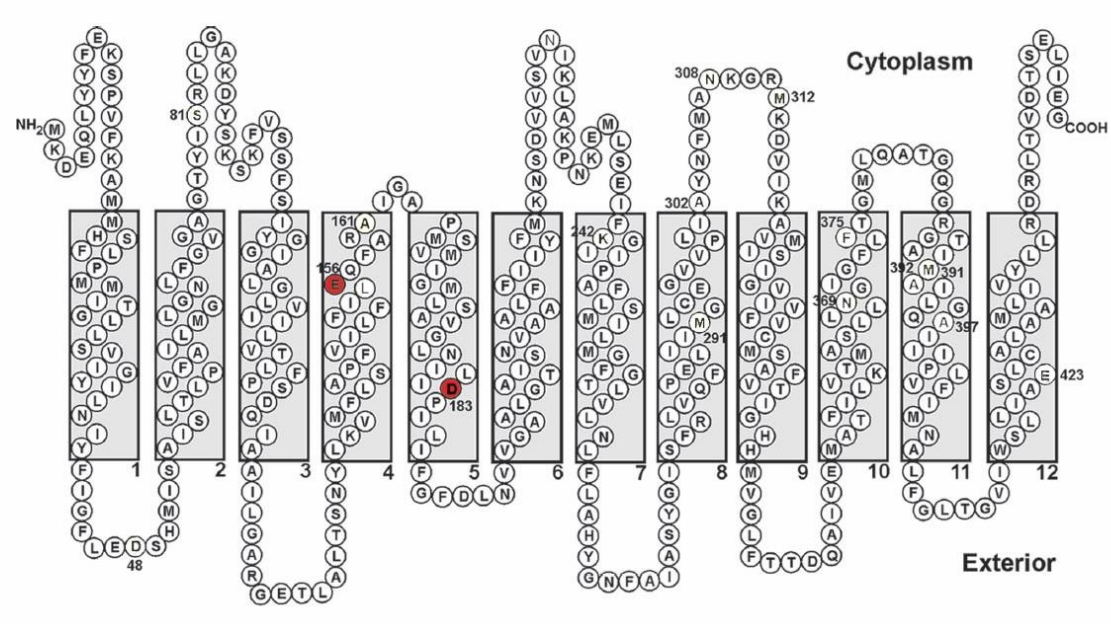

Figure 1.5: Schematic model of the *S. aureus* MepA multidrug transporter

Secondary structure of the MepA membrane transporter as predicted by hydropathy analysis. Essential residues identified in substrate binding are indicated in red. Figure modified from Schindler et al., (2013b).

Analysis of critical residues for substrate binding and/or translation, revealed that like many drug efflux pumps, negatively charged residues play a key role in the substrate translocation pathway, which in MepA are Glu156 and Asp183. However, the exact nature of their participation in this process is still to be elucidated (Banchs et al., 2014; Schindler et al., 2013b). The alternating–access mechanism, as discussed in reference to MFS proteins (Section 1.3.2.1), is also thought to be the mechanism of transport by the MepA MP, however further assessment is needed to fully elucidate this process (Schindler et al., 2013b).

1.3.5. Proteobacterial chlorhexidine efflux protein family

The proteobacterial chlorhexidine efflux (PCE) family is the newest addition to the five currently recognised families/superfamilies of drug transporters. Representing the first member to epitomize this family is the *Acinetobacter baumannii* AceI chlorhexidine transporter (Hassan et al., 2013). It was first reported following whole–genome transcriptomic analysis following treatment with subinhibitory concentrations of the biocide chlorhexidine. This revealed the overexpression of genes encoding the RND AdeAB multidrug efflux system, in addition to one other gene (A1S_2063–renamed AceI), annotated as encoding a hypothetical protein, shown to be overexpressed by more than 10–fold. Further analysis revealed this gene along with its orthologues from *P. aeruginosa* and *Burkholderia cenocepacia* to be a highly conserved component of the chlorhexidine response regulatory circuit. A Pfam database search identified that AceI contained two Bacterial Transmembrane Pair (BTP) family domains, which are found in a number of hypothetical conserved proteins shown to be mainly encoded by proteobacteria. Further analysis of the BTP family identified a number of orthologues of AceI, which when cloned and expressed in *E. coli* cells were shown to mediate various levels of chlorhexidine resistance (Hassan et al., 2013). Predicted to contain four TMS, functional assessment of the AceI transport protein revealed that a negatively charged residue, Glu15, plays an essential role in transport. Site–directed mutagenesis revealed that wild type AceI and an AceI mutant Glu15Gln were both able to directly and specifically bind chlorhexidine,

however, the Glu15Gln mutant was unable to confer chlorhexidine resistance or transport. These results indicate that Glu15, although vital for transport, does not play a role in substrate binding and may instead play a role in ion coupling.

The AceI substrate profile was analysed using Trp fluorescence assays with 11 antimicrobial compounds. As no significant fluorescence changes were observed, it was determined that AceI and members of the PCE family are substrate specific transporters. Recent studies assessing AceI homologs have revealed that like AceI many of these transport proteins conferred resistance to chlorhexidine, however, unlike AceI many also conferred resistance to other biocides like benzalkonium, dequalinium, proflavine and acriflavine (Hassan et al., 2015). These results indicate the expansion of the PCE family, which given its expanding substrate profile is now referred to as the proteobacterial antimicrobial compound efflux family of transporters.

1.4 Small multidrug resistance family

1.4.1 Diversity and evolution

Prior to the discovery of the PCE family, the SMR family of transporters represented the smallest proteins of the membrane transport system. The SMR family is divided into three subfamilies; these include the small multidrug pumps (SMR), paired small multidrug resistance pumps (PSMRs) and suppressors of *groEL* mutations (SUGs). Phylogenetic analysis of SMR proteins revealed that members of this family are subdivided into two phylogenetic clusters, with only one of these being able to catalyse drug export (Bay et al., 2008; Paulsen et al., 1996c). Of the two identified clusters, the largest one is composed of SMR proteins such as the *S. aureus* QacC and the *E. coli* EmrE, which are grouped with QacE, QacEΔ1, QacF, QacH and QacJ. The remaining cluster is primarily composed of SUG proteins. Amongst these are the *E. coli* SugE membrane transporter and the Gram–positive *Chlorobium ferrooxidans* SugE protein. As drug transport has not been demonstrated for a number of proteins within the PSMR subfamily, members of this subfamily are spread throughout the two clusters mentioned above. However, analysis of these proteins has revealed that certain members of the PSMR

subfamily share a relationship with proteins from both the SMR and SUG subfamilies, as on a phylogenetic tree generated using 101 SMR protein sequences, they are positioned towards the end of the SMR cluster, adjacent to the SUG branch (Figure 1.6), indicating that they share an evolutionary relatedness to both subfamilies. As a result of this, it has been suggested that the PSMR proteins should be designated to their own subclass designation (Bay et al., 2008; Bay and Turner, 2009).

Analysis of a multiple sequence alignment of representative SMR proteins revealed that they evolved under severe structural constraints, as seen by the absence of gaps, which are indicative of insertions and deletions within the sequence (Paulsen et al., 1996c). A search for a common SMR progenitor, assessing the diversity and evolution of the SMR family, revealed that this family underwent a high frequency of lateral gene transfer and rapid sequence divergence, which gave rise the current variety of SMR proteins (Paulsen et al., 1996c). The emergence of PSMR proteins is thought to have occurred through gene duplication events, explaining the shared relationship some of these proteins exhibit with both the SMR and SUG subfamilies. Additionally SMR family proteins display regions of conservation with other larger metabolite and multidrug membrane transporters. Fusions of TMS of SMR proteins and sequence rearrangements are believed to have contributed to the formation of the bacterial/archaeal transporter family, containing proteins with five TMS, which later led to the formation of the drug/metabolite efflux family (Bay and Turner, 2009; Saier, 2003).

Frequently identified on mobile elements such as integrons and plasmids, these proteins are typically composed of 100-150 amino acids and form four TMS (Bay and Turner, 2009). Members of this family confer low-level resistance to a variety of antimicrobial agents, including a number of QACs such as TPP and benzalkonium in addition to toxic lipophilic cations such as DNA intercalating agents. They have also demonstrated to have the ability to extrude potentially toxic metabolites such as nicotine intermediates out of the cell (Bay et al., 2008; Bay and Turner, 2009).

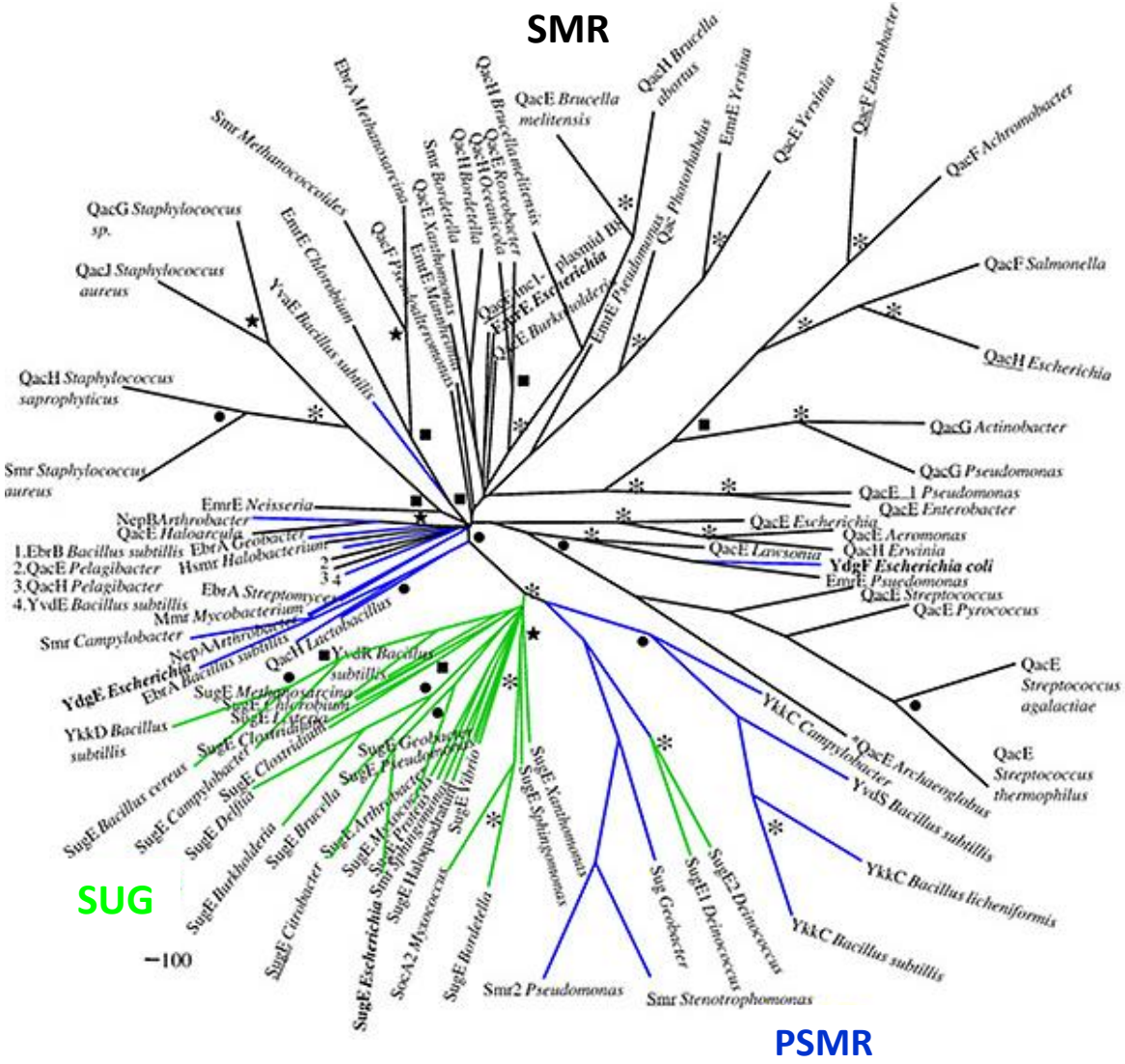

Figure 1.6: A phylogenetic tree of the SMR protein family

Based on 101 SMR protein sequences, this unrooted phylogenetic tress shows the distribution of the three SMR subfamilies, with PSMR protein braches (blue) spread throughout the SUG (green) and SMR (blue) clusters. Figure modified from Bay et al., (2008).

Transporters belonging to the SMR and the PSMRs subfamilies are characterised based on their ability to confer host resistance to a broad range of lipophilic drugs and QACs. Amongst the SMR subfamily the most highly characterised are the *E. coli* EmrE (Yerushalmi et al., 1995, 1996) and *S. aureus* QacC (Grinius et al., 1992; Littlejohn et al., 1991; Paulsen et al., 1995) multidrug resistance proteins. Amongst the PSMR class the *Bacillus subtilis* YkkCD (Jack et al., 2000) and EbrAB (Masaoka et al., 2000) membrane transporters represent those most experimentally characterised. Lastly, proteins belonging to the SUG subfamily are grouped based on their ability to phenotypically suppress the *groEL* mutation, discussed further in Section 1.4.2.1, and are thought to support cellular chaperones. Amongst these is the *E. coli* SugE membrane transporter which, when over expressed, has also been found to confer low levels of resistance against a very narrow range of compounds (Chung and Saier, 2002; Son et al., 2003).

1.4.2 Subfamilies: SUGs, PSMRs and SMRs

1.4.2.1 The SUG/SugE subfamily

Proteins belonging to the SUG subfamily are referred to as suppressors of *groEL* mutations (SUG or SugE), based on their conferred phenotype (Greener et al., 1993). The *E. coli* GroEL chaperonin, which forms part of the GroEL/GroES chaperone complex, promotes protein folding in bacteria (Radford, 2006). Unlike SMR multidrug transporters, members of the SUG family lack the capacity to transport a broad range of QAC or lipophilic dyes. Overexpression studies have, nonetheless, shown that these proteins are able to export a very narrow range of these substrates (Chung and Saier, 2002; Son et al., 2003). For SugE this is limited to antiseptics such as cetylpyridinium, cetyldimethylethyl ammonium and cetrimide (Chung and Saier, 2002). However, more recently, an *Enterobacter cloacae* SugE protein when overexpressed in an *E. coli* drug-hypersensitive strain conferred resistance to the aforementioned antiseptics in addition to TPP, benzalkonium chloride, ethidium bromide and sodium dodecyl sulphate (SDS) (He et al., 2011). In addition to these, studies assessing the mechanism of xenobiotic tributylin resistance in *Aeromonas molluscorum* Av27 identified a new SugE transporter,

Av27-SugE, which is involved in tributylin resistance. *E. coli* cells transformed with this protein were also shown to be resistant to ethidium bromide, chloramphenicol and tetracycline (Cruz et al., 2013).

Information pertaining to the relationship between the SugE protein and GroEL is sparse, however, it is believed that SugE possesses some chaperone like activity (Chung and Saier, 2002; Paulsen et al., 1996c; Saier and Paulsen, 2001). Furthermore, its overexpression and accumulation in the membrane has been demonstrated to suppress *groEL* mutations (Greener et al., 1993), suggesting that SugE proteins may play a role in the uptake of chaperone regulatory compounds (Bay et al., 2008).

As illustrated for SugE, which is comprised of four TMS (Figure 1.7) and utilises the PMF for transport, sequence analysis has identified four residues conserved specifically within the SUG subfamily; these include His24, Met39, Ile43 and Ala44. Mutagenesis studies focussing on these residues demonstrated that a His24Glu mutation resulted in hypersensitivity to more positively charged compounds indicating that this residue plays a possible role in drug binding. Mutations of the remaining three conserved residues generated a SugE derivative with an increased sensitivity to QACs such as methyltriphenylphosphonium and dyes such ethidium, methyl viologen and proflavin. The observed hypersensitivity of the SugE mutant poses the possibility that SugE functions as an importer (Son et al., 2003).

1.4.2.2 The PSMR subfamily

Found in both Gram-negative and Gram-positive bacteria, a number of PSMR proteins have been identified in *B. subtilis*; these include EbrAB, YkkCD, YvdRS and YvaDE. Of these only EbrAB and YkkCD mediated drug efflux. However, further investigation revealed that these MPs were able to extrude a wide range of antimicrobials, including cationic dyes such as ethidium bromide and crystal violet, and anionic and neutral antimicrobials such as phosphonomycin and chloramphenicol as well as streptomycin and tetracycline (Jack et al., 2000; Masaoka et al., 2000).

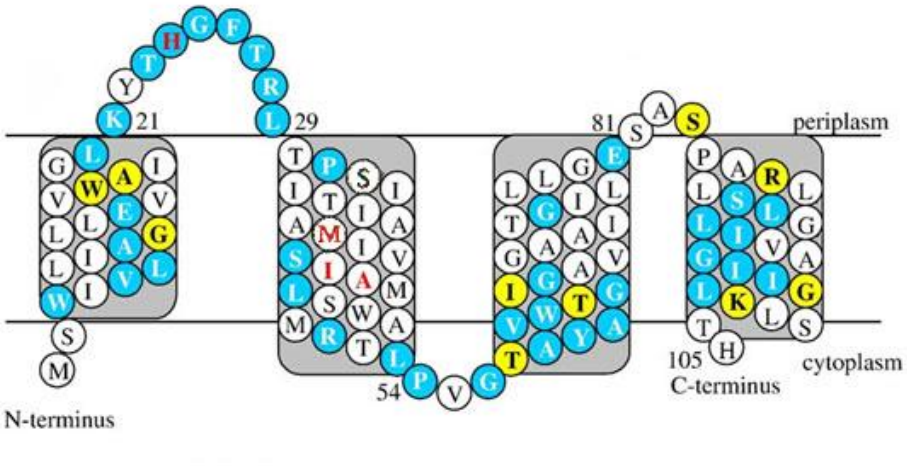

Figure 1.7: Topology of the *Citrobacter freundii* SugE transport protein

The secondary structure of the *C. freundii* SugE membrane transporter as determined by hydropathy analysis. Colours of residues are indicative of sequence conservation, with blue shaded amino acids representing conserved motifs found in SMR proteins, yellow conserved residues specific to SUGs and the highly conserved SUG residues His24, Met39, Iso43 and Ala44 shown in red. Figure modified from Turner et al., (2008).

Distinct from members of the SMR and SUG subfamilies, proteins belonging to the PSMR subfamily require the co-expression of two SMR homologues to confer host resistance to toxic metabolites and antimicrobial compounds (Bay and Turner, 2009). Genes encoding these proteins are located adjacently in a single operon at a genetic locus that is separate from other SMR proteins (Kikukawa et al., 2006). Assessment of these proteins has revealed that they function as hetero-oligomers, composed of one shorter (EbrA–105 residues), and one slightly longer protein (EbrB–117 residues) (Figure 1.8) (Jack et al., 2000; Masaoka et al., 2000). Mutants constructed of EbrA and EbrB which were lacking both the hydrophilic loops and the C–terminal regions were able to confer resistance as individual proteins. It was thus determined that the hydrophilic loops and the C–terminal regions act as constraints which lock them into their respective conformations upon the formation of the hetero-oligomer (Kikukawa et al., 2006). Nonetheless, hetero–oligomerisation appears to afford PSMR proteins the ability to confer resistance to a broader range of substrates when compared to homo–oligomeric SMR proteins (Jack et al., 2000).

Functional characterisation of the *B. subtilis* EbrAB proteins has revealed that conserved membrane embedded Glu residues, located in TMS1, are required for substrate efflux (Zhang et al., 2007). Sequence analysis revealed that as PSMRs like EbrAB function as hetero-oligomers, they carry two Glu residues, Glu15 in EbrA and Glu14 in EbrB. Following mutagenesis, it was found that both of these Glu residues are required for proper function. However, it was demonstrated that of these two, only the EbrA Glu15 is essential for substrate binding. This illustrates that although both are needed for drug efflux, their mechanistic roles in this process appears to be different (Zhang et al., 2007).

As determined for SMR multidrug efflux proteins, the negatively charged residue Glu14 is highly conserved in all SMR family members (reviewed in Bay et al., 2008). This essential residue, located within TMS1 has been revealed to be part of the substrate and proton binding site of SMRs like EmrE (Gutman et al., 2003; Yerushalmi et al., 2001; Yerushalmi and Schuldiner, 2000b) and QacC (Grinius and Goldberg, 1994), discussed further in Section 1.4.3.2.

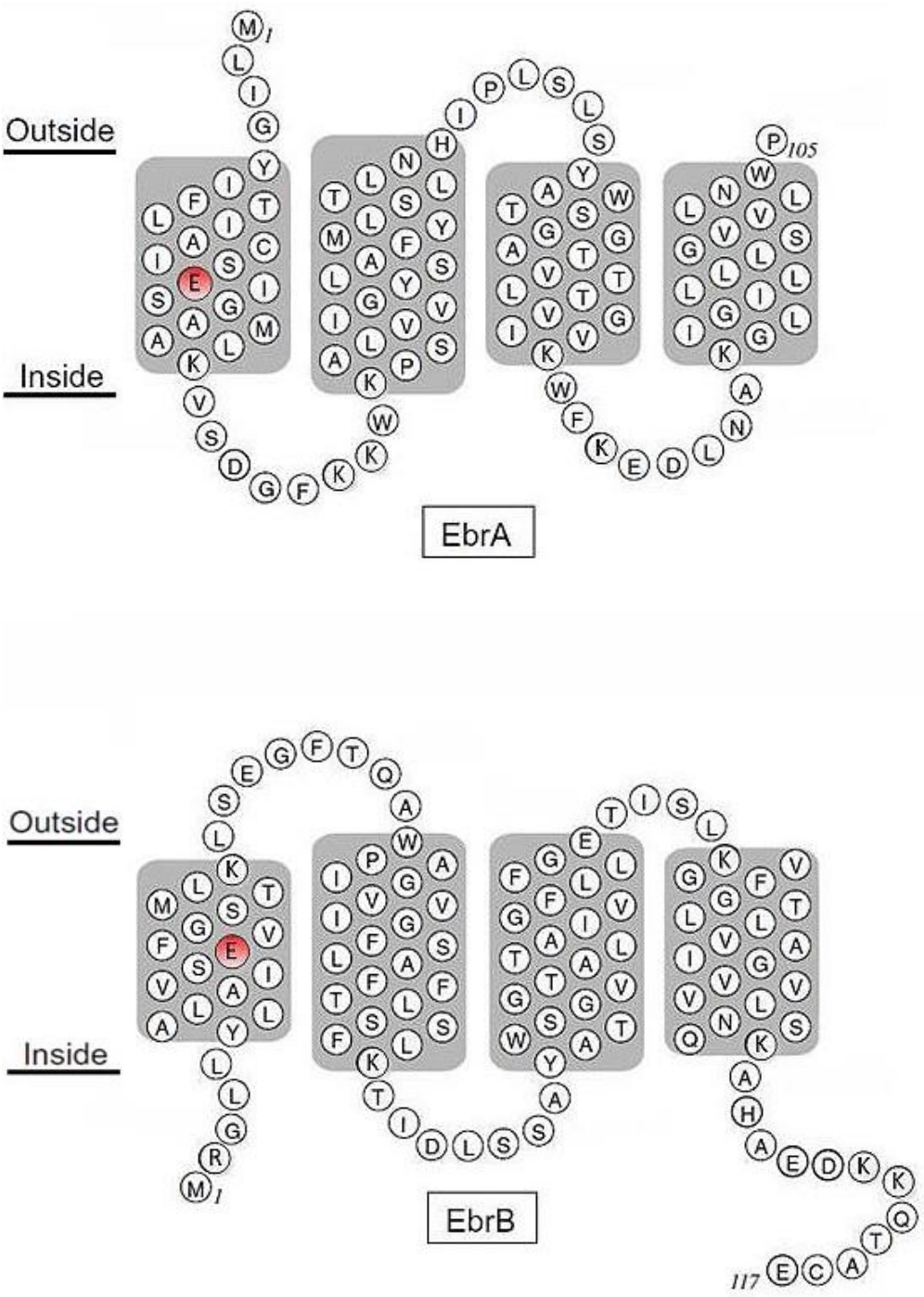

Figure 1.8: Schematic diagram of the secondary structure of the subunits of *B. subtilis* EbrAB paired small multidrug transporter

Secondary structure of the EbrAB paired small membrane transporter as predicted by hydropathy analysis. Essential residues identified in substrate binding are indicated in red. Figure modified from Nara et al., (2007).

In addition to drug transport, PSMR proteins have also been observed to function as exporters of methylamine, the end product of nicotine catabolism. Identified in *Arthrobacter nicotinovorans* a Gram–positive soil bacterium, the genes *nepAB* are part of the pAO1 plasmid encoded nicotine regulon, identified in nicotine catabolism. These encode the PSMR NepAB membrane transporter which when co-expressed exported methylamine, and the intermediate γ-*N*-methylaminobutyrate (Ganas et al., 2009; Ganas et al., 2007). Novel PSMR proteins in *Klebsiella pneumoniae* (KpnEF efflux pump) have also been revealed to play a role in capsule synthesis (Srinivasan and Rajamohan, 2013). Identified to be part of the *K. pneumoniae* cell envelope response regulon, *ΔkpnEF* mutants were shown to have increased sensitivity towards antimicrobial compounds such as benzalkonium chloride, chlorhexidine and triclosan.

1.4.3 SMR multidrug transport proteins

Proteins belonging to the SMR subfamily are characterised by their ability to confer multidrug resistance to both Gram-positive and Gram-negative bacteria, as well as Archaea (Bay et al., 2008). Members of this subfamily are proteins which are plasmid encoded, such as the *S. aureus* QacC membrane transporter (Grinius and Goldberg, 1994; Lyon and Skurray, 1987; Sasatsu et al., 1989), chromosomally encoded as exemplified by the *E. coli* EmrE protein (Morimyo et al., 1992; Purewal, 1991) or integron encoded, such as the *Klebsiella* QacE and QacEΔ1 (Paulsen et al., 1993). With more SMR proteins being identified in various bacteria, this subfamily has grown significantly to include the chromosomally encoded *A. baumannii* AbeS (Srinivasan et al., 2009) and the *Mycobacterium tuberculosis* TBsmr membrane transporter (Balganesh et al., 2012).

Amongst the Gram-negative members of this group the *E. coli* EmrE is the most extensively analysed SMR (Dutta et al., 2014; Schuldiner, 2009; Schuldiner et al., 2001b; Yerushalmi et al., 1995) and is currently referred to as the archetype of all SMR transporters (Cho et al., 2014; Ubarretxena-Belandia and Tate, 2004). Homologues of EmrE have been identified in other Gram-negative bacteria including *P. aeruginosa* (Pau-EmrE) and *Bordetella pertussis* (Bpe-EmrE) (Ninio et

al., 2001), with a sequence identity of approximately 45% and 50%, respectively (Bay et al., 2008). Whilst among Gram-positive bacteria the first to be identified and characterised is the *S. aureus* QacC multidrug transporter (Littlejohn et al., 1991; Paulsen et al., 1995; Poget et al., 2010). Besides these, the only other SMR protein to be characterised is the *M. tuberculosis* Mtu-smr transporter (De Rossi et al., 1998), which shares 32% sequence identity with QacC, and 41% with EmrE (Bay et al., 2008). Finally, representing Archaeabacteria is the *Halobacterium salinarium* Hsmr membrane transporter, which has also been both functionally and structurally characterised (Ninio and Schuldiner, 2003; Rath et al., 2006).

1.4.3.1 Conserved motifs

As found for all other membrane transport families, a multiple sequence alignment of proteins belonging to the SMR subfamily revealed a number signature motifs (Figure 1.9). These motifs lie in each of the predicted TMS with the highest residue conservation observed in TMS1 (Bay et al., 2008). Initial sequence analysis identified three signature motifs specific to the SMR subfamily. These are; motif 1; (TMS1) (W,Y)(L,I,V)XLXXA (I,G)XXE(L,I,V)(L,I,V); motif 2; (TMS1-TMS2) KX(A,S,T)(E,H) GF(S,T)(R,K)(L,F)XPS; and motif 3; (TMS3) PX(G,N)X(A,T)YAXW(A,T,S)G(L,I,V,M)G (Paulsen et al., 1996c). In these, X denotes any residue whilst those which are ambiguous are shown in parentheses. An extension of these, based on sequence alignments with newly assigned SMR transporters, includes three loop motifs, as well as an additional TMS motif, located in TMS4. Motif 4; identified to be (A,G,S)XX(G,S,L,F)(I,L,V,M,F) (I,V,L,M)XX(G,A,S)(V,I,A,L)(I,V,A,L)(V,L,I), is also referred to as the SMR dimerisation motif, as it is thought to mediate TMS4 based SMR protein dimerisation (Poulsen et al., 2009) (Section 7.3.1). In addition to TMS1, high sequence conservation was also determined in this motif, emphasising the importance of both motifs (Bay et al., 2008).

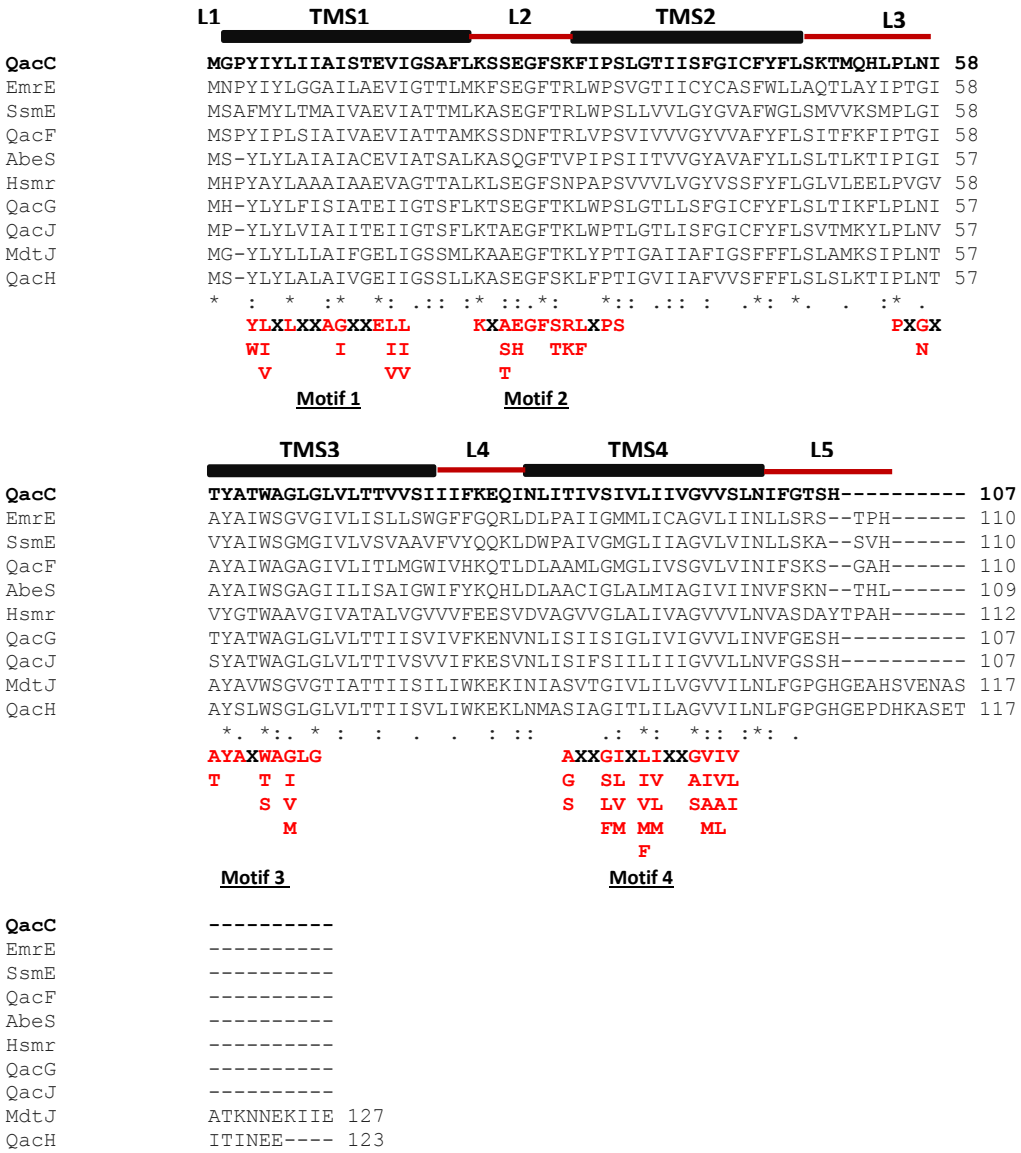

Figure 1.9: Multiple sequence alignment of QacC and related SMR efflux proteins

Multiple sequence alignment of QacC and other SMR family proteins was generated using ClustalW 2.1 (Section 2.17.1). Each of the 4 proposed QacC TMSs (Section 5.7) are indicated as black bars above the sequence, whilst loops (L) are depicted in red thin bars. Asterisks indicate fully conserved residues, colons indicate strongly similar residues and dots indicate weakly similar residues. Dashes represent gaps. SMR motif sequences are shown in red and are listed below the sequences. A black X denotes any residue. Sequences were obtained from the GenBank database by the use of the following protein accession numbers: QacC (P14319.1), EmrE (P23895.1), SsmE (BAF80121.1), QacF (YP_358817), AbeS (ACO55443.1), QacG (WP_002477903.1), QacJ (CAD55144), MdtJ (WP_015240978.1) and QacH (CAA76544). The Hsmr (B0R6K7) sequence was obtained from the Universal Protein Resource website (http://www.uniprot.org).

1.4.3.2 SMR substrate binding site–conserved residue cluster in TMS1

Sequence analysis and functional assessment of residues found in TMS1 revealed that a single negatively charged residue (QacC-Glu13, EmrE-Glu14), shown to be highly conserved in SMR members, was essential for drug transport (Grinius and Goldberg, 1994; Gutman et al., 2003; Muth and Schuldiner, 2000; Yerushalmi and Schuldiner, 2000). Replacing Glu13 in QacC with the smaller acidic residue Asp resulted in host cell susceptibility to benzalkonium (Grinius and Goldberg, 1994). In EmrE replacements of Glu14 with Asp or Cys abolished EmrE mediated resistance against ethidium, acriflavine and methyl viologen (Yerushalmi and Schuldiner, 2000). In addition to this residue, Cys-scanning mutagenesis of residues in TMS1 of EmrE revealed that Glu14 is surrounded by a cluster of residues, located on the same face as Glu14, that form the substrate and proton binding site (Gutman et al., 2003). Amongst these, mutations of Leu7 and Gly17 led to the inability of EmrE to bind substrates, whilst mutations of Thr18 and Ile11 resulted in lower substrate binding affinity and finally a mutation of Ala10 modified EmrE affinity for H^+, resulting in uncoupled transport (Gutman et al., 2003).

1.4.3.3 Molecular mechanism

Charged residues have been found in TMS of numerous transport proteins, although their insertion is energetically unfavourable (Yerushalmi and Schuldiner, 2000b). Based on their location, within a transmembrane domain, it is postulated that these not only play a role in substrate binding, but also in proton transfer. Such “time-sharing”, whereby a single common binding site is used for both substrates and protons, has been observed for EmrE (Yerushalmi and Schuldiner, 2000a, 2000b).

Members of the SMR family are known as proton-coupled transporters, and are proposed to function as dimers (Bay et al., 2008; Korkhov and Tate, 2009; Schuldiner, 2009). Structural and functional analysis of EmrE reveals that dimeric EmrE exchanges two protons/substrate/transport cycle (Rotem and Schuldiner, 2004). This suggests that the substrate is in contact with both Glu14 residues as part of a dimer and that proton binding to each of these residues leads to the

release of the substrate at the periplasmic surface (Yerushalmi and Schuldiner, 2000a, 2000b). Substrate binding has been proposed to occur at the monomer-monomer interface, at the centre of the protein, from the inner leaflet of the cytoplasmic membrane, and results with the movement of at least one TMS (Tate et al., 2003).

As determined for membrane transporters from the MFS, MATE and RND families, the single-site alternating access model has also been proposed for the mechanism of SMR proteins (Fleishman et al., 2006; Yerushalmi and Schuldiner, 2000b). Thus each state in the transport cycle is only open to one side of the membrane (outward-facing conformation) and only when a substrate or 2 protons are bound does the protein interconvert to an inward facing state, as illustrated in Figure 1.10 (Henzler-Wildman, 2012; Morrison et al., 2012).

1.5 The *Escherichia coli* EmrE – a paradigm SMR

As mentioned in Section 1.4.1, the *E. coli* EmrE multidrug transporter, initially known as the MvrC protein, is considered to be an archetypical SMR efflux pump which has been extensively studied using numerous experimental techniques (Korkhov and Tate, 2009; Morimyo et al., 1992; Yerushalmi et al., 1995). This 12 kilodaltons (kDa) protein consists of 110 residues and extrudes positively charged aromatic compounds such as ethidium, acriflavine, methyl viologen, benzalkonium, TPP, 1-methyl-4-phenylpyridinium, tetracycline, erythromycin and sulfadiazine (Morrison and Henzler-Wildman, 2014; Schuldiner et al., 1997; Schwaiger et al., 1998; Yerushalmi et al., 1995; Yerushalmi et al., 2001). Its small size and stability and its ability to retain its function when solubilised in detergent has allowed for many biochemical and biophysical studies to be carried out, providing valuable insight into the nature of proton coupling of SMRs (Sharoni et al., 2005; Yerushalmi and Schuldiner, 2000a, 2000b) (Section 1.4.4.3).

Structural and functional analysis of EmrE indicates that it is functional as a dimer (Butler et al., 2004; Elbaz et al., 2004; Tate et al., 2001). The organisation of EmrE monomers within the membrane was initially a source of great debate, as studies proposed a parallel organisation of monomers within the membrane

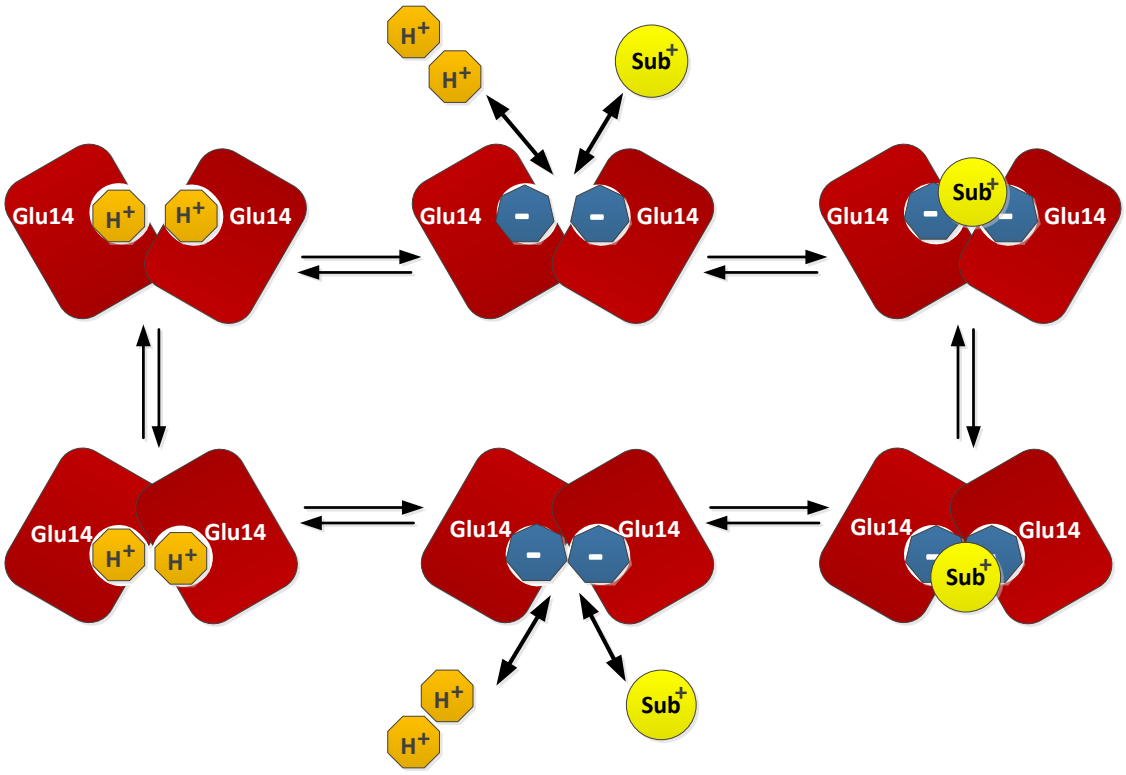

Figure 1.10: Transport mechanism used by the *E. coli* EmrE multidrug transporter

A proposed model of the mechanism of transport by EmrE (Schuldiner, 2012). According this model, functioning as a dimer, binding of the substrate (Sub^+) is carried out by two negatively charged Glu14 residues located within TMS1 of both EmrE monomers. Substrate binding shifts the equilibrium of the protein which leads to the release of 2 protons (H^+) from the binding site. Substrate binding also induces a change in the conformation of the transporter that opens the binding site to the opposite side of the membrane. Proton binding displaces the substrate resulting in its release, and reorients the binding site. Protons are shown as orange octagons whilst the substrate is depicted by a yellow ball. Figure modified from Schuldiner, (2012).

(Soskine et al., 2006; Steiner-Mordoch et al., 2008) whilst others believed that EmrE monomers took on an antiparallel orientation (Lloris-Garcera et al., 2012; Nara et al., 2007; Rapp et al., 2007). Nonetheless, continued research into EmrE topology has shown growing consensus for the antiparallel topology, with the term "dual-topology protein" becoming more and more synonymous with EmrE (Rapp et al., 2006; Rapp et al., 2007; Seppala et al., 2010; von Heijne, 2006). Much of this support comes from X-ray structures of EmrE, which have been determined both in the apo form and bound with TPP (Chen et al., 2007).

1.5.1 Structural assessment of EmrE

The structure of EmrE has been determined by electron cryomicroscopy (cryo-EM) and X-ray crystallography. The first to be determined was the 3D structure of EmrE bound to TPP by cryo-EM of 2D crystals, at a resolution of 7.0 Å (Ubarretxena-Belandia et al., 2003). According to this structure EmrE is an asymmetric homodimer that consists of a bundle of eight TMS with one substrate molecule that is bound near the centre (Figure 1.11). This TMS arrangement and asymmetry was also shown in the crystal structure of EmrE in complex with TPP, determined at a resolution of 3.8 Å. According to this structure, the asymmetrical arrangement appeared to stem from an antiparallel topology within the homodimer (Chen et al., 2007). However, given this unusual topology, it was proposed that the arrangement of this structure was in fact an artefact of the crystallisation conditions (Schuldiner, 2007).

The antiparallel orientation proposed for EmrE was surprising, as MPs generally insert in one particular orientation and follow the "positive-inside rule". This rule states that loops which are rich in positively charged residues such as Lys and Arg, tend to orient towards the cytoplasm (Heijne, 1986) (Section 5.5). However, SMR proteins like EmrE appear to have a weak charge bias, and as such may not conform to this rule (Rapp et al., 2007). The plasticity of EmrE was demonstrated by a study which altered the orientation of EmrE monomers by the addition of a single charged residue at the C-terminal end (Seppala et al., 2010). EmrE reconstituted into various membrane-mimetic environments appears to support the antiparallel

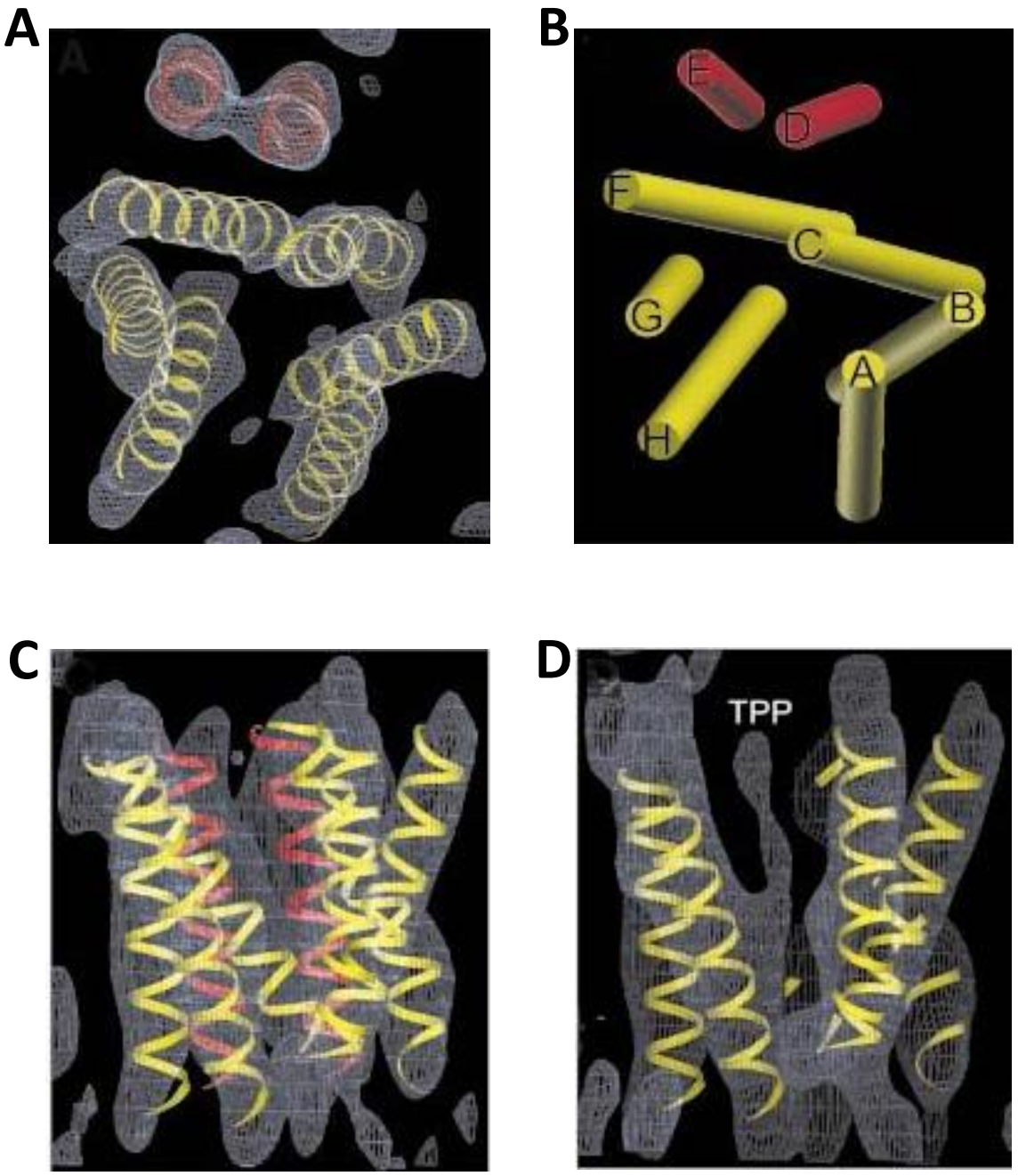

Figure 1.11: Three dimensional density map of EmrE

Density maps of EmrE in its dimeric form as shown from the top and side views. (A) Top view of the 3D EmrE structure perpendicular to the membrane plane. (B) Schematic view of the EmrE structure shown perpendicular to the membrane plane with all helices (A-H) shown as straight cylinders. (C) Side view of the EmrE 3D structure as shown along the membrane plane. (D) The location of TPP, an EmrE substrate, as shown from the side view of a slice shown along the membrane plane of the density map. Helices are represented in two sets, those shown in yellow form the substrate binding pocket, whilst those in red are separated from the pocket by helices F and C. Figure modified from Ubarretxena-Belandia et al., (2003).

orientation of EmrE (Cho et al., 2014; Morrison et al., 2012). However, more analysis is being continually carried out assessing the role of the lipids in EmrE folding, function and stability, with studies showing that dimer stability and affinity is greatly influenced by the lipid membrane (Dutta et al., 2014). According to these studies, lipid curvature, lateral pressure and charge of lipid may ultimately affect the insertion of a protein, its folding, oligomerisation form, dynamics and finally, its overall function, as has been demonstrated for other proteins (Dutta et al., 2014; Fink et al., 2012; Iscla and Blount, 2012; Mitchell, 2012; Soubias and Gawrisch, 2012).

Given the plethora of biochemical and structural studies performed on EmrE, findings for EmrE were used as a point of comparison to the results obtained throughout this study for QacC. In Chapter 3, structural data obtained from solved EmrE structures was used to generate and evaluate models of QacC. This was extended in Chapters 6 and 7, where the topology and oligomerisation of QacC was examined and results thereof compared with those for EmrE. Finally, QacC functionally relevant residues, as determined in Chapter 5, were compared with those shown to be critical for function for EmrE.

1.6 SMR family transporters in *S. aureus*

1.6.1 QacC (QacD, Smr, Ebr)

The QacC multidrug transporter was the first SMR family member identified. It was originally cloned from *S. aureus* strains that were shown to be resistant to disinfectants such as ethidium and benzalkonium chloride (Littlejohn et al., 1991). The *qacC* determinant was found to be located on a small 2.4 kb plasmid identified as pSK89 which was isolated from an Australian hospital (Lyon and Skurray, 1987). Homology with pSK89 was also identified with plasmids pJE1 and pSK41, suggesting that these too carried resistance determinants which were similar or identical to the determinant which was responsible for the resistance attributed to QacC, as such this determinant was referred to as *qacD.* Although initially assumed that the *qacC* gene stemmed from *qacD*, it was later shown that that these genes were

identical and thus encoded one protein only, QacC (Grinius et al., 1992). In addition to QacD, the QacC MP has also been referred to as the smr (staphylococcal multidrug resistance) protein (Grinius et al., 1992) as well as the antiseptic-resistance gene ebr (ethidium bromide resistance protein) (Sasatsu et al., 1989; Sasatsu et al., 1995). Recent publications have favoured the name 'smr' in reference to QacC (Poget et al., 2010), however, this has the potential to be confusing as QacC belongs to the SMR family, and as such will continue to be referred to as QacC in this study.

Comparative analysis of the QacC amino acid sequence with other proteins such as the *S. aureus* QacA, the *B. subtilis* Bmr and the *E. coli* TetA proteins revealed low sequence identity implying an independent origin of the *qacC* gene. Further analysis revealed that QacC shared high sequence similarity with proteins belonging to a family of small hydrophobic proteins, which were all encoded by short open reading frames (ORFs) (~330 bp). Of these, QacC was shown to be most closely related to the *E. coli* MvrC protein, which is now known as EmrE (Section 1.5), showing a 42% sequence identity (Figure 1.12) (Grinius et al., 1992; Yerushalmi et al., 1995). In addition to displaying high homology with EmrE, QacC was also shown to contain a number of highly conserved residues which were identified in other SMR proteins belonging to bacteria such as *P. vulgaris*. One such residue was the highly conserved Glu13 residue, which was proposed to serve as a binding site for lipophilic cations and protons.

Topological assessment of QacC, initially investigated by the construction and analysis of a series of *qacC-phoA* and *qacC-lacZ* fusions, supported the proposed four TMS topology seen for SMR proteins (Paulsen et al., 1995). In addition, the orientation of the N terminus was shown to be cytoplasmic, however, the location of the C terminus could not be unequivalently determined. Further analysis, using nuclear magnetic (NMR) spectroscopy revealed that QacC is composed of four TMS, with residues 6-23 contained in TMS1, 32-40 TMS2, 60-71 TMS3 and 94-101 TMS4 (Poget et al., 2010). As with other members of the SMR subfamily, the QacC functional unit is postulated to be dimeric (Chapter 7).

```
QacC  2    PYIYLIIAISTEVIGSAFLKSSEGFSKFIPSLGTIISFGICFYFLSKTMQHLPLNITYAT  61
           PYIYL  AI  EVIG+  +K SEGF++  PS+GTII +   F+ L++T+ ++P  I YA
EmrE  3    PYIYLGGAILAEVIGTTLMKFSEGFTRLWPSVGTIICYCASFWLLAQTLAYIPTGIAYAI  62

QacC  62   WAGLGLVLTTVVSIIIFKEQINLITIVSIVLIIVGVVSLNIFGTS  106
           W+G+G+VL +++S   F ++++L  I+ ++LI  GV+ +N+   S
EmrE  63   WSGVGIVLISLLSWGFFGQRLDLPAIIGMMLICAGVLIINLLSRS  107
```

Figure 1.12: A pairwise sequence alignment between QacC and EmrE

A BLASTp (NCBI) pairwise sequence alignment between QacC (P14319.1) and EmrE (P23895.1), showing a 42% sequence identity. Conserved residues are listed; positive residues are denoted with a plus whilst residues with no similarity are indicated with a space.

Functional analysis of the QacC multidrug transporter revealed that it conferred host resistance to a broad spectrum of antimicrobials including QACs such as cetyltrimethylammonium bromide, cetylpyridinium chloride, benzalkonium and dequalinium as well as dyes such as ethidium, proflavine, crystal violet, rhodamine 6G (Paulsen et al., 1995) and pyronin Y (Figure 1.13). In addition to the essential Glu13 residue, site-directed mutagenesis of residues located in motif 3 (Section 1.4.4.1, Figure 1.9), revealed two other highly conserved residues which were essential for function, Tyr59 and Trp62 (Paulsen et al., 1995) as substitutions of these resulted in a significant decrease in the ability of these mutants to confer resistance against any of the compounds tested. As residues in motif 3 have been proposed to play a role in substrate recognition (Paulsen et al., 1993), it was suggested that Tyr59 and Trp62 play an essential role in the structure and/or function of the QacC transport protein. Pro31 was also shown to be conserved and essential for full function amongst members of the SMR family, as substitutions of this residue with either Gly or Ala residues, which carry the similar side chain volume, resulted in a reduction in resistance to both ethidium and crystal violet, suggesting that this residue may reside in an region of the protein that may play a role in determining QacC substrate specificity (Paulsen et al., 1995).

1.6.2 Prevalence of *qacC* genes in clinical isolates of *S. aureus*

Analysis of antiseptic resistance determinants in *S. aureus* strains revealed that *qacC* genes are carried by MRSA strains which have been isolated from clinical settings, and as such contribute to the ongoing problems associated with such resistance. These strains have been identified worldwide including Malaysia, with 6.3% of MRSA shown to carry *qacC* genes (Ghasemzadeh-Moghaddam et al., 2014). Recent investigations in Mustafa Kemal University Research Hospital in Turkey into the prevalence of *qacA/qacB* and *qacC* in staphylococci isolated from surgical site infections (Duran et al., 2014) revealed 36% of MRSA isolates were positive for *qacC*. Previous studies of clinical MRSA isolated from various hospitals have reported a *qacC* gene carriage rate of 31.6% in India, down to 1.9% in samples from Asia (Noguchi et al., 1999).

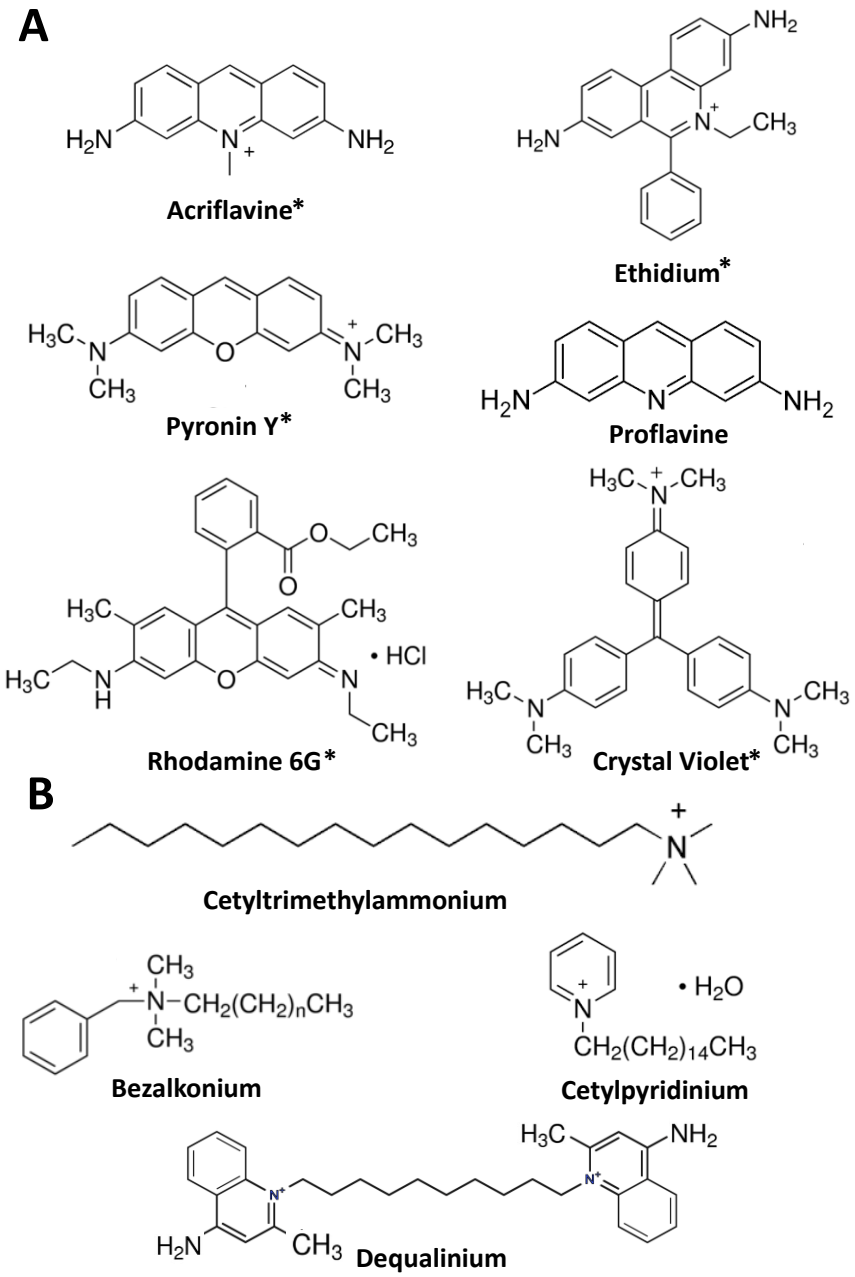

Figure 1.13: Selected substrates of the *S. aureus* QacC multidrug efflux protein

QacC mediates resistance to a number of monovalent lipophilic compounds, which can be classified as (A) dyes and (B) quaternary ammonium compounds. Of these, substrates noted with an asterisk were used throughout this study to further assess QacC function and topology.

Although varying from country to country, the continued presence of the *qacC* and *qacA/qacB* genes in clinically relevant MRSA strains represents an enormous challenge as QACs are disinfectants which are frequently used for hospital infection control. Conferred resistance to these by QacC and other *S. aureus* multidrug efflux pumps represents difficulties in both treatment and disinfection of hospital equipment which may represent a potential source of infection.

1.6.3 Other SMR resistance genes in staphylococci

In addition to QacC, proteins such as QacE, QacEΔ1, QacF, QacG, QacH and QacJ have been identified in a number of antibiotic resistant bacteria, including various staphylococci strains isolated from humans (Correa et al., 2008), animals (Bjorland et al., 2005; Bjorland et al., 2003) and in the food industry (Heir et al., 1998, 1999a, b). The Qac nomenclature given to these proteins is based on their ability to confer host resistance against QACs (Bay et al., 2008).

Assessment of resistance has shown that despite conferring resistance to a wide range of QACs and cationic dyes, Qac proteins display their own unique resistance profiles for specific compounds. High levels of resistance to ethidium have been determined for QacH, QacE and QacEΔ1, but not for QacG. However, QacG has been shown to confer higher levels of resistance to cetyltrimethylammonium bromide than QacH (reviewed in Bay et al., 2008). Displaying distinct substrate ranges, analysis of human staphylococcal clinical isolates have shown that several *qac* genes coexist in some isolates and appear to act synergistically to remove different compounds from the bacterial cells (Correa et al., 2008).

1.6.4 SepA – a chromosomally encoded SMR

Thus far only one SMR like MP, the SepA transporter, has been identified in *S. aureus* to be chromosomally encoded (Narui et al., 2002). The antiseptic resistance gene, *sepA* (staphylococcal efflux pump gene), was first identified in antiseptic resistant *S. aureus* strains isolated in Japan. The SepA protein is unlike any other currently identified SMR protein; consisting of 157 amino acids it is much larger

than any current SMR member. Analysis of this membrane transporter revealed that it confers resistance to acriflavine, but not to the common SMR substrate, ethidium. Limited biochemical assessment of this protein exists and as such it is difficult to classify this protein as a multidrug transporter, or as a member of the SMR family. Phenotypic characterisation of SepA is further discussed in Chapter 8.

1.7 Scope of thesis

Multidrug transporters have been identified in pathogenic bacteria and contribute to resistance against a broad spectrum of antimicrobials, which includes both antibiotics and disinfectants many of which are used in the healthcare setting. Despite the plethora of knowledge gained from studies on the *E. coli* EmrE protein, limited analyses have been undertaken on other SMR proteins. As such, the broad aim of the studies presented in this thesis was to expand on the knowledge of the *S. aureus* QacC multidrug transporter. Given its small size QacC represents an ideal model with which extensive analysis, such as site-directed mutagenesis can be carried out.

The studies presented in this thesis include both structural and functional assessment of the QacC protein. Structural analysis was carried out by computer modelling and site-directed mutagenesis as these provide an effective way with which protein topology can be assessed. Moreover, using Cys-scanning mutagenesis and interaction with various thiol-reactive reagents allows not only for the analysis of the relative position of each residue within the protein, but also for the determination of the physical environment in which the residue resides. As such, Cys-scanning mutagenesis was used to assess every residue within QacC. Assessment included structural analysis, where each QacC mutant was assessed in solvent accessibility studies utilising fluorescein-5-maleimide. Results from these, in combination with computer modelling allowed for the refinement of the topology of QacC. A number of QacC mutants were also evaluated to determine their possible role in substrate binding and/or translocation. In addition to structure, functional assessment was carried out on each mutant. The ability of each mutant to efflux ethidium out of the cell was thus assessed as was the resistance profile

against three common SMR substrates. Finally, the oligomeric conformation of QacC was evaluated using formaldehyde cross-linking and blue native gel electrophoresis (BN-PAGE). In combination, the results obtained in the studies on QacC as outlined in this thesis provide a detailed insight into both the structure and function of QacC and essential residues therein.

Finally, as the *S. aureus* SepA MP represents an unusual SMR, it was further investigated in this study. With only one study carried out SepA (Narui et al., 2002), its amino acid sequence was analysed to evaluate its homology with other members of the SMR family, to ascertain whether it was correctly classified or if it could possibly be a MP which belongs to a as yet unknown membrane transporter family. Finally, functional assays were performed in order to confirm SepA functions as a multidrug transporter, and to clarify its substrate profile.

CHAPTER 2 – MATERIALS AND METHODS

2.1 Solutions and buffers

All reagents listed in Table 2.1 were of molecular grade and obtained from Sigma-Aldrich Australia unless otherwise specified. Powered compounds were dissolved according to the manufacturer's recommendations or as described in the Merck index (Budavari, 1996). All other reagents were of analytical grade and obtained through commercial sources. Carbonyl-cyanide m-chlorophenylhydrazone (CCCP) was prepared as a solution in 1 mg/ml ethanol and then diluted to 100 µg/ml in 10 mM NaOH as previously described (Hassan, 2007). Solutions were sterilised by either autoclaving at 121°C for 15 minutes or passing through a filter with a 0.20 µm pore width.

2.2 Bacterial strains

The bacterial strains used in this study are listed in Table 2.2, along with their relevant genetic characteristics. The plasmids used in this study are listed in Table 2.3, along with a brief description of their properties as well as their source of origin.

2.3 Bacterial culture media, growth conditions and storage

Typically Luria-Bertani (LB) medium contained 1% [w/v] tryptone (Oxoid, Australia), 0.5% [w/v] yeast extract (Oxoid, Australia) and 1% [w/v] NaCl; the pH was adjusted to 7.5. Mueller-Hinton (MH) medium was prepared according to the manufacturer's recommendations (BD, Bacto Laboratories Pty Ltd, Australia). Terrific Broth (TB) contained 1.2% [w/v] tryptone (Oxoid, Australia), 2.4% [w/v] yeast extract (Oxoid, Australia) and 0.4% [v/v] glycerol. The TB medium was supplemented with a final concentration of 0.17 M KH_2PO_4 and 0.72 M K_2HPO_4. For solid media, 1% [w/v] J3 grade agar (Gelita, Australia) was added. All media were sterilised by autoclaving at 121°C for 15 minutes.

Table 2.1: Buffers and solutions

Buffers/solution	Constituents
15% SDS-PAGE gel	2.5 ml of 1.5 M Tris-HCl (pH 8.8), 5 ml of 30% (v/v) acrylamide, 0.1 ml of 10% (v/v) sodium dodecyl sulphate (SDS), 0.1 ml of 10% (v/v) ammonium persulphate, 4 µl of *N,N,N′, N′* – tetramethylethylene diamine (TEMED) and 2.3 ml of dH_2O
Bead wash buffer	50 mM Tris HCl (pH 8.0), 10% (v/v) glycerol, 1% (w/v) n-dodecyl β-D-maltoside (DDM), 50 mM imidazole (pH 8.0)
Binding buffer	50 mM Tris HCl (pH 8.0), 10% (v/v) glycerol, 0.1% (w/v) DDM, 50 mM imidazole (pH 8.0)
Blocking buffer	10% (w/v) skim milk power in T-TBS
Chemiluminescence detection solution	Solution 1: 2% (22 mM 4-aminopthalhydrazide, 42.6 m-Coumaric acid in dimethyl sulfoxide [DMSO]) on 0.1 M Tris buffer (pH 9.35) Solution 2: 0.036% H_2O_2 in 0.1 M Tris buffer (pH 9.35)
Coomassie stain	30% (v/v) methanol, 10% (v/v) glacial acetic acid, 15% (w/v) Coomassie Brilliant Blue R-250
Crushing buffer	50 mM Tris-HCl (pH 7.5), 10% (v/v) glycerol, 300 mM NaCl
Destain solution	30% (v/v) methanol, 10% (v/v) glacial acetic acid
Dialysis buffer	15 mM Tris-HCl (pH 7.5), 10% (v/v) glycerol
Electrophoresis loading buffer	50% (w/v) sucrose, 50 mM ethylenediaminetetra acetic acid (EDTA) (pH 7.0), 0.05% (w/v) bromophenol blue
Elution buffer	50 mM Tris-HCl (pH 7.5), 400 mM imidazole, 0.1 % (w/v) DDM and 10% (v/v) glycerol
KPO_4 10 mM (pH 7.2)	0.717 ml 1 M K_2HPO_4 mixed with 0.283 ml 1 M KH_2PO_4 up to a final volume of 100 ml
Membrane resuspension buffer	20 mM Tris-HCl (pH 7.5), 10% (v/v) glycerol, 1% (w/v) DDM
NEM-quenching buffer	20 mM N-ethylmaleimide (NEM), 50 mM Tris HCl (pH 8.0), 10% (v/v) glycerol
SDS-PAGE running buffer (10X)	250 mM Tris-base (pH 8.3), 1.9 M glycine and 1% (w/v) SDS
SDS-PAGE sample buffer (2X)	0.05 M Tris-HCl (pH 6.8), 5% (v/v) β-mercaptoethanol, 2% (w/v) SDS, 10% (v/v) glycerol, 0.1% (w/v) bromophenol blue

Buffers/solution	Constituents
Stacking SDS-PAGE gel	0.63 ml of 1.0 M Tris-HCl (pH 6.8), 0.83 ml of 30% (v/v) acrylamide, 0.05 ml of 10% (v/v) SDS, 0.05 ml of 10% (v/v) ammonium persulphate, 0.005 ml of TEMED and 3.4 ml of dH_2O
TAE buffer (50x)	24.2% (w/v) Tris-base. 50 mM ethylenediaminetetra acetic acid (pH 8.0), 5.71 (v/v) glacial acetic acid
Transfer buffer	25 mM Tris, 192 mM glycine, 10% (v/v) MeOH (pH 8.3–8.8)
TBS	10 mM Tris-HCl (pH 7.5), 150 mM NaCl
T-TBS	10 mM Tris-HCl (pH 7.5), 150 mM NaCl, 0.1% (v/v) tween
TFBI buffer	30 mM potassium acetate, 100 mM potassium chloride, 100 mM calcium chloride dihydrate, 50 mM manganese(II) chloride tetrahydrate and 15% (v/v) glycerol, (pH 5.8)
TFBII buffer	10 mM 4-morpholinepropanesulfonic acid (MOPS), 75 mM calcium chloride, 100 mM potassium chloride and 15% (v/v) glycerol, (pH 6.5)

Table 2.2: Bacterial strains used in this study

Strain	Genotype	Reference/Source
***E. coli* DH5α**	*supE*44 Δ*lacU*169(ϕ80*lacZ*ΔM15) *hsdR*17 *recA*1 *endA*1 *gyrA*96 *thi*-1 *relA*1	(Hanahan, 1983)
***E. coli* BL21 (DE3)**	B *dcm ompT hsdS*(rB-mB-) *gal* λ(DE3)	(Studier and Moffatt, 1986)
***E. coli* TOP10**	*mcr*A, Δ(*mrr-hsd*RMS-*mcr*BC), Δ*lac*X74, *deo*R, *rec*A1, *ara*D139Δ(*ara-leu*)7697, *gal*K, *rps*L, *end*A1, *nup*G	Invitrogen

Table 2.3: Plasmids used in this study

Plasmid	Description[a]	Reference/ Source
pBAD30	AmpR, pUC-based expression system containing the *araBAD* promoter	(Guzman et al., 1995)
pBluescript II SK	AmpR, pUC-based expression vector containing the T7 promoter	Stratagene
pBADQacC	AmpR, codon optimized *qacC* cloned in pBAD into *Nco*I and *Xba*I downstream of the *araBAD* promoter, carrying the 1D4 and Hisx6 tags	(Dunham et al., 2001)
pBADQacC_C43T	AmpR, C43T QacC derivative of pBADQacC	(Dunham et al., 2001)
pBADQacC_P3C	AmpR, P3C QacC derivative of pBADQacC_C43T	(Dunham et al., 2001)
pBADQacC_Y4C	AmpR, Y4C QacC derivative of pBADQacC_C43T	This study
pBADQacC_I5C	AmpR, I5C QacC derivative of pBADQacC_C43T	This study
pBADQacC_Y6C	AmpR, Y6C QacC derivative of pBADQacC_C43T	This study
pBADQacC_L7C	AmpR, L7C QacC derivative of pBADQacC_C43T	This study
pBADQacC_I8C	AmpR, I8C QacC derivative of pBADQacC_C43T	This study
pBADQacC_I9C	AmpR, I9C QacC derivative of pBADQacC_C43T	This study
pBADQacC_A10C	AmpR, A10C QacC derivative of pBADQacC_C43T	This study
pBADQacC_I11C	AmpR, I11C QacC derivative of pBADQacC_C43T	This study
pBADQacC_S12C	AmpR, S12C QacC derivative of pBADQacC_C43T	This study
pBADQacC_T13C	AmpR, T13C QacC derivative of pBADQacC_C43T	This study
pBADQacC_E14C	AmpR, E14C QacC derivative of pBADQacC_C43T	This study
pBADQacC_V15C	AmpR, V15C QacC derivative of pBADQacC_C43T	This study
pBADQacC_I16C	AmpR, I16C QacC derivative of pBADQacC_C43T	This study

Plasmid	Description[a]	Reference/ Source
pBADQacC_G17C	AmpR, G17C QacC derivative of pBADQacC_C43T	This study
pBADQacC_S18C	AmpR, S18C QacC derivative of pBADQacC_C43T	This study
pBADQacC_A19C	AmpR, A19C QacC derivative of pBADQacC_C43T	This study
pBADQacC_F20C	AmpR, F20C QacC derivative of pBADQacC_C43T	This study
pBADQacC_L21C	AmpR, L21C QacC derivative of pBADQacC_C43T	This study
pBADQacC_K22C	AmpR, K22C QacC derivative of pBADQacC_C43T	This study
pBADQacC_S23C	AmpR, S23C QacC derivative of pBADQacC_C43T	(Dunham et al., 2001)
pBADQacC_S24C	AmpR, S24C QacC derivative of pBADQacC_C43T	This study
pBADQacC_E25C	AmpR, E25C QacC derivative of pBADQacC_C43T	(Dunham et al., 2001)
pBADQacC_G26C	AmpR, G26C QacC derivative of pBADQacC_C43T	This study
pBADQacC_F27C	AmpR, F27C QacC derivative of pBADQacC_C43T	(Dunham et al., 2001)
pBADQacC_S28C	AmpR, S28C QacC derivative of pBADQacC_C43T	This study
pBADQacC_K29C	AmpR, K29C QacC derivative of pBADQacC_C43T	(Dunham et al., 2001)
pBADQacC_F30C	AmpR, F30C QacC derivative of pBADQacC_C43T	(Dunham et al., 2001)
pBADQacC_I31C	AmpR, I31C QacC derivative of pBADQacC_C43T	This study
pBADQacC_P32C	AmpR, P32C QacC derivative of pBADQacC_C43T	This study
pBADQacC_S33C	AmpR, S33C QacC derivative of pBADQacC_C43T	This study
pBADQacC_L34C	AmpR, L34C QacC derivative of pBADQacC_C43T	This study
pBADQacC_G35C	AmpR, G35C QacC derivative of pBADQacC_C43T	This study
pBADQacC_T36C	AmpR, T36C QacC derivative of pBADQacC_C43T	This study
pBADQacC_I37C	AmpR, I37C QacC derivative of pBADQacC_C43T	This study

Plasmid	Description[a]	Reference/ Source
pBADQacC_I38C	AmpR, I38C QacC derivative of pBADQacC_C43T	This study
pBADQacC_S39C	AmpR, S39C QacC derivative of pBADQacC_C43T	This study
pBADQacC_F40C	AmpR, F40C QacC derivative of pBADQacC_C43T	This study
pBADQacC_G41C	AmpR, G41C QacC derivative of pBADQacC_C43T	This study
pBADQacC_I42C	AmpR, I42C QacC derivative of pBADQacC_C43T	This study
pBADQacC_F44C	AmpR, F44C QacC derivative of pBADQacC_C43T	This study
pBADQacC_Y45C	AmpR, Y45C QacC derivative of pBADQacC_C43T	This study
pBADQacC_F46C	AmpR, F46C QacC derivative of pBADQacC_C43T	This study
pBADQacC_L47C	AmpR, L47C QacC derivative of pBADQacC_C43T	This study
pBADQacC_S48C	AmpR, S48C QacC derivative of pBADQacC_C43T	This study
pBADQacC_K49C	AmpR, K49C QacC derivative of pBADQacC_C43T	This study
pBADQacC_K49A	AmpR, K49A QacC derivative of pBADQacC_C43T	This study
pBADQacC_K49R	AmpR, K49R QacC derivative of pBADQacC_C43T	This study
pBADQacC_T50C	AmpR, T50C QacC derivative of pBADQacC_C43T	This study
pBADQacC_M51C	AmpR, M51C QacC derivative of pBADQacC_C43T	This study
pBADQacC_Q52C	AmpR, Q52C QacC derivative of pBADQacC_C43T	This study
pBADQacC_H53C	AmpR, H53C QacC derivative of pBADQacC_C43T	This study
pBADQacC_L54C	AmpR, L54C QacC derivative of pBADQacC_C43T	This study
pBADQacC_P55C	AmpR, P55C QacC derivative of pBADQacC_C43T	This study
pBADQacC_L56C	AmpR, L56C QacC derivative of pBADQacC_C43T	This study
pBADQacC_N57C	AmpR, N57C QacC derivative of pBADQacC_C43T	This study

Plasmid	Description[a]	Reference/ Source
pBADQacC_I58C	AmpR, I58C QacC derivative of pBADQacC_C43T	This study
pBADQacC_T59C	AmpR, T59C QacC derivative of pBADQacC_C43T	This study
pBADQacC_Y60C	AmpR, Y60C QacC derivative of pBADQacC_C43T	This study
pBADQacC_A61C	AmpR, A61C QacC derivative of pBADQacC_C43T	This study
pBADQacC_T62C	AmpR, T62C QacC derivative of pBADQacC_C43T	This study
pBADQacC_W63C	AmpR, W63C QacC derivative of pBADQacC_C43T	This study
pBADQacC_A64C	AmpR, A64C QacC derivative of pBADQacC_C43T	This study
pBADQacC_G65C	AmpR, G65C QacC derivative of pBADQacC_C43T	This study
pBADQacC_L66C	AmpR, L66C QacC derivative of pBADQacC_C43T	This study
pBADQacC_G67C	AmpR, G67C QacC derivative of pBADQacC_C43T	This study
pBADQacC_L68C	AmpR, L68C QacC derivative of pBADQacC_C43T	This study
pBADQacC_V69C	AmpR, V69C QacC derivative of pBADQacC_C43T	This study
pBADQacC_L70C	AmpR, L70C QacC derivative of pBADQacC_C43T	This study
pBADQacC_T71C	AmpR, T71C QacC derivative of pBADQacC_C43T	This study
pBADQacC_T72C	AmpR, T72C QacC derivative of pBADQacC_C43T	This study
pBADQacC_V73C	AmpR, V73C QacC derivative of pBADQacC_C43T	This study
pBADQacC_V74C	AmpR, V74C QacC derivative of pBADQacC_C43T	This study
pBADQacC_S75C	AmpR, S75C QacC derivative of pBADQacC_C43T	This study
pBADQacC_I76C	AmpR, I76C QacC derivative of pBADQacC_C43T	This study
pBADQacC_I77C	AmpR, I77C QacC derivative of pBADQacC_C43T	This study
pBADQacC_I78C	AmpR, I78C QacC derivative of pBADQacC_C43T	This study

Plasmid	Description[a]	Reference/ Source
pBADQacC_F79C	AmpR, F79C QacC derivative of pBADQacC_C43T	This study
pBADQacC_K80C	AmpR, K80C QacC derivative of pBADQacC_C43T	This study
pBADQacC_E81C	AmpR, E81C QacC derivative of pBADQacC_C43T	This study
pBADQacC_Q82C	AmpR, Q82C QacC derivative of pBADQacC_C43T	This study
pBADQacC_I83C	AmpR, I83C QacC derivative of pBADQacC_C43T	This study
pBADQacC_N84C	AmpR, N84C QacC derivative of pBADQacC_C43T	This study
pBADQacC_I85C	AmpR, I85C QacC derivative of pBADQacC_C43T	This study
pBADQacC_I86C	AmpR, I86C QacC derivative of pBADQacC_C43T	This study
pBADQacC_T87C	AmpR, T87C QacC derivative of pBADQacC_C43T	This study
pBADQacC_I88C	AmpR, I88C QacC derivative of pBADQacC_C43T	This study
pBADQacC_V89C	AmpR, V89C QacC derivative of pBADQacC_C43T	This study
pBADQacC_S90C	AmpR, S90C QacC derivative of pBADQacC_C43T	This study
pBADQacC_S90G	AmpR, S90G QacC derivative of pBADQacC_C43T	This study
pBADQacC_S90I	AmpR, S90I QacC derivative of pBADQacC_C43T	This study
pBADQacC_I91C	AmpR, I91C QacC derivative of pBADQacC_C43T	This study
pBADQacC_V92C	AmpR, V92C QacC derivative of pBADQacC_C43T	This study
pBADQacC_L93C	AmpR, L93C QacC derivative of pBADQacC_C43T	This study
pBADQacC_I94C	AmpR, I94C QacC derivative of pBADQacC_C43T	This study
pBADQacC_I95C	AmpR, I95C QacC derivative of pBADQacC_C43T	This study
pBADQacC_V96C	AmpR, V96C QacC derivative of pBADQacC_C43T	This study
pBADQacC_G97C	AmpR, G97C QacC derivative of pBADQacC_C43T	This study

Plasmid	Description[a]	Reference/ Source
pBADQacC_G97I	AmpR, G97I QacC derivative of pBADQacC_C43T	This study
pBADQacC_V98C	AmpR, V98C QacC derivative of pBADQacC_C43T	This study
pBADQacC_V99C	AmpR, V99C QacC derivative of pBADQacC_C43T	This study
pBADQacC_S100C	AmpR, S100C QacC derivative of pBADQacC_C43T	This study
pBADQacC_L101C	AmpR, L101C QacC derivative of pBADQacC_C43T	This study
pBADQacC_N102C	AmpR, N102C QacC derivative of pBADQacC_C43T	This study
pBADQacC_I103C	AmpR, I103C QacC derivative of pBADQacC_C43T	This study
pBADQacC_F104C	AmpR, F104C QacC derivative of pBADQacC_C43T	This study
pBADQacC_G105C	AmpR, G105C QacC derivative of pBADQacC_C43T	This study
pBADQacC_K22C_K29C	AmpR, K22C_K29C K29C derivative of pBADQacC_K22C	This study
pBADQacC_K22C_K49C	AmpR, K22C_K49C K49C derivative of pBADQacC_K22C	This study
pBADQacC_K22C_K80C	AmpR, K22C_K80C K80C derivative of pBADQacC_K22C	This study
pBADQacC_K29C_K49C	AmpR, K29C_K49C K49C derivative of pBADQacC_K29C	This study
pBADQacC_K29C_K80C	AmpR, K29C_K80C K80C derivative of pBADQacC_K29C	This study
pBADQacC_E25C_K49A	AmpR, E25C_K49A K49C derivative of pBADQacC_E25C	This study
pBADQacC_K49C_K80C	AmpR, K49C_K80C K80C derivative of pBADQacC_K49C	This study
pBADQacC_K49A_G105C	AmpR, K49A_G105C G105C derivative of pBADQacC_K49A	This study
pBADQacC_S90C_G97C	AmpR, S90C_G97C G97C derivative of pBADQacC_S90C	This study
pBSQacC_C43T	AmpR, C43T QacC derivative of pBADQacC cloned into *Xho*I and *Eco*RI downstream of the T7 promoter, carrying the 1D4 and Hisx6 tag	This study
pBSQacC_C43THis6	AmpR, codon optimized *qacC* in the pBS expression vector, cloned into *Nco*I and *Eag*I downstream of the T7 promoter carrying the Hisx6 tag	This study

Plasmid	Description[a]	Reference/ Source
pBSQacC_Y4C	AmpR, Y4C QacC derivative of pBSQacC_C43T	This study
pBSQacC_L7C	AmpR, L7C QacC derivative of pBSQacC_C43T	This study
pBSQacC_I11C	AmpR, I11C QacC derivative of pBSQacC_C43T	This study
pBSQacC_S12C	AmpR, S12C QacC derivative of pBSQacC_C43T	This study
pBSQacC_G17C	AmpR, G17C QacC derivative of pBSQacC_C43T	This study
pBSQacC_S18C	AmpR, S18C QacC derivative of pBSQacC_C43T	This study
pBSQacC_A19C	AmpR, A19C QacC derivative of pBSQacC_C43T	This study
pBSQacC_K22C	AmpR, K22C QacC derivative of pBSQacC_C43T	This study
pBSQacC_S23C	AmpR, S23C QacC derivative of pBSQacC_C43T	This study
pBSQacC_E25C	AmpR, E25C QacC derivative of pBSQacC_C43T	This study
pBSQacC_K49C	AmpR, K49C QacC derivative of pBSQacC_C43T	This study
pBSQacC_T50C	AmpR, T50C QacC derivative of pBSQacC_C43T	This study
pBSQacC_Q52C	AmpR, Q52C QacC derivative of pBSQacC_C43T	This study
pBSQacC_H53C	AmpR, H53C QacC derivative of pBSQacC_C43T	This study
pBSQacC_Q82C	AmpR, Q82C QacC derivative of pBSQacC_C43T	This study
pBSSepA	AmpR, pBluescript II SK-derived expression plasmid containing *sepA* cloned into *Nco*I and *Xba*I sites downstream of the T7 promoter, carrying the Hisx6 tag	This study
pBSSepAΔ1	AmpR, SepA derivative of pBSSepA	This study
pBSSepAΔ1QacCTMS4	AmpR, SepA derivative of pBSSepAΔ1	This study
pSK7201	AmpR, pBluescript II SK-derived expression plasmid containing *qacA* cloned into the *Pst*I site downstream of the T7 promoter, carrying the Hisx6 tag	(Xu, 2005)

a: AmpR, ampicillin resistance

Both *E. coli* and *S. aureus* were incubated at 37°C, LB medium was used for standard overnight culturing unless stated otherwise. Liquid bacterial cultures were grown on a gyratory shaker at 220 rpm. For antibiotic selection, the media were allowed to cool to approximately 50°C after autoclaving. In experiments with *E. coli* this was followed by the addition of filter sterilised ampicillin at a final concentration of 100 µg/ml (LBAmp). For long term storage, 1 ml of overnight (16 hour) bacterial culture, grown from a single colony, was mixed with 22 µl of sterile 80% glycerol and stored at -80°C.

2.4 Transformation of *E. coli*

2.4.1 Preparation of chemically competent *E. coli* cells

Overnight cultures of the strain of interest were grown from single, well isolated colonies, in 10 ml of LB. These were used to inoculate fresh LB at a 1:20 dilution. Cells were grown to OD_{600}=0.6 and incubated on ice for 10-50 minutes. The cells were then pelleted by centrifugation (8,000 x *g* for 10 minutes at 4°C), resuspended in TFBI buffer (Table 2.1) and incubated on ice for 15 minutes. Cells were pelleted by centrifugation (8000 x *g* for 10 minutes at 4°C), resuspended in TFBII buffer (Table 2.1) and incubated on ice for 10 minutes. Single aliquots of competent cells (100 µl) were made and stored at -80°C.

2.4.2 Transformation of competent *E. coli* cells

Aliquots of 100 µl of competent *E. coli* cells (Section 2.4.1) were thawed on ice prior to use (10-15 minutes). Typically 50 ng of purified plasmid DNA (Section 2.5.1) or 20 ul of ligation mixture (Section 2.6.5), equating to approximately 50 to 100 ng of DNA, was added to a 100 µl aliquot of competent cells and incubated on ice for 30 minutes. The cells were then heat shocked by rapidly heating to 42°C for 1 minute and then allowed to recover on ice for 2 minutes. Eight hundred µl of LB broth was added to each transformation mix and cells were incubated at 37°C for 60 minutes, after which they were plated onto solid LB agar media containing 100 µg/ml of ampicillin. Plates were incubated overnight at 37°C and single, well-isolated colonies were chosen for further analysis.

2.5 Molecular Procedures

2.5.1 Small-scale purification of plasmid DNA

Plasmid DNA for screening, sequencing and cloning reactions was isolated using the Real Genomics HiYield™ plasmid miniprep kit following the manufacturer's instructions. Briefly, single, well-isolated bacterial clones were used to inoculate 10 ml of sterile LB broth in 50 ml tubes and cultures were incubated overnight at 37°C with shaking. These were harvested by centrifugation (8,000 x *g* for 5 minutes at room temperature [RT]) and resuspended in 400 µl of PD1 buffer, then vortexed to resuspend cells. Cells were lysed by the addition of 400 µl of PD2 Buffer with gentle inversion. After 2 minutes, the lysis mix was neutralised by the addition of 500 µl of PD3 buffer with gentle inversion. Cellular debris was pelleted by centrifugation (17,000 x *g* for 3 minutes at RT) and the supernatant passed through a Real Genomics HiYield™ plasmid miniprep column via centrifugation (17,000 x *g* for 30 seconds at RT). Columns were initially washed with 400 µl W1 buffer during centrifugation, followed by 600 µl of Wash buffer (ethanol added) and further centrifugation (17,000 x *g* for 30 seconds at RT). Columns with bound DNA were dried by centrifugation (17,000 x *g* for 3 minutes at RT) and transferred to new 1.5 microfuge tubes. Following a 2 minute incubation at RT, the DNA was eluted in 50 µl of elution buffer by centrifugation (17,000 x *g* for 2 minutes at RT). All DNA concentrations were measured by a NanoDrop 1000 spectrophotometer (Biolab) using 2 µl of eluted DNA.

2.5.2 Agarose gel electrophoresis

DNA fragments were electrophoresed through horizontal agarose gels following standard molecular biological methods (Sambrook and Russell, 2001). Agarose (Bioline) was dissolved in 0.5X TAE buffer (Table 2.1) at a concentration of 1% for visualisation of DNA fragments of >500 bps and 2% for fragments less than 500 bps. Prior to electrophoresis, 0.2 volumes of Bioline 5X DNA electrophoresis loading buffer (Table 2.1) were added to DNA samples. Molecular weight (MW) standards (1 kb DNA ladder, Bioline) covering the size of the DNA fragments of

interest were co-electrophoresed with samples. Electrophoresis was conducted in 0.5X TAE buffer (Table 2.1) at 100 V until the loading buffer dye front had migrated approximately three quarters of the length of the gel. Gels were stained in 3X GelRed Nucleic Acid Gel Stain (Biotium) for approximately 5 minutes. Stained DNA fragments were visualised and photographed using a BioRad Gel Doc™ EZ system.

2.5.3 Purification of DNA fragments from agarose gels

DNA fragments to be purified were electrophoresed and stained as described in Section 2.5.2, then visualised using a BioRad Gel Doc™ EZ system. Fragments of interest were excised with a sterile scalpel blade, placed in a sterile 1.5 ml microfuge tube and purified using the Real Genomics HiYield™ Gel/PCR DNA extraction kit, according to the manufacturer's specifications. Briefly, 500 µl of HiYield DF buffer was added to 200-300 mg of agarose pieces and incubated at 55°C for 10 to 15 minutes. DNA fragments were then bound to HiYield™ DF columns, washed with 500 µl of wash buffer, and eluted in 15 to 30 µl of elution buffer.

2.5.4 DNA sequencing

Plasmid DNA was prepared using the Real Genomics HiYield™ plasmid miniprep kit, as described in Section 2.5.1. Typically, sequencing samples were prepared by mixing 500 ng of plasmid DNA with 6 pmol of appropriate sequencing primer (Table 2.4) in a 12 µl volume. DNA sequencing reactions were performed at the Australian Genome Research Facility. Nucleotide sequences were compiled analysed and stored using CLC Sequence Viewer 6.8.2 (CLC bio).

2.6 Recombinant DNA procedures

2.6.1 Polymerase chain reaction

Typically, 50 µl polymerase chain reaction (PCRs) contained approximately 100 ng of plasmid DNA template or genomic DNA, 0.4 mM of each dNTP (Finnzymes), 200 nM of each primer complementary to the 3' and 5' ends of the region to be complemented, 2.5 U *Velocity* (Bioline) DNA polymerase and 1X native *Velocity*

Table 2.4: Primer names and nucleotide sequence

Name	Nucleotide sequence (5' → 3')[a]	Location[b]
QacC mutagenesis primers		
Y4C_For	GGAATTAACCATGGGTCCA**TGC**ATCTACCTGATCATCGCG	-1-31
I5C_For	GGAATTAACCATGGGCCCGTAC**TGC**TACCTGATCATCGCG	-1-31
Y6C_For	CCATGGGTCCGTACATA**TGC**CTGATCATCGCGATTAGTACTG	-1-41
L7C_For	GGTCCGTACATCTAC**TGC**ATCATCGCGATTAGTACTG	4-41
I8C_For	CCGTACATCTACCTA**TGC**ATCGCGATTAGTACTGAAGTTATCGG	7-51
I9C_For	GGTCCGTACATCTACCTCATA**TGC**GCGATTAGTACTGAAGTTATCGG	4-51
A10C_For	CCGTACATCTACCTGATCATC**TGC**ATTAGCACTGAAGTTATCGG	7-51
I11C_For	CCGTACATCTACCTGATCA**TCG**CATGCAGTACTGAAGTTATCGGTAGC	7-55
S12C_For	CTGATCATCGCGATT**TGT**ACAGAAGTTATCGGTAGC	19-55
T13C_For	CTGATCATCGCGATTAGT**TGT**GAAGTTATCGGTAGC	19-55
E14C_For	CTGATCATCGCGATTAGTACA**TGT**GTTATCGGTAGCGCT	19-58
V15C_For	GCGATTAGTACTGAA**TGC**ATCGGTAGCGCTTTCCTG	28-64
I16C_For	GCGATTAGTACTGAAGTT**TGC**GGATCCGCTTTCCTGAAGAGCTCC	28-73
G17C_For	GATTAGTACTGAAGTTA**TCT**GCAGCGCTTTCCTGAAGAGC	30-70
S18C_For	GTACTGAAGTTATCGGT**TGC**GCATTCCTGAAGAGCTCC	36-73
A19C_For	GAAGTTATCGGTAGC**TGT**TTCCTGAAGAGCTCC	40-73
F20C_For	GTTATCGGTAGCGCA**TGC**CTGAAGAGCTCCGAAGG	43-77
L21C_For	GTTATCGGTAGCGCTTTC**TGT**AAGAGCTCCGAAGG	43-77
K22C_For	GTTATCGGTAGCGCTTTCCTC**TGC**AGCTCCGAAGGTTTC	43-81
S24C_For	GCTTTCCTGAAGAGC**TGC**GAAGGTTTCTCCAAATTC	53-90

Name	Nucleotide sequence (5' → 3')[a]	Location[b]
G26C_For	GCTTTCCTGAAGAGCTCCGAA**TGC**TTCTCCAAATTCATCCC	53-95
S28C_For	GAGCTCCGAAGGTTTC**TGC**AAATTCATCCCGTCC	66-99
I31C_For	GGTTTCTCCAAATTC**TGC**CCGTCCCTGGGTACC	76-108
P32C_For	TTCTCCAAATTCATC**TGC**TCCCTGGGTACCATC	79-114
S33C_For	TTCTCCAAATTCATCCCG**TGC**TTAGGTACCATCATCTCC	79-120
L34C_For	CCAAATTCATCCCGTCC**TGC**GGTACCATCATCTCC	83-117
G35C_For	CATCCCGTCCCTG**TGT**ACCATCATCTCCTTCGG	90-122
T36C_For	CCCGTCCCTGGGT**TGC**ATCATCTCCTTCGG	93-122
I37C_For	CCGTCCCTGGGTACC**TGC**ATCTCCTTCGGTATC	94-126
I38C_For	CCCTGGGTACCATC**TGC**TCCTTCGGTATCACC	98-129
S39C_For	CTGGGTACCATCATC**TGC**TTCGGTATCACCTTC	100-132
F40C_For	CCCTGGGTACCATCATCAGC**TGC**GGTATCACCTTCTAC	98-135
G41C_For	CCCTGGGTACCATCATAAGCTTC**TGT**ATCACCTTCTACTTC	98-138
I42C_For	CCATCATCTCCTTCGGT**TGT**ACATTCTACTTCTTAAGC	107-144
F44C_For	CCATCATCTCCTTCGGTATCACA**TGT**TACTTCTTAAGCAAAACC	107-150
Y45C_For	CGGTATCACCTTCTGCT**TCT**TAAGCAAAACC	120-150
F46C_For	GGTATCACCTTCTAC**TGC**TTAAGCAAAACCATC	121-153
L47C_For	GGTATCACCTTCTACTTC**TGC**AGCAAAACCATGCAGC	121-157
S48C_For	CACCTTCTACTTCTTA**TGC**AAAACCATGCAGCACC	126-160

Name	Nucleotide sequence (5' → 3')[a]	Location[b]
K49C_For	TTCTACTTCTTAAGC**TGT**ACAATGCAGCACCTGCCG	130-165
T50C_For	TACTTCTTAAGCAAA**TGC**ATGCAGCACCTGCCG	133-165
M51C_For	CTACTTCTTAAGCAAAACA**TGT**CAGCACCTGCCGCTG	132-168
Q52C_For	CTTAAGCAAAACCATG**TGC**CACCTGCCGCTGAACATC	138-174
H53C_For	CTTAAGCAAAACCATGCAG**TGT**CTACCGCTGAACATCACCTAC	138-180
L54C_For	GCAAAACCATGCAGCAC**TGC**CCGCTGAACATCACCTACGCC	143-183
P55C_For	GCAAAACCATGCAGCACCTG**TGC**CTGAACATCACCTACGCC	143-183
L56C_For	CCATGCAGCACCTGCCG**TGC**AACATCACCTACGCC	149-183
N57C_For	GCAGCACCTGCCGCTA**TGC**ATCACCTACGCCACC	153-186
I58C_For	GCACCTGCCGCTGAAC**TGC**ACATATGCCACCTGGGCTGGTCTGG	156-199
T59C_For	GCACCTGCCGCTGAACATA**TGC**TACGCCACCTGGGCTGG	156-194
Y60C_For	CCGCTGAACATCACC**TGC**GCCACCTGGGCTGGTCTG	163-198
A61C_For	CTGAACATCACCTAC**TGC**ACCTGGGCTGGTCTG	166-198
T62C_For	CTGCCGCTGAACATCACGTACGCC**TGC**TGGGCTGGTCTGGGTC	160-202
W63C_For	GAACATCACCTACGCCACATG**TGC**TGGTCTGGGTCTG	168-204
A64C_For	ACCTACGCCACCTGG**TGC**GGTCTGGGTCTGGTG	175-207
G65C_For	CACCTACGCCACCTGGGCA**TGC**CTGGGTCTGGTGTTA	174-210
L66C_For	GCCACCTGGGCTGGT**TGC**GGTCTGGTGTTAACC	181-213
G67C_For	CACCTGGGCTGGTCTG**TGT**CTGGTGTTAACCACC	183-216
L68C_For	GGGCTGGTCTGGGT**TGT**GTGTTAACCACCGTTG	188-220

Name	Nucleotide sequence (5' → 3')[a]	Location[b]
V69C_For	GCTGGTCTGGGTCTG**TGT**TTAACCACCGTTGTTTCC	190-225
L70C_For	CCTGGGCTGGTCTGGGACTAGTG**TGC**ACCACCGTTGTTTCC	185-225
T71C_For	CTGGGTCTGGTGTTA**TGC**ACCGTTGTTTCCATC	196-228
T72C_For	GGTCTGGTGTTAACC**TGC**GTTGTTTCCATCATC	199-228
V73C_For	CTGGTGTTAACCACC**TGT**GTTTCCATCATCATC	202-234
V74C_For	GTGTTAACCACCGTT**TGT**TCCATCATCATCTTC	204-237
S75C_For	GTGTTAACCACCGTTGTA**TGC**ATCATCATCTTCAAAGAAC	204-244
I76C_For	GTTAACCACCGTTGTTTCA**TGC**ATCATCTTCAAAGAACAG	206-246
I77C_For	GTTAACCACCGTTGTTTCCATA**TGC**ATCTTCAAAGAACAG	206-246
I78C_For	CCACCGTTGTTTCCATCATA**TGC**TTCAAAGAACAGATT	212-244
F79C_For	GTTGTTTCCATCATCATA**TGC**AAAGAACAGATTAATC	217-253
K80C_For	CCATCATCATCTTC**TGC**GAACAGATTAATCTG	223-255
E81C_For	CCATCATCATCTTCAAA**TGC**CAGATTAATCTGATC	223-258
Q82C_For	CATCATCTTCAAAGAA**TGC**ATTAATCTGATCACC	228-261
I83C_For	CATCTTCAAAGAACAG**TGT**AATCTGATCACCATC	231-264
N84C_For	CAAAGAACAGATT**TGT**CTGATCACCATCG	237-265
L85C_For	CAAAGAACAGATTAAT**TGT**ATCACCATCGTTTCA	237-270
I86C_For	GAACAGATTAATCTG**TGT**ACCATCGTTTCAATTG	241-274
T87C_For	GAACAGATTAATCTGA**TAT**GCATCGTTTCAATTGTTC	241-277
I88C_For	GATTAATCTGATCACCTGCGTTTCAATTGTTCTG	246-279

Name	Nucleotide sequence (5' → 3')[a]	Location[b]
V89C_For	CAGATTAATCTGATCACCATA**TGT**TCAATTGTTCTGATC	244-282
S90C_For	CTGATCACCATCGTA**TGC**ATTGTTCTGATCATCGTTGG	253-290
I91C_For	CTGATCACCATCGTTTCA**TGT**GTTCTGATCATCGTTGG	253-290
V92C_For	CTGATCACCATCGTTTCCATA**TGT**CTGATCATCGTTGGTG	253-292
L93C_For	CACCATCGTTTCAATTGTA**TGC**ATCATCGTTGGTGTTG	258-295
I94C_For	CATCGTTTCAATTGTTCTA**TGC**ATCGTTGGTGTTGTAAG	261-299
I95C_For	GTTTCAATTGTTCTGATC**TGC**GTTGGTGTTGTAAGC	265-300
V96C_For	GTTTCAATTGTTCTGATCATA**TGT**GGTGTTGTAAGCTTG	265-303
G97C_For	GTTCTGATCATCGTTTG**TGT**TGTAAGCTTGAAC	274-306
V98C_For	CTGATCATCGTTGGT**TGT**GTAAGCTTGAACATC	277-309
V99C_For	GATCATCGTTGGTGTC**TGC**AGCTTGAACATCTTCGG	279-314
S100C_For	CATCGTTGGTGTTGTA**TGC**TTGAACATCTTCGGT	282-315
L101C_For	GTTGGTGTTGTAAGC**TGC**AACATCTTCGGTACTTCC	286-321
N102C_For	GTTGGTGTTGTAAGCTTA**TGC**ATCTTCGGTACTTCC	286-321
I103C_For	GTTGTAAGCTTGAAC**TGC**TTCGGTACTTCCCAC	292-324
F104C_For	GTTGTAAGCTTGAACATA**TGC**GGTACTTCCCACGGT	292-327
G105C_For	GTAAGCTTGAACATCTTC**TGT**ACTTCCCACGGTCTA	295-330

Name	Nucleotide sequence (5' → 3')a
Cloning primers	
C43T_For	GAGACTCGAGTTTTTGGGCTAACAGGAGGAATTAACCATG
C43T_Rev	GAGAGAATTGTCAATGATGATGATGATGATGGTC
C43T_RevHis6	GAGACGGCCGTCAATGATGATGATGATGATGTTCTAGACCGTGGGAA
SepAFor	GAGAGGTACCGGAGGCC
SepARev	TCTCTCTAGACTAGTGATGGTGAG
SepA Truncated	GAGAGGTACCAGGAGGTATGGCATGTATATAATCTTTGGGTACTTC
SepATMS3Rev	GAGACTGCAGTTCAGGGTAAAATTGAAGCTT
C43TTMS4For	GAGACTGCAGATTAATCTGATCACCATC
Sequencing primers	
pBAD_For	ATGCCATAGCATTTTTATCC
pBAD_Rev	GATTTAATCTGTATCAGG
M13_For	GTAAAACGACGGCCAGT
M13_Rev	GGAAACAGCTATGACCATG

a: Primers are presented in a 5’ to 3’ direction. Only forward (For) mutagenesis primers are listed. The sequences to reverse (Rev) primers were complementary. The codon substitution produced by each primer is indicated in bold.
b: The intended location of the primers annealing to the DNA sequence of *qacC* shown in Figure 4.2.

(Bioline) buffer. Reactions were performed in thin-walled 0.5 ml tubes and cycled using a Multigene PCR thermal cycler (Labnet Int.). Cycling conditions included a 5 minute initialising step at 95°C, followed by repeated cycles of a 1 minute denaturation step at 95°C, a 1 minute annealing step at 55°C, a 4 minute extension step at 72°C and a 10 minute single final extension step at 72°C. PCR products were visualised by agarose electrophoresis (Section 2.5.2) and were typically purified from agarose gels (Section 2.5.3) or using the Invitrogen PureLink™ PCR purification kit according to the manufacturer's specifications. Briefly, 4 volumes of PureLink™ binding buffer was added to 1 volume of PCR product and transferred to a PureLink™ spin column in a 1.5 ml centrifuge tube and centrifuged (17,000 x *g* for 1 minute at RT). Bound DNA was washed with 650 µl of wash buffer, centrifuged (17,000 x *g* for 3 minutes at RT) and eluted in 50 µl of elution buffer.

2.6.2 Site-directed mutagenesis

Site-directed *qacC* mutants were made using the QuickChange™ (Stratagene) method with *Velocity* DNA polymerase (Bioline). Pairs of complimentary oligonucleotide primers were designed to incorporate the desired codon change(s) and a silent mutation, which either removed or included an endonuclease restriction site in the plasmid template to aid in screening (Table 2.4). Reactions typically contained approximately 100 ng of plasmid DNA template, 0.4 mM dNTPs (Finnzymes), 1.25 U *Velocity* DNA polymerase (Bioline), 1 x native *Velocity* (Bioloine) Buffer and 200 nM of each primer. Reactions were performed in thin-walled 0.5 ml tubes and cycled using a Multigene PCR thermal cycler (Labnet Int). Cycling conditions included a 5 minute initialising step at 95°C, followed by repeated cycles of a 30 seconds denaturation step at 95°C, a 45 seconds annealing step at 55°C, a 7 minute extension step at 72°C and a 21 minute final extension step at 72°C. PCR products were visualised after agarose electrophoresis (Section 2.5.2), and treated with *Dpn*I (Section 2.6.4) prior to being transformed into competent *E. coli* cells (Section 2.4.2).

2.6.3 Colony PCR screening

Colonies were picked up with a sterilised toothpick and transferred to sterilised dH_2O to be used as DNA templates. Optimised colony PCR mixture contained template DNA 1X native *EconoTaq®* (Lucigen) Buffer, 3 mM $MgCl_2$, 0.4 mM dNTPs (Finnzymes), 200 nM of each primer and 1.2 U *EconoTaq®* DNA polymerase. Reactions were performed as in Section 2.6.2. Cycling conditions included a 5 minutes initialising step at 95°C, followed by 30 cycles of a 30 seconds denaturation step at 95°C, a 45 seconds annealing step at 55°C, a 1 minute extension step at 72°C and a 4 minute final extension step at 72°C. PCR products were visualised after agarose electrophoresis (Section 2.5.2).

2.6.4 Restriction endonuclease digestion

Restriction digestions were performed with endonucleases incubated as per manufacturer's recommendations (New England Biolabs, Genesearch, Australia). Endonucleases were heat-inactivated if possible as per manufacturer's recommendations, alternatively, DNA fragments were purified, as described in Section 2.5.1. DNA fragments for cloning were excised from a gel and column purified (Section 2.5.3).

2.6.5 DNA ligation

Ligations were performed in a 1:3 ratio (vector: insert) overnight at 4°C using T4 DNA ligase and ligation buffer (Promega). T4 DNA ligase was heat-inactivated by incubation at 70°C for 10 minutes prior to transformation (Section 2.4.2).

2.7 Protein overexpression and purification

2.7.1 Overexpression of QacC

Overnight cultures of *E. coli* TOP10 cells harbouring plasmids encoding QacC and QacC variants were diluted and grown at 37°C until OD_{600}=0.55. Cultures were then supplemented with 0.02% [w/v] of L-arabinose and grown for a further 2 hours at 37°C. Alternative induction conditions of 0.02% [w/v] L-arabinose was added when cells reached OD_{600}=1.00 and grown for a further 30 minutes.

Overexpression of *E. coli* BL21 (DE3) cells harbouring QacC, QacC variants, SepA, ΔSepA and ΔSepAQacCTM3 were diluted and grown at 37°C until OD_{600}= 0.55, these were then supplemented with 0.1 mM isopropyl-β-D-thiogalactopyranoside (IPTG) and grown for a further 2 hours at 37°C.

2.7.2 Whole cell lysis

Overnight cultures of *E. coli* cells were pelleted (4,100 x *g* for 3 minutes at RT) and resuspended in 1X SDS-PAGE loading buffer (Table 2.1). Resuspended cells were heated at 37°C for 30 minutes prior to separation by SDS-PAGE (Section 2.8.1).

2.7.3 Disruption of *E. coli* membrane vesicles

Harvested *E. coli* cells (Sections 2.7.1), were washed in 1X Tris-buffered saline (TBS) (Table 2.1) centrifuged (8,000 x *g* for 20 minutes at RT) and resuspended in 25 ml of crushing buffer (Table 2.1). Resuspended cells were disrupted by explosive decompression at 30,000 pounds per square inch (PSI) using the Constant Systems Cell Disruptor (Constant Systems Ltd). Remaining intact cells and cellular debris were removed by centrifugation (27,200 x *g* for 20 minutes at RT) and the membrane fraction collected by ultracentrifugation using a swinging-bucket SW32Ti rotor (134,000 x *g* for 1 hour at 4°C) and resuspended in 300 µl of membrane resuspension buffer (Table 2.1). Samples were used in SDS-PAGE for detection of expression by Western blot (Section 2.8.1, 2.9), or purified by affinity chromatography (Section 2.7.4) followed by labelling of QacC exposed Cys residues by fluorescein-5-maleimide (Section 2.14.1). Samples were stored at -20°C.

2.7.4 QacC purification by affinity chromatography

Membrane fractions of *E. coli* TOP10 cells containing QacC and variant QacC proteins were prepared as described in Section 2.7.3. Membranes from each litre of starting culture were solubilised for an hour at 4°C with slow rotation in 5 ml of membrane resuspension buffer supplemented with 1% [w/v] DDM (Table 2.1). One-fifth volume of Pro-Bond™ Nickel-Chelating Resin, equilibrated in bead wash buffer (Table 2.1), was added to solubilised membranes, and left for an hour at 4°C

with slow rotation. The column was washed once with 6 bed volumes of bead wash buffer (Table 2.1) followed by a further three times with 3 bed volumes of bead wash buffer (Table 2.1). QacC proteins were eluted in 500 µl fractions of elution buffer (Table 2.1). Fractions containing QacC protein were dialysed twice against 2 litres of 15 mM Tris-HCl (pH 7.5), 10% [v/v] glycerol. To increase the concentration of QacC, samples were concentrated using Amicon Ultra 15-centrifugal filter units (Merck Millipore) as per manufacturer's specifications.

2.8 Protein polyacrylamide gel electrophoresis

2.8.1 SDS-PAGE

Protein samples in membrane resuspension buffer (Section 2.7.3) or elution buffer (Section 2.7.4) were incubated at 37°C for 30 minutes with 1 volume of 2X SDS-PAGE sample buffer (Table 2.1). SDS-PAGE (15%) gels overlayed with 4% stacking gels were loaded with pre-incubated samples and electrophoresed in 1X SDS-PAGE running buffer (Table 2.1) at 120 V until the dye front had migrated to the bottom of the gel. Pre-stained Precision Plus Protein Standards™ (Bio-Rad) with molecular masses of 10, 15, 20, 25, 37, 50, 75, 100, 150 and 250 kDa were co-electrophoresed with protein samples to give an indication of protein molecular mass. Gels were stained for their protein content and, for protein detection by Western blot transferred to a polyvinylidine difluoride (PVDF) membrane as described in Sections 2.9.1 and 2.9.2, respectively.

2.8.2 Blue-Native PAGE

Blue-Native (BN)-PAGE was conducted using the NativePAGE™ Bis-Tris Gel System according to manufacturer's instructions. QacC was purified as described in Section 2.7.4, and samples were supplemented with the NativePAGE™ G-250 Sample Additive such that the final G-250 concentration in the sample was one quarter of the detergent concentration. Following the addition of 1X NativePAGE™ Sample Buffer, samples were loaded into separate lanes and electrophoresed at 100 V for 1 hour at 4°C. NativeMark™ Unstained Protein Standards with molecular

masses of 20, 66, 146, 242, 480, 720, 1,048 and 1,236 kDa were co-electrophoresed to provide an indication of protein or protein complex molecular masses.

2.9 Protein detection

2.9.1 Coomassie Brilliant Blue R-250 staining of polyacrylamide gels

Polyacrylamide gels were removed from the PAGE vertical gel electrophoresis systems and stained for 45 minutes to 1 hour in Coomassie stain (Table 2.1) followed by destaining until protein bands were clearly visible. Gels were scanned using a BioRad Gel Doc™ EZ system.

2.9.2 Protein transfer to PVDF membrane for Western blot analysis

Pieces of PVDF membrane (Amersham Hybond-P™, GE Healthcare) were cut to the size of the gel, soaked in 100% methanol for 10 seconds, washed in 1X transfer buffer (Table 2.1) and placed on top of the polyacrylamide gel after a brief rinsing with transfer buffer. Two pieces of Whatman paper presoaked in 1X transfer buffer (Table 2.1) were used to sandwich the gel and PVDF membrane. The assembly was loaded into a transfer system (PAGEgel) with the membrane facing the anode. Proteins were transferred to the PVDF membrane by applying a current of 25 mV for 16 hours (overnight) to allow the proteins to migrate to the membrane. The membrane was immediately probed with a specific antibody preparation (Section 2.9.3).

2.9.3 Immunological detection of transferred proteins

PVDF membranes with bound proteins were incubated in 1X T-TBS (Table 2.1) for 5 minutes, then in blocking buffer (Table 2.1) with gentle agitation for 1 hour. Membranes were washed once in 1X T-TBS for 5 minutes and incubated at 37°C for 2 hours in a 1:5,000 dilution of rabbit anti-Hisx6 antibody (Rockland) in 1X T-TBS. They were then washed 3 times for 10 minutes each in 1X T-TBS and incubated at 37°C for 1 hour in a 1:10,000 dilution of goat-anti-rabbit antibody conjugated to horseradish peroxidase (BioRad) in 1X T-TBS (Table 2.1). Membranes were washed 6 times for 5 minutes in 1X T-TBS and the location of the bands detected by

incubation in equal volumes of freshly prepared chemiluminescence Western detection agents, Solution A and B (Table 2.1). Membranes were wrapped in plastic film and exposed to Amersham Hyperfilm™ ECL for 10 minutes in a Kodak X-Omatic cassette. The film was developed using the Kodak X-Omat 1000 processor.

2.9.4 Protein quantification

Protein quantification of purified proteins or total membrane fractions used the DC protein (Bio-Rad) assay, which is a modified Lowry assay, compatible with detergents including DDM. The assay reagents were used according to the manufacturer's specifications. Assays were performed in microtiter F-plates, with samples of bovine serum albumin (BSA) at known concentrations as reference standards. Protein samples were analysed with the Multiskan spectrophotometer (Labsystems MultisKan EX).

2.10 Minimum inhibitory concentration analysis

2.10.1 Microtiter-plate dilution method for pBAD-based constructs

Minimum inhibitory concentration (MIC) analysis was conducted using the microbroth dilution method using 96 well sterile microtiter plates. Fifty µl of fresh MH broth, supplemented with 0.0004% L-arabinose, was added to 84 wells. To the remaining 12 wells, 100 µl of an appropriate antimicrobial compound, at double the highest concentration, was added and subsequently used in a serial dilution. Overnight cultures of freshly transformed *E. coli* TOP10 cells harbouring plasmids of interest were diluted 1:20 in 10 ml of fresh LBAmp broth and grown to OD_{600}=0.55. QacC protein expression was induced by the addition of 0.02% L-arabinose, and grown for a further 2 hours. Cells were diluted 1:100 in MH and 50 µl of each dilution was added to 78 wells with the remaining wells being used as controls. Plates were incubated at 37°C for 18 hours and the MIC determined as the concentration at which 90% growth inhibition was observed. Growth was measured at OD_{600} by the Multiskan spectrophotometer (Labsystems MultisKan EX).

2.10.2 Microtiter-plate dilution method for pBluescript II SK–based constructs

Overnight cultures of *E. coli* BL21 (DE3) cells carrying the pBluescript II SK-based constructs were diluted 1:20 into 10 ml of fresh LBAmp and grown till OD_{600}=0.55. These were then induced with 0.1 mM IPTG and grown for a further hour at 37°C. Cells were diluted again to OD_{600}=0.06 in MH media and samples of 50 µl aliquoted into a 96-well microtiter plate containing 50 µl of increasing concentrations of substrate to be tested. Plates were incubated at 37°C with shaking, and the cell density measured after 5 hours of growth at OD_{600} using the Multiskan spectrophotometer (Labsystems MultisKan EX).

2.10.3 Plate-dilution method

The plate-dilution method was conducted on *E. coli* BL21 (DE3) cells carrying pBluescript II SK–based constructs. Overnight cultures were diluted 1:20 in 10 ml of fresh LBAmp broth and grown to OD_{600}=0.55 when they were induced with 0.1 mM IPTG. After one hour further growth at 37°C, cells were diluted and standardised to OD_{600}=0.6. A series of 10-fold dilutions was prepared and 5 µl of each dilution were spotted on plates containing different concentrations of the test antimicrobial. Colony growth was recorded after 16 hours of incubation at 37°C.

2.11 Transport assays

2.11.1 Ethidium transport assays

E. coli TOP10 or DH5α cells harbouring plasmids of interest were grown in overnight LBAmp cultures from freshly transformed colonies. Cells were diluted 1:20 and grown to OD_{600}=0.6. Where appropriate, cells were induced at OD_{600}= 0.55, with 0.02% L-arabinose or 0.1 mM IPTG, and grown for a further 2 hours. Five ml of cultures was removed, pelleted by centrifugation (7,000 x *g* for 5 minutes at RT) and washed three times in 20 mM HEPES (pH 7.0). Cells were resuspended in 5 ml of 20 mM HEPES (pH 7.0) and kept on ice until required. CCCP was added to a final concentration of 10 µM to 1 ml of washed bacteria, which were loaded with

15 µM of ethidium bromide in a 1.5 ml microfuge tube. Tubes were covered with aluminium foil and incubated at 37°C for 45 minutes with gentle agitation. Loaded cells were harvested by centrifugation (4,100 x *g* for 3 minutes at RT) and washed 3 times in 1 ml of 20 mM HEPES (pH 7.0). Pellets were resuspended in 1 ml of 20 mM HEPES (pH 7.0) and warmed to 37°C immediately before fluorimetric analysis. Ethidium bromide efflux was energised by the addition of sodium formate to a final concentration of 160 mM.

To determine the effect of NEM modification on QacC–Cys mutants, *E. coli* TOP10 cells expressing QacC–Cys mutants were pre-treated with 5 mM NEM for 20 minutes at 37°C. Cells were then washed 2 times with 20 mM HEPES (pH 7.0) prior to being loaded with 15 µM ethidium (see above).

Fluorimetric measurements were performed at 37°C using a Perkin Elmer LS 55 fluorescence spectrophotometer with slit width of 10 nm. The excitation and emission wavelengths for ethidium were 530 nm and 590 nm, respectively. Data were analysed and plotted by Microsoft Office Excel 2007.

2.11.2 Determination of K_m and V_{max}

Analysis of the kinetics of QacC-mediated ethidium efflux was carried out in *E. coli* TOP10 cells. One ml aliquots of cells were prepared as described in Section 2.11.1 and loaded with 0.5 to 16 µM ethidium. Using these concentrations a series of efflux curves were generated and the initial velocity calculated by taking the average of the linear part of each curve. A minimum of 9 substrate concentrations [S] were utilised whose range was selected to be spread out equally above and below the predicted K_m. Using the initial velocity, a Michaelis-Menton graph and a double–reciprocal plot was determined in Microsoft Excel Office 2007, where 1/velocity (v) was plotted as a function of 1/[S]. This was then used to determine the values for both the K_m and the V_{max} using the equation: $1/v = (1/V_{max}) + (K_m/V_{max}) \times 1/[S]$.

2.12 Pyronin Y transport assays

E. coli TOP10 cells harbouring plasmids of interest were grown in overnight LBAmp cultures from freshly transformed colonies. Cells were diluted 1:20 and grown to OD_{600}=0.6. Where appropriate, cells were induced at OD_{600}=0.55 with 0.02% L-arabinose or 0.1 mM IPTG, and grown for a further 2 hours. Five ml of cultures was removed, pelleted by centrifugation (7,000 x *g* for 5 minutes at RT) and washed three times in 20 mM HEPES (pH 7.0). Cells were resuspended in 5 ml of 20 mM HEPES (pH 7.0) and kept on ice until required. CCCP was added at a final concentration of 10 µM to 1 ml of washed bacteria. Cells were loaded with 49 nM pyronin Y in a 1.5 ml centrifuge tube. Tubes were covered with aluminium foil and incubated at 37°C for 45 minutes with gentle agitation. Loaded cells were harvested by centrifugation (4,100 x *g* for 3 minutes at RT) and washed 3 times in 1 ml of 20 mM HEPES (pH 7.0). Pellets were resuspended in 1 ml of 20 mM HEPES (pH7.0) warmed to 37°C immediately before fluorimetric analysis. Pyronin Y efflux was energised by the addition of sodium formate to a final concentration of 160 mM.

Fluorimetric measurements were performed at 37°C using a Perkin Elmer LS 55 fluorescence spectrophotometer with slit width of 10 nm. The excitation and emission wavelengths for pyronin Y were 529 nm and 553 nm, respectively. Data were analysed and plotted by Microsoft Office Excel 2007.

2.13 Acriflavine transport assays

E. coli TOP10 cells harbouring plasmids of interest were grown in overnight LBAmp cultures from freshly transformed colonies. Cells were diluted 1:20 and grown to OD_{600}=0.6. Where appropriate, cells were induced at OD_{600}=0.55, with 0.02% L-arabinose or 0.1 mM IPTG, and grown for a further 2 hours. Five ml of cultures was removed, pelleted by centrifugation (7,000 x *g* for 5 minutes at RT) and washed three times in 20 mM HEPES (pH 7.0). Cells were resuspended in 5 ml of 20 mM HEPES (pH 7.0) and kept on ice until required. Acriflavine at a final concentration of 96 nM was added to 1 ml of cells mixed with 1 ml of 20 mM HEPES (pH 7.0). CCCP, trialled at concentrations of 10, 12 and 48 µM, was added to the

assay cuvette followed by sodium formate at 160 mM once accumulation reached a stable level.

Fluorimetric measurements were performed at 37°C using a Perkin Elmer LS 55 fluorescence spectrophotometer with slit width of 10 nm. The excitation and emission wavelengths for acriflavine were 416 nm and 514 nm, respectively. Data were analysed and plotted by Microsoft Office Excel 2007.

2.14 Fluorescein-5-maleimide labelling of exposed Cys residues

2.14.1 Solvent accessibility analysis of QacC residues

Fluorescein-5-malemide was added to 125-250 µg of membrane fractions (Section 2.7.3) to a final concentration of 0.50 mM and incubated at RT for 30 minutes. The labelling reaction was stopped by the addition of 850 µl of NEM-quenching buffer (Table 2.1). Following an incubation of 1 hour at 4°C, samples were centrifuged (9,600 X *g* for 20 minutes at 4°C) to remove debris. Labelled QacC and QacC variant proteins were then purified by the addition of 100 µl of ProBond™ Nickel-Chelating Resin (Life Technologies) and incubated for 2 hours at 4°C. The resin beads were washed 4 x with 750 µl of bead wash buffer (Table 2.1)

Elution of proteins was carried out by the addition of 60 µl of 400 mM imidazole in 2X SDS-PAGE sample buffer (Table 2.1) and incubated at RT for 30 minutes before being placed at 37°C for 30 minutes. Samples were resolved using SDS-PAGE with 15% acrylamide resolving gels (Section 2.8.1) and visualised using a Bio-Rad Gel Doc EZ imager (Bio-Rad). Following visualisation gels were stained with Coomassie Brilliant Blue R-250 solution (Section 2.9.1). The degree of fluorescein-5-malemide labelling per sample was calculated and expressed as a percentage of the highly reactive Glu25Cys mutant protein.

To determine the effect of ethidium and pyronin Y binding on the labelling of the QacC Cys mutant proteins by fluorescein-5-maleimide, membrane fractions (Section 2.7.3) containing QacC mutant proteins were first incubated at 37°C for 5 minutes in the absence or presence of 1 mM ethidium or up to 1.5 mM pyronin Y; labelling of QacC Cys mutant proteins was performed as described above.

2.14.2 Fluorescein-5-maleimide and eosin-5-maleimide labelling of QacC derivatives from whole cells

E. coli TOP10 cells harbouring plasmids of interest were grown in overnight LBAmp cultures from freshly transformed colonies. Cells were diluted 1:20, grown till OD_{600}=0.55, induced with 0.02% L-arabinose and grown for a further 2 hours. One ml of cells were pelleted at 9,600 x *g* for 3 minutes at RT and washed once with 15 mM Tris-HCl (pH 7.5) and 10% [v/v] glycerol, pelleted and resuspended in the same buffer. Fluorescein-5-maleimide or eosin-5-maleimide was added to a final concentration of 0.25 mM and cells incubated at 37°C for 25 min. Labelled cells were pelleted by centrifugation (9,600 x *g* for 5 minutes at RT) and washed once with 15 mM Tris-HCl (pH7.5) and 10% [v/v] glycerol, centrifuged again at 9,600 x *g* for 5 minutes at RT and resuspended in 850 µl of NEM-quenching buffer (Table 2.1). Solubilisation of QacC was carried out by the addition of 1% [w/v] SDS and 0.5% [w/v] DDM. Following an incubation time of 1 hour at 4°C, QacC was purified using the ProBond^{TM} Nickel-Chelating Resin, following steps outlined in Section 2.14.1.

2.15 Measuring QacC–rhodamine 6G binding affinity by fluorescence polarisation

Fluorescence polarisation-based binding reactions were carried out at 25°C in QacC dialysis buffer (Table 2.1) (Section 2.7.4) supplemented with 0.05% [w/v] DDM. Rhodamine 6G was dissolved in the dialysis buffer and used at a final concentration of 4 nM. Readings were carried out on titrations of concentrated QacC protein and fluorescence polarisation was measured using a Perkin Elmer LS 55 fluorescence spectrophotometer with a slit width of 5 nm. Samples were excited at 525 nm and emissions were read at 555 nm. Curves were fitted to plots of the polarisation data from saturation binding experiments. To determine K_D values the following equation was used:

$$P = \frac{(P_b - P_f)\,[\text{QacC}]}{K_D + [\text{QacC}]} + P_b$$

Where P is the polarisation measure at a given QacC concentration [QacC], P_b is the maximum polarisation of the QacC-bound rhodamine 6G, i.e., after complete saturation, P_f is the polarisation constant for the interaction. Since saturation binding experiments used rhodamine 6G at a concentration 500 times lower than the K_D value for the QacC-rhodamine 6G interaction, the concentration of free QacC protein was assumed to be the same as the total QacC protein concentration. All data were analysed and plotted by Microsoft Office Excel 2007.

2.16 Cross-linking of QacC monomers

2.16.1 Formaldehyde crosslinking

Overnight cultures of *E. coli* TOP10 cells expressing the pBAD based-QacC and QacC Ser90 and Gly97 mutant, were diluted 1:20 grown till OD_{600}=0.55, induced with 0.02% L-arabinose and grown until they reached an OD_{600}=1.0. Cells were harvested by centrifugation and washed once in ice-cold 10 mM KPO_4 buffer (pH 7.2) (Table 2.1) and then resuspended in the same buffer. Formaldehyde 37% [w/v] was added to a final concentration of 0.5 and 1% and samples were incubated for 1 hour at RT. Cross-linked samples were pelleted at 17,000 x *g* for 1 minutes at RT, washed once in ice-cold 10 mM KPO_4 (pH 7.2) and resuspended in 100 µl of 2X sample buffer (Table 2.1). Samples were subjected to SDS-PAGE (Section 2.8.1) and Western immunoblotting (Section 2.9.2 and 2.9.3).

2.17 Bioinformatic analyses

2.17.1 DNA and amino acids sequence alignments

All nucleotide and amino acid sequences were obtained from the National Center for Biotechnology Information (NCBI) website (http://www.ncbi.nlm.nih.gov/). Alignment and manipulation of DNA sequences, of plasmids and sequenced fragments were performed with Sequencher™ 4.9 (GeneCodes) and viewed on CLC sequence viewer 6.8.2 (CLC bio). Multiple sequence alignments of DNA and amino acid sequences were generated using the

on-line version of ClustalW and ClustalW2 (Larkin et al., 2007; Thompson et al., 2002).

2.17.2 Software employed for QacC structural analysis

Protein data bank (PBD) files were viewed and analysed using Pymol (DeLano, 2002) and Chimera version 1.9 (Pettersen et al., 2004). QacC secondary structural analysis was carried out using TMHMM (Krogh et al., 2001), PHDhtm (Rost et al., 1996a) and TOPPRED (Claros and von Heijne, 1994). Two prediction algorithms were used for QacC *de novo/ab initio* modelling; QUARK server (Xu and Zhang, 2012) and the CABS-folds server (Blaszczyk et al., 2013). Homology modelling used MODELLER (Sanchez and Sali, 2000) and Phyre2 (Kelley and Sternberg, 2009).

2.17.3 Validation and assessment of QacC models

The overall stereochemical quality of each QacC model was assessed by generating Ramachandran plots using PROCHECK (Laskowski et al., 1996). Quality factors for each QacC model were calculated using Verify3D (Luthy et al., 1992) and Errat2 (Colovos and Yeates, 1993).

2.17.4 Other software

HeliQuest (Gautier et al., 2008) was used to construct the α-helical wheel representation of QacC, whilst the WebLogo 3.0 server (Crooks et al., 2004) accessed online at http://weblogo.berkeley.edu, was used to generate sequence logos.

CHAPTER 3 – QACC MODELLING

3.1 Introduction

The unique three-dimensional (3D) structure of a protein is determined by numerous factors including its amino acid sequence, hydrogen bonding, van der Waals forces and its interaction with the surrounding solvent/environment (Bradley et al., 2005; Floudas et al., 2006; Khoury et al., 2014). Although our knowledge pertaining to a protein's sequence and its effect on structure and folding continues to expand, structure prediction remains one of the most challenging fields of proteomics (Khoury et al., 2014). With the advent of large scale genome sequencing, the number of protein sequences have increased, however, experimental structural analysis of these sequences such as X-ray crystallography and NMR, is limited by factors such as time and cost, not to mention the experimental challenges which are inherent to such analyses (Cozzetto and Tramontano, 2008; Eswar et al., 2008). As a result of this, computational determination of protein structure has become an important tool for structural analysis, and can lead to the generation of hypotheses which can be further analysed experimentally.

Computational protein structure prediction can be divided into three approaches; comparative modelling, threading and *de novo* folding. Comparative modelling and threading generate protein models based on homology or analogy, by aligning the query protein sequence onto a solved template. When high sequence homology between a target and template exists, a high-resolution model can be generated utilising both approaches. However, all too frequently, especially with regards to MPs, no structural information (of a solved homologue) is available for the target protein, and in such cases structure prediction needs to be carried out *de novo*/*ab initio*. This process utilises protein like biases which have been established from solved structures and relies on the dogma that the native state of the protein is the thermodynamically most stable state, under the physiological conditions surrounding it, which also corresponds to the minimum free energy conformation (Bradley et al., 2005; Cozzetto and Tramontano, 2008).

The initial stages of protein structure prediction starts with comparative sequence analysis to identify related sequences which homology modelling can be based on. Once possible homologues are identified, a pairwise sequence alignment of the target protein and homologues is used to generate all possible matches between the proteins, with the highest scoring alignment identifying the best homologue for modelling (Punta et al., 2007). The most commonly used matrix used for protein sequence alignment is known as the blocks of amino acids substitution matrix (BLOSUM), which is a substitution matrix used to score the alignment between different protein sequences, based on differences in conserved, ungapped regions of a protein family (Henikoff and Henikoff, 1992). Of the various BLOSUM matrices, the BLOSUM62 score matrix is used on the NCBI website to score pairs of aligned sequences (BLASTp). This scoring matrix has been shown to be the most effective in detecting known members of a protein family, and is based on sequences that match with at least 62% identity (Eddy, 2004).

Sequence analysis, based on homology, also allows for the possible identification of structurally and functionally relevant residues. Conserved amino acid residues, identified at a specific position within members of a family of related proteins, can be an indication of selective evolutionary pressures at that site. These can then be marked as potential targets for both functional and structural analyses. Variability between sequences of members of the same family on the other hand, can conceivably be used to identify sub-families, or different functional characteristics, such as different binding sites (Punta et al., 2007). Variability between residues can also aid in the prediction of a protein's secondary structure, as high sequence variability and insertions have been observed to occur in exposed regions and can be used to identify residues that lie on the surface of the polypeptide, as opposed to those that lie buried within the protein (Zvelebil et al., 1987).

Secondary protein structure prediction can be divided into two aims, the first looks at the whole-protein prediction, and is used in structural genomics projects to specifically select or exclude a protein from a target list. The second is based on a per-residue prediction, which examines each residue and determines which

residues span the membrane and which make up loop regions (Punta et al., 2007). This last aim is specific for MPs and varies greatly for water-soluble proteins, as the environment in which integral proteins lie is largely lipophilic and has a limiting potential for hydrogen-bonding (Forrest et al., 2006).

Early approaches to secondary structure prediction for MPs utilised a "propensity scale" which was based on the properties of the amino acids found in the TMS. An example of such a plot, which uses a hydropathy scale composed of the hydrophilic and hydrophobic properties of each of the 20 amino acid side-chains, is the Kyte-Doolittle plot (Kyte and Doolittle, 1982). Although useful and simplistic, more research and the addition of new structural data has led to this method evolving with more modern approaches using Hidden Markov Models (HMM) and neural networks. Suitable for representing sequence heterogeneity, HMM look at global patterns in the structure, whereas neural networks focus on local sequence patterns, which are composed of several residues and are thus capable of recognising subtle patterns such as the length of a TMS or the amphipathicity in helices (Punta et al., 2007). Assessment of these models has shown that the prediction accuracy varies from model to model, with the best models achieving a prediction accuracy of about 80%, when using homologous sequence information (Martin et al., 2006).

As with the determination of the secondary structure, 3D homology modelling relies on the availability of template sequences which are highly similar to the target sequence. As mentioned above, with the number of solved MP structures being so low, there is a lack of available modelling templates (Section 4.1) to use. Thus, the challenge with 3D modelling is to generate a good model using template structures which share low sequence identity. A number of methods exist to generate 3D models, of these the MODELLER program (Sali and Blundell, 1993) has been shown to generate statistically better results in comparison to others (Reddy Ch et al., 2006). Once a 3D model has been obtained, it needs to be evaluated and validated. This is best carried out by implementing a number of validation protocols, as each protocol may focus on a different aspect of the model, and then compare all results obtained. Of the various protocols available, PROCHECK

(Laskowski et al., 1996) which includes the Ramachandran validation plot (Ramachandran et al., 1963), Verify 3D (Luthy et al., 1992) and Errat2 (Colovos and Yeates, 1993) are able to provide an overall assessment of a predicted structure.

In this chapter, both secondary and 3D structures of QacC were developed using both the *de novo*/*ab initio* approach and by comparative modelling. Each model was then validated by a number of tests to assess their overall quality and to determine how well they represent the target protein.

3.2 Secondary structural analysis of QacC

The QacC input sequence used for structure prediction was that carried on the staphylococcal plasmid pSK89 (Littlejohn et al., 1991; Paulsen et al., 1995) (Accession number: P14319) (Section 1.6.1). Two methods were used to obtain the QacC secondary structure. The first of these utilised three popular topology prediction methods; TMHMM (Krogh et al., 2001), TOPPRED (Claros and von Heijne, 1994; von Heijne, 1992) and PHDhtm (Rost et al., 1996b) (Section 2.17.2), and were used in their single-sequence mode, thus not including information from homologous proteins, with the QacC input sequence used as the search parameter (Figure 3.1). All three prediction methods identified four possible TMS. Results from the TMHMM model identified 74.3 amino acids to be located within a TMS and gave a total probability of 0.6077 for the N-terminus to lie on the cytoplasmic side of the membrane. Using the GES-scale (Engelman et al., 1986), TOPPRED identified four putative TMS, allocating a certainty score of 1.157 (TMS1), 1.812 (TMS2), 1.945 (TMS3) and 2.092 (TMS4) to each segment. Scores with a hydrophobicity value of 1.0 or above are deemed a “certain” TMS (Zhang and Skolnick, 2004). In contrast to TMHMM, TOPPRED predicted the N-terminal to reside in the periplasmic as opposed to the cytoplasm, although the model also predicts the first amino acid to reside in the TMS. Lastly, results of the PHDhtm model were in accordance with TOPPRED, in predicting the first loop to be located in the periplasm. The best model and the reliability of the model are given a value of 1, which on a scale of 0 (low) to 9 (high), is very low. Each predicted membrane helix

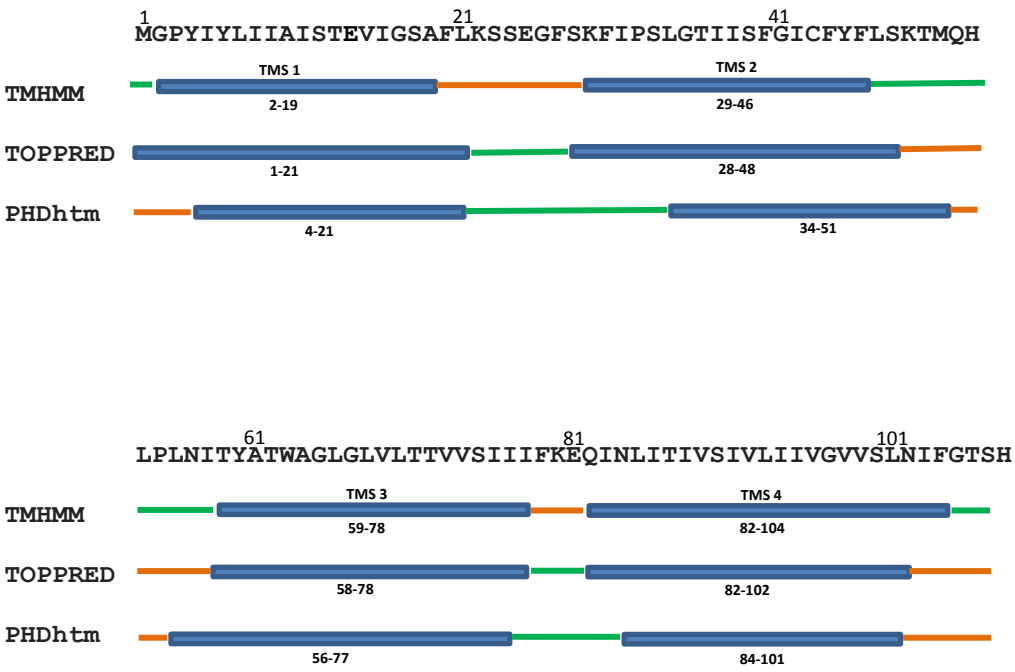

Figure 3.1: Prediction of TMS within QacC using TMHMM, TOPPRED and PHDhtm.

Graphical representation of the predicted locations of the TMS in QacC generated by TMHMM, TOPPRED and PHDhtm. TMS predicted using a single-sequence mode with default settings. The QacC input sequence (accession number: P14319.1) is given above the location of the predicted TMS (blue bars) with the starting and end residue number indicated beneath each segment. Green bars indicate loops located in the cytoplasm; orange bars denote loops in the periplasm.

is sorted by strength, (strongest indicating best prediction), with TMS3 listed as the strongest followed by TMS4, TMS2 and lastly TMS1.

The second method employed to generate the secondary structure of QacC was based on a consensus prediction method. This method predicts TMS based on an alignment generated between the input sequence and other members of a specific family (Bernsel et al., 2009; Cuff et al., 1998). In order to obtain the secondary structure of QacC using this method two different approaches were used and compared; Jpred3 (Cole et al., 2008) and TOPCONS (Bernsel et al., 2009). Using the same QacC input sequence as before, predictions made with both methods were produced with all parameters left at the default settings (Figure 3.2 and 3.3). Results generated by Jpred3 incorporated the Jnet 2.0 algorithm (Cole et al., 2008), whilst results from TOPCONS combined an arbitrary number of topology predictions and used a "majority-vote" principle, which takes into account the level of agreement between each prediction algorithm used to generate one consensus prediction (Bernsel et al., 2009). Both methods predicted a secondary structure of QacC with four TMS, with TOPCONS identifying the N-terminus to lie in the cytoplasm. Whilst results generated by Jpred3 also predict certain residues to form a beta sheet structure, as determined by the Jpssm scoring matrix, the final allocation of these residues, determined by the Jnet algorithm, assigns these to an alpha helix. According to this algorithm, which also predicts relative solvent accessibility using various cut-offs (0, 5 and 25%) and assesses whether each residue is buried or exposed at each of these cut-offs, the predicted QacC model contains a large number of buried residues which have less than 25% solvent accessibility.

Analysis of QacC used in this study, used a derivative with a number of tags added to the C-terminus to aid in purification (see Chapter 4) therefore the addition of these tags was evaluated to assess their possible effect on QacC structure. To this end, an input sequence of QacC was used which included both the 1D4 and Hisx6 tags. Results generated by TMHMM (Krogh et al., 2001) and TOPPRED (Claros and von Heijne, 1994) were in accordance with those which were

Figure 3.2: Secondary structure prediction of QacC using Jpred3.

The top row shows the QacC input sequence with the position number given in the row above. This sequence was used in a multiple sequence alignment, from which the secondary structure was then predicted using the Jnet neural network. The Jnet algorithm uses the position-specific scoring matrix (Jpssm) and the hidden Markov model (Jhmm) to predict protein secondary structure. The relative predicted position of each amino acid is denoted by either: (H) alpha helix, (E) beta sheet or (-) random coil. Solvent accessibility is predicted by Jnet using 0% (Jnet_0), 5% (Jnet_25) and 25% (Jnet_5) cut-offs, with residues which are predicted to be buried (B) or exposed (-) shown at each of the three relative solvent accessibilities (prediction accuracy shown by Jnet Rel, ranging from 0 to 9, with higher numbers representing higher accuracy). Multiple sequence alignment results generated by Jpred3 are listed in Appendix 1.

A

Predicted TMS postions:

SCAMPI-seq	1. 3-23, 2. 30-50, 3. 59-79, 4. 85-105
SCAMPI-msa	1. 3-23, 2. 31-51, 3. 59-79, 4. 85-105
PRODIV	1. 3-23, 2. 28-48, 3. 58-78, 4. 83-103
PRO	1. 3-23, 2. 28-48, 3. 60-80, 4. 85-105
OCTOPUS	1. 3-23, 2. 31-51, 3. 59-79, 4. 85-105

B

Consensus prediction:

TOPCONS **1.** 3-23, **2.** 30-50, **3.** 59-79, **4.** 85-105

C

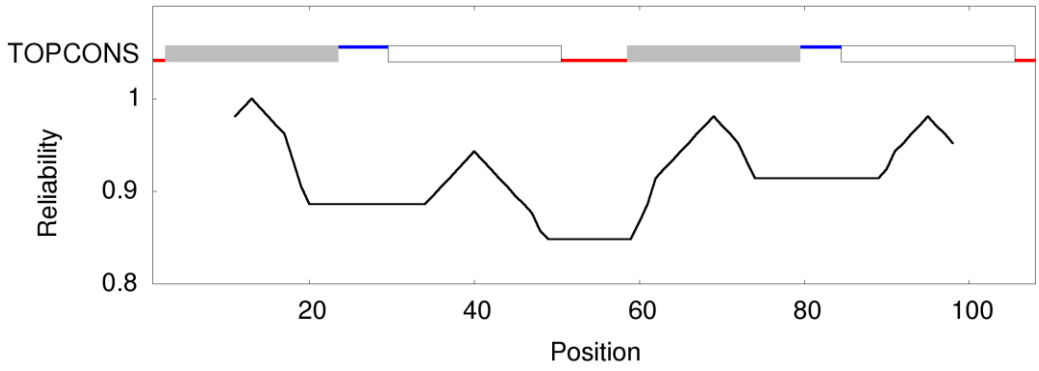

Figure 3.3: Secondary structure prediction of QacC using TOPCONS.

A QacC homology sequence profile from BLASTp was used as an input for all other methods (SCAMPI, OCTOPUS, ΔG-Scale, ZPRED and PRO/PRODIV-TMHMM). (A) Individual results are produced and consolidated to generate the consensus predicition. The name of the method is followed by TMS number showing the amino acids within that segment. (B) The concensus prediction showing the final allocation of residues to TMS. (C) A graphical representation of the TOPCONS results for QacC. Bars indicate the predicted TMS; grey bar orientation - TMS IN → OUT, white bar - TMS OUT → IN, red lines indicate loops on the inside, blue lines loops on the outside. Residue positon is indicated on the botton of the graph, whilst the reliability of residue assignment is given on the left hand side of the graph.

based on a QacC sequence lacking the tags (data not shown). Further assessment carried out with Jpred3 (Cole et al., 2008) confirmed that the addition of the 1D4 and Hisx6 tags did not alter the predicted secondary structure of QacC (Figure 3.4).

3.3 De novo modelling of QacC using QUARK and CABS

To date, no high-resolution crystal structure exists of QacC and there is no published 3D model. Thus, in order to produce a predictive 3D model of QacC, and given the lack of any structural data, the *de novo*/*ab initio* approach was undertaken. Of the various methods available for *de novo*/*ab initio* modelling two prediction algorithms were chosen to model QacC; the QUARK computer algorithm (Xu and Zhang, 2012) and the CABS-folds server (Blaszczyk et al., 2013) (Section 2.17.2). Both web servers ranked high in Critical Assessment of Protein Structure Prediction (CASP) experiments and as such were utilised in this study (Blaszczyk et al., 2013; Jayaram, 2013). CASP experiments take place every two years and serve as independent assessors of state of the art in *de novo* protein modelling thus solving protein structure from the amino acid sequence (Moult, 2005; Moult et al., 2014). For both platforms, the QacC input sequence (Section 3.2.1) was submitted in FASTA format, with all settings left at default. Both servers returned a number of possible models; QUARK produced 10 models and CABS-produced 9 models. The QUARK server included a template modelling score (TM-score) for the best model, whilst the CABS-folds server included clustering data, the C^{α} root-mean-square deviation (RMSD) between predicted models as well as the C^{α} global distance test, also written as GDT_TS to represent total score, between predicted models. Both platforms include a secondary structure prediction, whilst the QUARK server also includes a predicted solvent accessibility model. Of the 10 models predicted, Figure 3.5 shows model 1, which was identified as the best model with a TM-score of 0.51 ± 0.08, whilst the secondary structure prediction and solvent accessibility model, supporting results obtained in Section 3.2, are shown in Figure 3.6.

Results generated by the CABS-fold web server predicted 9 possible models of QacC, with the best model (model 1) shown in Figure 3.7. The CABS-fold server uses

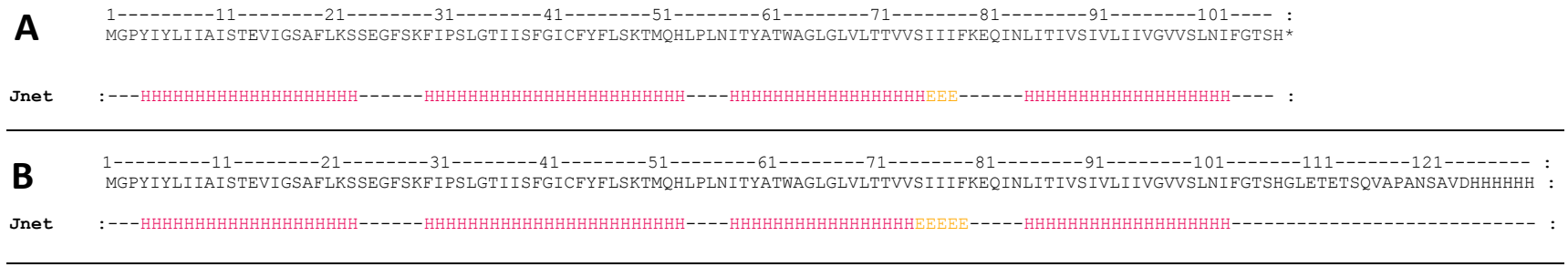

Figure 3.4: Comparison of the secondary structure prediction of QacC and QacC/1D4/Hisx6 using Jpred3

Summary results of the predicted secondary structure of QacC and QacC/1D4/His6 as generated by the Jpred3 server. (A) Top row shows the QacC input sequence with the amino acid position number displayed in the row above (B) Is that of the QacC/1D4/Hisx6 protein. Both sequences were used in a multiple sequence alignment, from which the secondary structure was then predicted using a neural network called Jnet. The relative predicted position of each amino acid is denoted by either: (H) alpha helix, (E) beta sheet or (-) random coil.

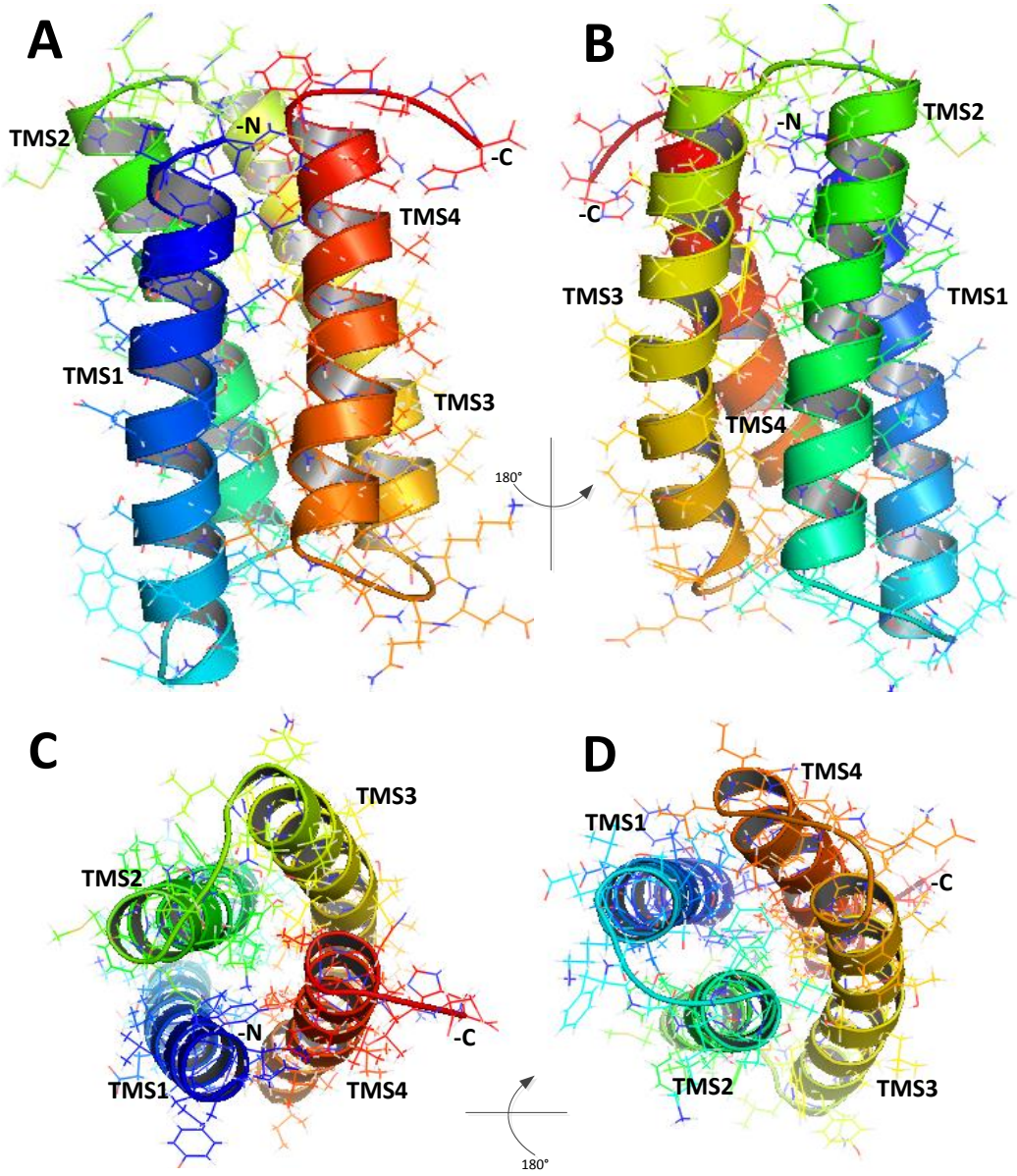

Figure 3.5: The 3D model of QacC as generated by QUARK

Cartoon representation of the predicted 3D model of QacC with the highest TM-score. TMS are indicated by different colours (TMS1-blue; TMS2-green; TMS3-yellow; TMS4-orange). Of the ten models of QacC that were generated with the QUARK algorithm, model 1, represented as the best, is shown. The model is shown in side view (A and B) as well as in the top (periplasmic) and bottom (cytoplasmic) view (C and D). The estimated TM-score of model 1 is 0.51±0.08. Models with a TM-score of > 0.5 are predicted to have a correct fold.

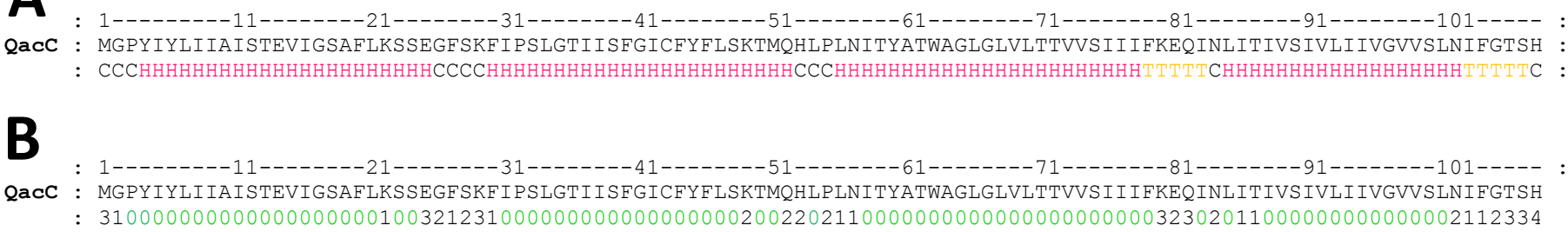

Figure 3.6: Predicted secondary structure and solvent accessibility of QacC as generated by QUARK

(A) Shows the QacC input sequence with residue position number given above. This sequence was used as an input sequence on the QUARK online *ab initio* protein structure prediction server. Results obtained included the predicted secondary structure of QacC, given in the top row, with each residue being assigned to a (H) alpha helix, (T) beta turn or (C) random coli. (B) Predicted solvent accessibility was based on a score for each residue, ranging from 0 to 9, with 0 indicating buried residues and 9 exposed residues.

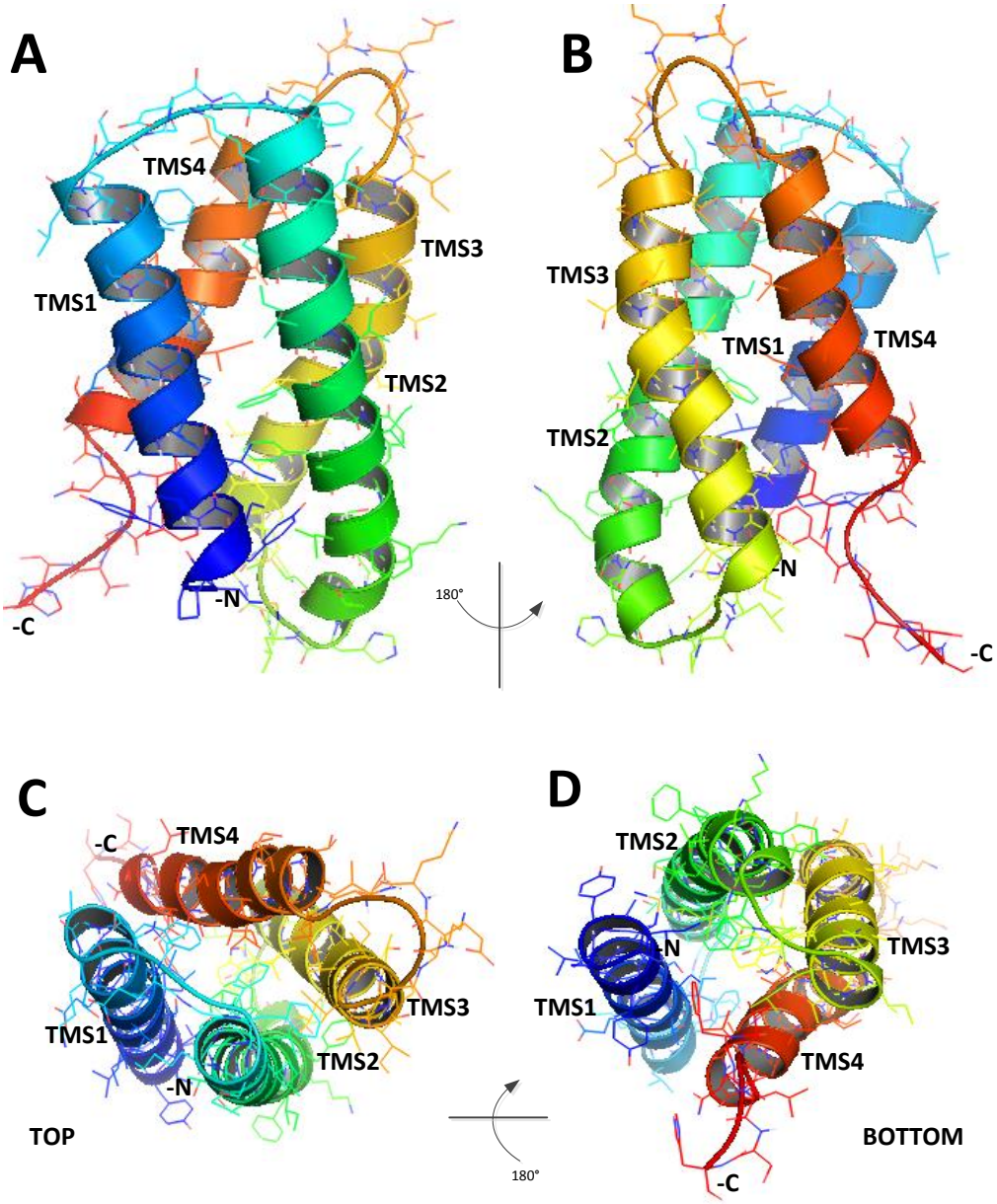

Figure 3.7: The 3D model of QacC as generated by the CABS-fold server

Cartoon representation of the best model of QacC as predicted by the CABS-fold server. TMS are indicated by different colours (TMS1-blue; TMS2-green; TMS3-yellow, TMS4-orange). (A) Model as shown from the front. (B) Rear (rotated at 180°C from panel A). (C) Top (periplasmic). (D) Bottom (cytoplasmic).

a number of methods to evaluate the predicated models and rank them from best to worst. Such models are identified by cluster density, size and average cluster RMSD, which for model 1 were 24.4, 61 and 2.5, respectively. Clustering of protein models is carried out to separate a set of protein models into groups designated clusters. From these a representative is chosen, which is always the model that has a minimal dissimilarity to all models in a cluster. The clusters are then ranked according to cluster density values, with the most dense (model 1) to the least dense (model 9) (Blaszczyk et al., 2013). The RMSD values, calculated on the C^{α} atoms, were used to measure the average distance between the atoms of the predicted models, and thus to measure the similarity in the 3D structures. A large value is indicative of dissimilar models, whilst a value close to zero indicates identical conformation (Maiorov and Crippen, 1994). The last measure of similarity between predicted models is the C^{α} GDT_TS, which counts the number of C^{α} pairs which have a distance of <1, 2, 4 and 8 Å following an optimal superposition (Zemla, 2003). Results are given in cut-offs which range from 0.05 to 1.0 nm.

The two best predicted QacC 3D models as generated by QUARK (model 1) and the CABS-fold server (model 1) were evaluated by PROCHECK (Laskowski et al., 1996), Verify3D (Luthy et al., 1992) and Errat2 (Colovos and Yeates, 1993), see section 3.5.

3.4 Homology modelling of QacC using MODELLER and Phyre2

Homology modelling of QacC was carried out using the programs MODELLER (Sanchez and Sali, 2000) and Phyre2 (Kelley and Sternberg, 2009) (Section 2.17.2). To determine a model of QacC using MODELLER an initial search of the database of known protein structures (Protein Data Bank, PDB) was peformed using the QacC input sequence as the query (Section 3.2). This resulted in four structures being identified as homologues of QacC (Table 3.1), all of which were of the EmrE protein from *E. coli* (Section 1.5.1). For each of those, a FASTA format sequence was downloaded and a multiple sequence alignment was generated between these and QacC. This alignment was then used by the MODELLER to predict a model of QacC (Figure 3.8). As can be seen in this figure, the predicted model of QacC was

Table 3.1: PDB search results using the QacC protein sequence[a]

PDB entry	Name of Entry	Res. (Å)	Reference
2I68	Cryo-EM based theoretical model structure of transmembrane domain of the multidrug-resistance antiporter EmrE from *E. coli*	7.5	(Fleishman et al., 2006)
3B5D	EmrE multidrug transporter in complex with TPP, C2 crystal form	3.8	(Chen et al., 2007)
3B61	EmrE multidrug transporter, apo crystal form	4.5	(Chen et al., 2007)
3B62	EmrE multidrug transporter in complex with P4P, P21 crystal form	4.4	(Chen et al., 2007)

a: Genbank accession number: P14319.1

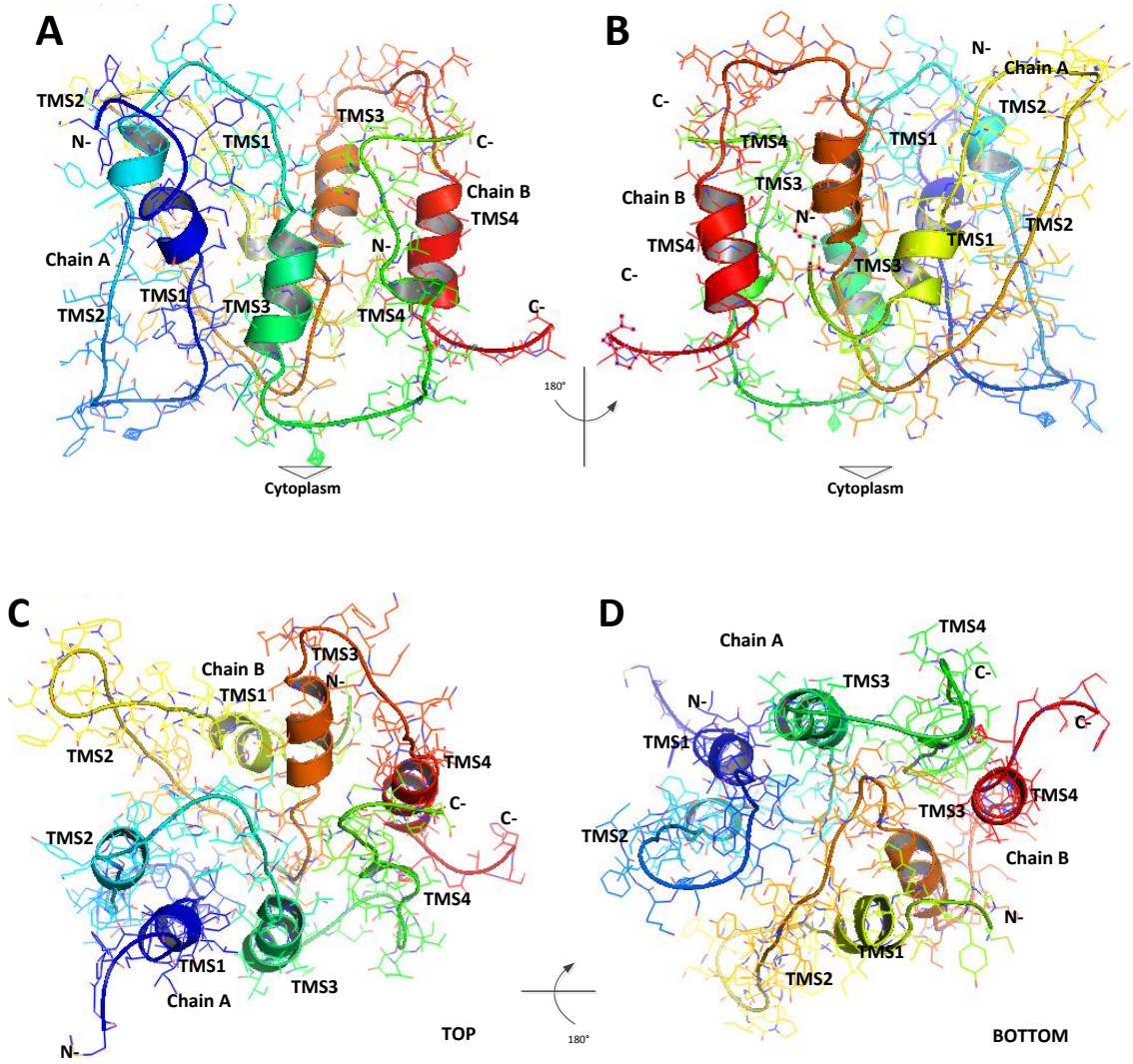

Figure 3.8: The 3D QacC prediction model generated by the MODELLER program

Ribbon drawings of the predicted model of QacC as generated by the MODELLER program. Chains A and B, representing the dimeric form of QacC, are shown with each TMS segment depicted with a different colour (Chain A: TMS1-blue; TMS2-light blue; TMS3-green; TMS4-light green. Chain 2: TMS1-yellow; TMS2-light yellow; TMS3-orange; TMS4-red). Sections of TMS depicted as ribbons indicate areas of homology with solved structures of EmrE, whereas those sections that have a lack of homology are depicted as string (TMS4 in chain A and TMS2 in chain B). (A) Model as shown from the front. (B) Rear (rotated 180° from panel A). (C) Top (periplasmic). (D) Bottom (cytoplasmic).

determined to be in a dimeric form, which is also the predicted oligomeric structure of EmrE (Chen et al., 2007; Ubarretxena-Belandia et al., 2003; Yerushalmi et al., 1996) (Section 1.5.1).

In addition to MODELLER, the Phyre2 web server was used for the prediction of a 3D model of QacC. As with previous protein structure prediction servers, the same QacC input sequence was used to generate all possible models. Of the five Phyre2 predicted models two returned high confidence scores (homologous scores). Of these, the model with the highest confidence value was shown to be based on template 1S7B (PDB ID code 2I68) which according to the PBD server is a retracted structure of EmrE (Chang et al., 2006) due to inconsistencies in the computer method used to analyse the data obtained. Therefore, the second model predicted by the Phyre2 server, the cryo-EM structure of EmrE (Fleishman et al., 2006), was used for the 3D structure of QacC (Figure 3.9). Similar to the model generated by MODELLER, the overall stereochemical quality of the Pyre2 predicted 3D model of QacC was assessed using PROCHECK (Laskowski et al., 1996), Verify3D (Luthy et al., 1992) and Errat2 (Colovos and Yeates, 1993) (Section 3.5).

3.5 Evaluation of predicted QacC models

The 3D QacC models predicted by *de novo* methods or comparison modelling were assessed by a Ramachandran plot, carried out by PROCHECK (Laskowski et al., 1996), Verify3D (Luthy et al., 1992) and Errat2 (Colovos and Yeates, 1993) (Section 2.17.3). Assessment was performed to assess the stereochemical quality and the side-chain environment of each model. Stereochemical quality assessed by PROCHECK used the Ramachandran plot to display the phi (ϕ) and psi (ψ) torsion angles for all residues in the predicted structures. The Ramachandran plot provides an overview of allowed and disallowed regions of torsion angle values, which serve as important indicators of the quality of the proteins 3D structure (Kleywegt and Jones, 1996). Models which are shown to have a high stereochemical quality are those which have 90% of the residues in the “core” region, which represents the most favourable combination of the ϕ and ψ (Morris et al., 1992). Verfiy3D was used to assess the compatibility between the amino acid sequence in the predicted

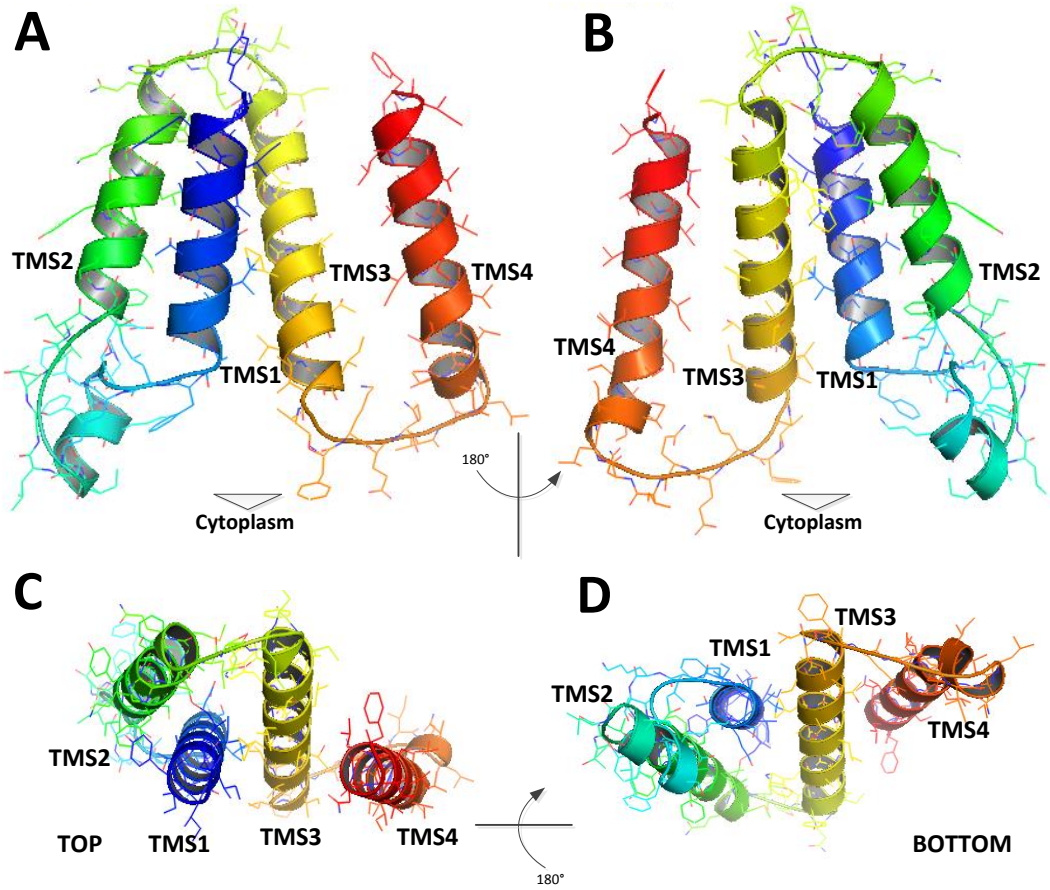

Figure 3.9: The 3D QacC prediction model generated by the Phyre2 web server

Ribbon drawings of the predicted model of QacC as generated by the Phyre2 web server using the EmrE PBD 2I68 structure. Each TMS is shown in a different colour (TMS1-blue; TMS2-green; TMS3-yellow; TMS4-red/orange). The predicted 3D structure is shown as a monomer only. Sections of the TMS depicted as ribbons indicate areas of homology with the structure of EmrE, sections not predicted due to a lack of homology with EmrE are depicted as string (A) Model as shown from the front. (B) Rear (rotated 180° from panel A). (C) Top (periplasmic). (D) Bottom (cytoplasmic).

model and the environment of the amino acid side chains. Compatibility scores above zero corresponded to side-chain environments which were acceptable (Luthy et al., 1992), with results given shown as a percentage of the acceptable side-chain environments. Lastly, Errat2 was used to assess the overall quality factor of the predicted models by looking at the distribution of non-bonded atoms with respect to one another in each protein model; higher scores (>50) represent higher quality structures (Colovos and Yeates, 1993).

The results of the assessment for each predicted QacC 3D model are summarised in Table 3.2. For comparison, predicted structures of EmrE, modelled using the same algorithms and web servers as QacC, were also assessed with the same validation procedures (Table 3.2). From this table it is apparent that models of QacC generated by the *de novo* methods (QUARK and CABS-fold server) were assessed to be of higher quality than those based on comparative modelling, with the QUARK algorithm predicting a model of QacC which satisfied all the validation criteria on the basis of PROCHECK, Verify3D and Errat2 (Figure 3.10, 3.11 and 3.12).

3.6 Model-template: assessment of similarity

To determine the structural similarity of the QacC models as obtained by the MODELLER program and the Phyre2 server, with the template used to generate said models, the TM-score and RMSD were calculated, and an alignment between the model and template made. This was performed to ascertain whether the predicted models share the same folds as the template and to determine structure similarity. Using the TM-align algorithm (Zhang and Skolnick, 2005) each predicted 3D QacC model was compared with the template used to generate each model, which in this study was the solved EmrE structure (PBD id code: 2I68) (Fleishman et al., 2006).

Alignment of the MODELLER predicted QacC 3D model with the EmrE template (Figure 3.13) produced a final TM-score of 0.518 (chain A) and 0.651 (chain B) indicating that both structures share the same fold, whilst a RMSD value of 2.97 demonstrates that the structures also share structural similarity. The results of the Phyre2 server predicted QacC model alignment with the EmrE template

Table 3.2: Validation results for QacC models

	Server	Ramachandran plot (%)[a]	Verify 3D (%)[b]	Errat2 (%)[c]
QacC	QUARK	95.7	82.82 (P)	87.88
QacC	CABS	90.4	72.22 (W)	96.97
QacC	MODELLER	63.8	17.51 (F)	29.81
QacC	Phyre2	87.8	18.63 (F)	72.04
EmrE	QUARK	93.4	100 (P)	100
EmrE	CABS	89	73.87 (W)	100
EmrE	Phyre2	85.7	7.8 (F)	58.06

a: Percentage of residues with ϕ, ψ conformation in the "most favoured" regions in the Ramachandran plot.
b: Percentage of residues that had an averaged 3D-1D score of >0.2 (P- Pass, W-Worth pursuing, F-Fail).
c: Expressed as a percentage of the protein for which the calculated error value falls below the 95% rejection limit. Good, high resolution structures generally produce values around 95% or higher. For lower resolutions (2.5 to 3 Å) the average overall quality factor is around 91%.

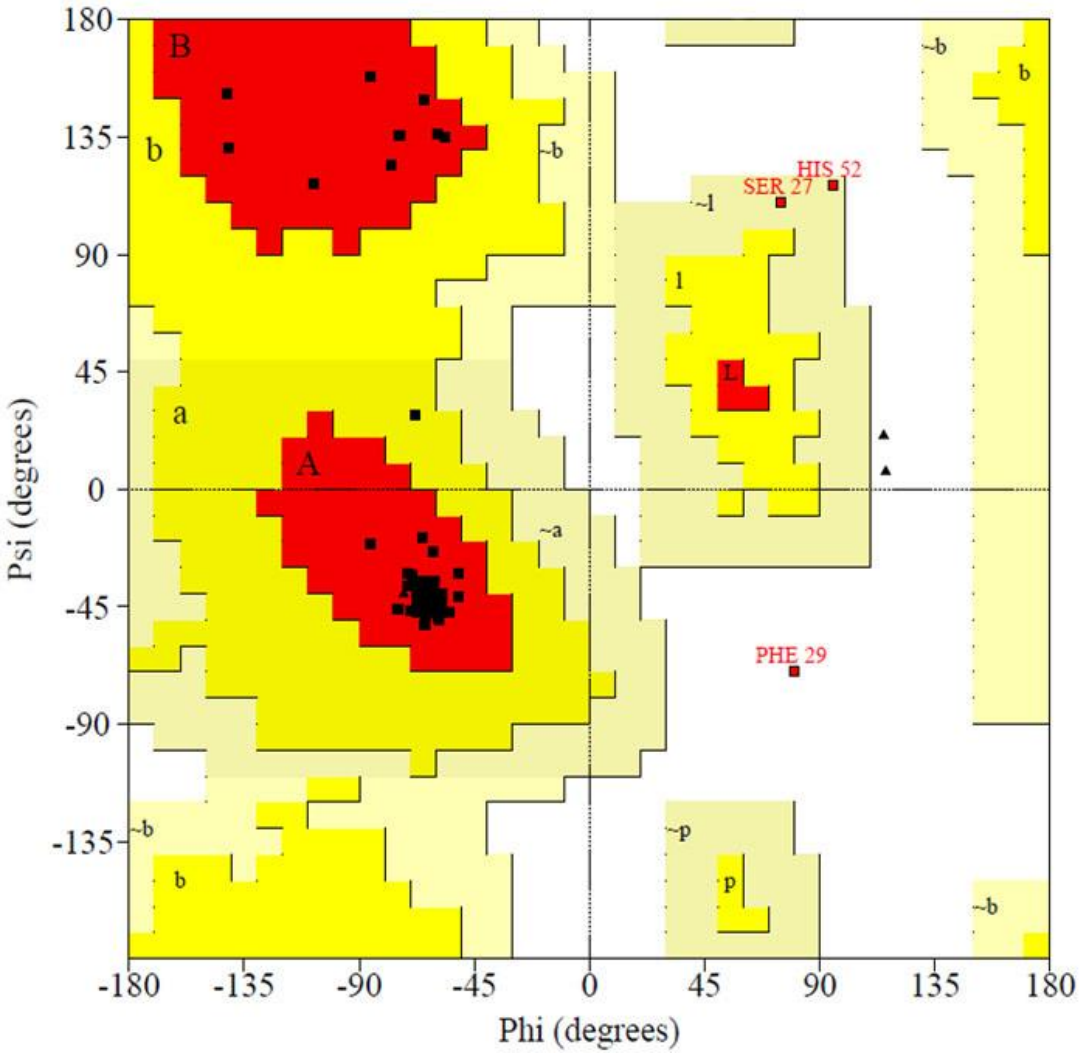

Figure 3.10: QacC 3D model generated by the QUARK algorithm assessed with PROCHECK.

Ramachandran plot of the QacC protein derived from PROCHECK. The ϕ/ψ angles of 95.7% residues are in the most favoured regions (red regions), 1.1% residues in the additional allowed region (brown regions), 2.1% residues in generously allowed region (yellow regions) and 1.1% residues are in the disallowed region (white regions).

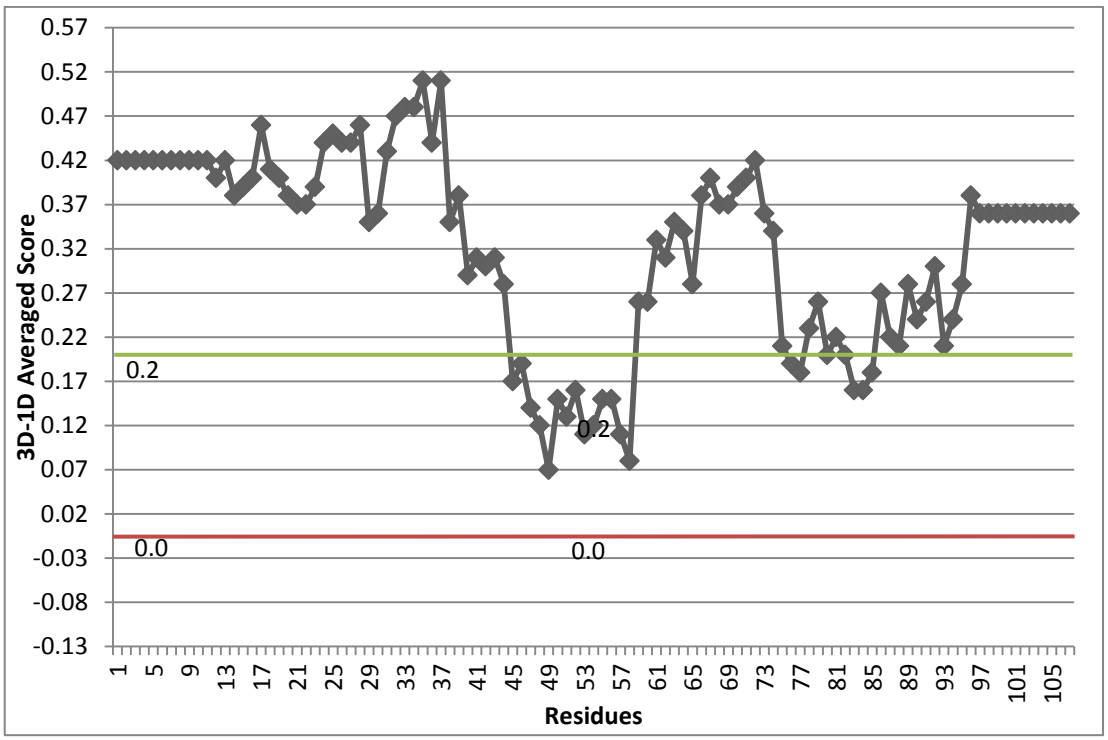

Figure 3.11: Verify3D plot of the 3D model of QacC with residues having the highest 3D-1D average score (0.51)

Verify3D plot showing the vertical axis, which represents the average 3D-1D profile score for all residues, and the horizontal axis indicating residue position. The score of the first 9 and last 9 residues is not allocated. The plot shows that 82.41% of residues had an average 3D-1D score of > 0.2 (green line), indicating that the model is compatible with its surrounding environment.

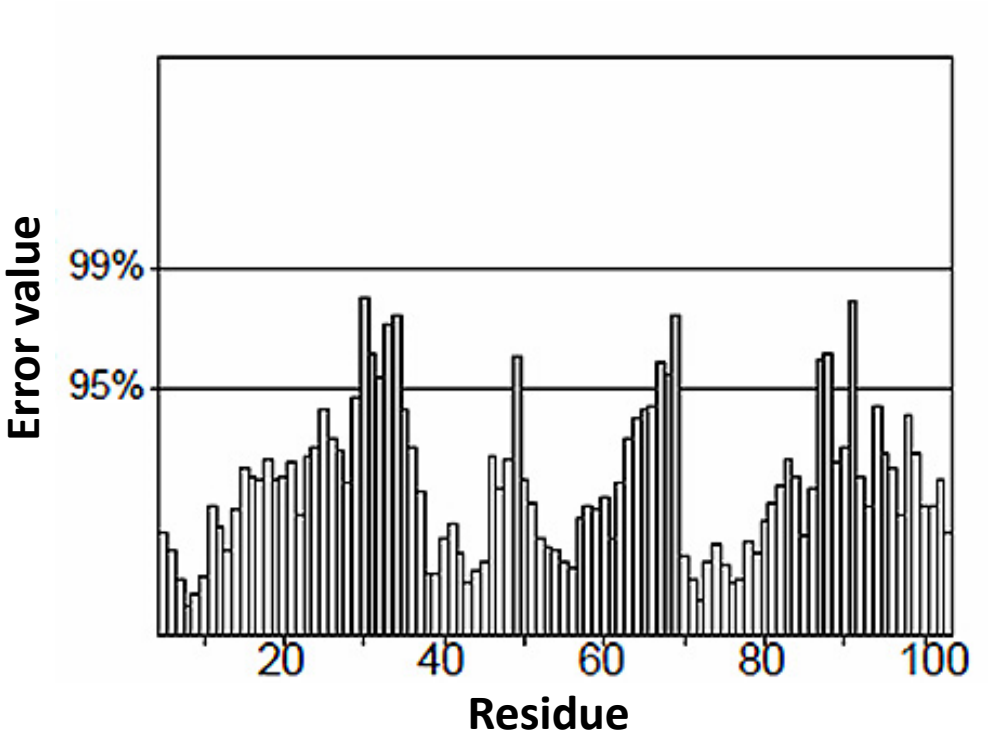

Figure 3.12: QacC 3D model generated by the QUARK algorithm assessed with Errat2.

Errat2 plot for the QUARK predicted QacC model, with the y-axis showing two lines indicating the confidence (99% and 95% confidence levels) with which it is possible to reject regions that exceed the error value. The overall quality factor of the residues of the QacC protein as predicted by the QUARK algorithm (87.89%), expressed as a percentage of the protein for which the calculated error value falls below the 95% rejection limit. The darker columns identify problem areas which appear in the predicted TMS2, TMS3 and also in TMS4 of QacC.

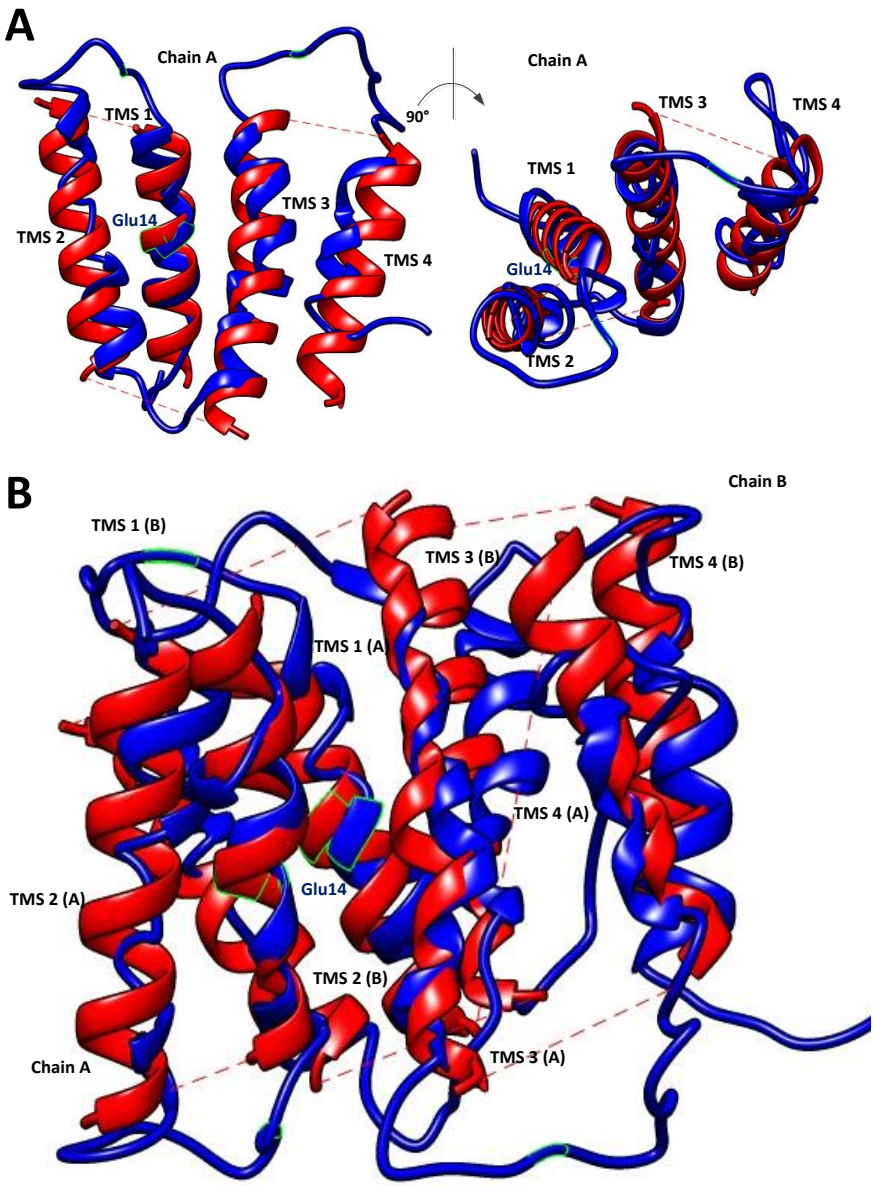

Figure 3.13: MODELLER predicted QacC model aligned with EmrE

Structural alignment of the 3D QacC model as predicted by the MODELLER program with the EmrE template (PDB id code: 2I68). (A) Ribbon 3D structures are presented in single chain-monomeric. (B) Dimeric form. Blue ribbons represent the QacC model whilst red ribbons represent the EmrE model. The conserved Glu14 residue (native Glu13 in QacC and EmrE) (Section 1.6.1), identified in substrate binding is shown in light blue, with the residues outlined in green. Due to the lack of detail in the available solved EmrE structures, discussed in Section 3.7, dashes are used to join TMSs in each monomer.

(Figure 3.14) also show that both structures share the same fold with a TM-score of 0.505 (chain A) and 0.605 (chain B). The calculated RMSD value of 3.19, which is somewhat higher than for the MODELLER predicted QacC model, also indicates that the structures share similarity.

3.7 Discussion

Although crucial for understanding protein function, the determination of protein structures by X-ray crystallography, electron microscopy or NMR remains challenging, as discussed in Section 3.1. As such, alternative methods utilising computer modelling have been applied to generate possible protein models, allowing for structural analysis and the formation of hypotheses which can be experimentally analysed. Due to the limitations associated with the analysis of MPs, as discussed in Section 4.1, and the lack of solved MP structures, computer modelling has been used in the assessment of a number of MPs, including the *E. coli* MdfA multidrug transporter. Using the X-ray structure of LacY and GlpT, an *E.* coli glycerol-3-phosphate transporter, a structural model of MdfA was constructed and analysed (Sigal et al., 2005). This model was then further refined, with the use of mutational analysis and experimental design, and was subsequently used to identify essential residues within MdfA, which were previously unknown. Like MdfA, there is no solved structure for QacC, as such, computer modelling was used to predict both the secondary and tertiary structure of QacC. Models obtained were then evaluated (Section 3.5) and, following further refinement by Cys-scanning mutagenesis (Chapter 5), a final structural model of QacC proposed.

Initial modelling of the QacC multidrug transporter focused on the determination of the secondary structure using a majority-vote principle (Section 3.2). Two different methods were used to predict possible secondary structures of QacC (Section 3.2), results of which were compiled and compared to generate a final 2D structure of QacC (Figure 3.15). The first method employed a single-sequence mode using the QacC amino acid sequence only to identify residues which were exposed or buried within the membrane. The second method was based on a consensus prediction method, which predicts TMS based on an

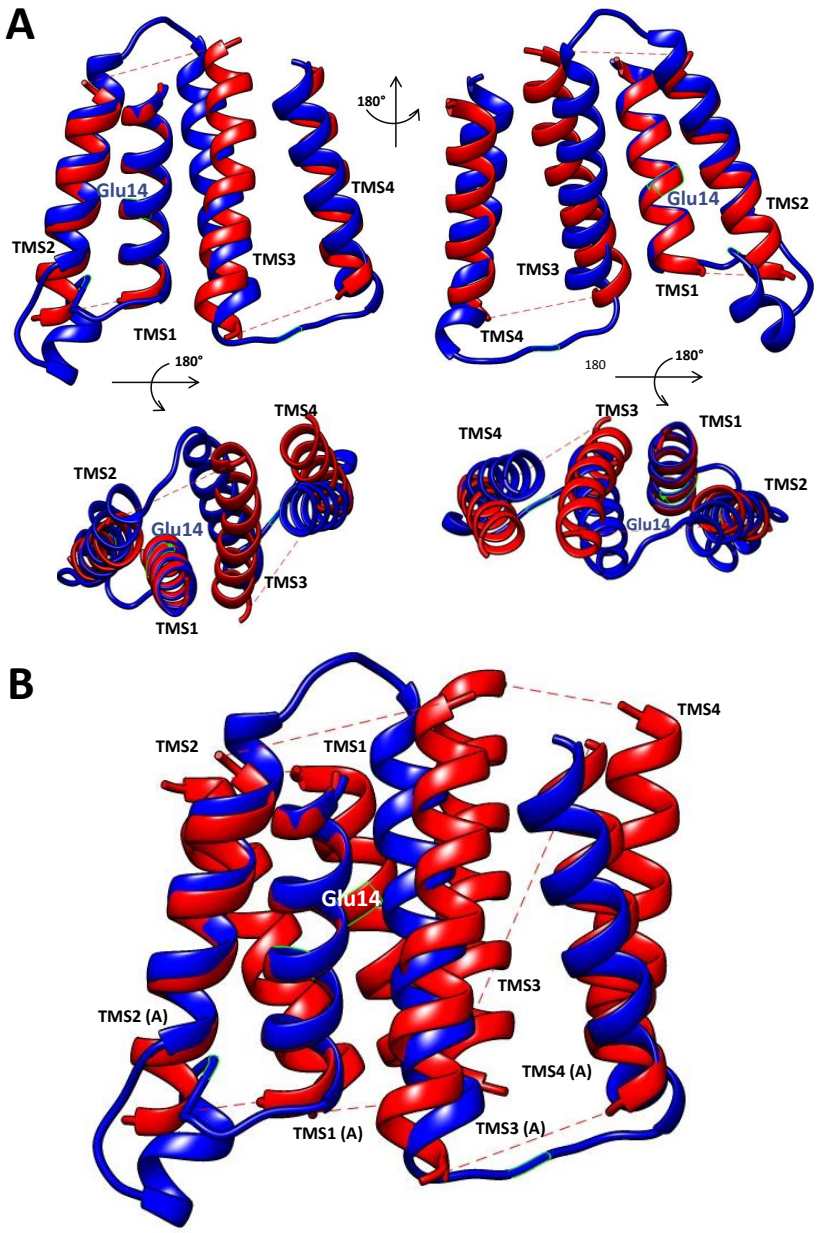

Figure 3.14: Phyre2 predicted QacC model aligned with EmrE

Structural alignment of the 3D QacC model as predicted by the Phyre2 server with the EmrE template (PDB id code: 2I68). (A) Ribbon 3D structures are presented in single chain-monomeric. (B) Dimeric form. Given that the Phyre2 server predicted a monomeric model only, the dimeric form shows only one chain of the QacC model. Blue ribbons represent the QacC model whilst the red ribbon represents the EmrE model. The conserved Glu14 residue, identified in substrate binding is shown in light blue (A) and white (B), with the residues outlined in green.

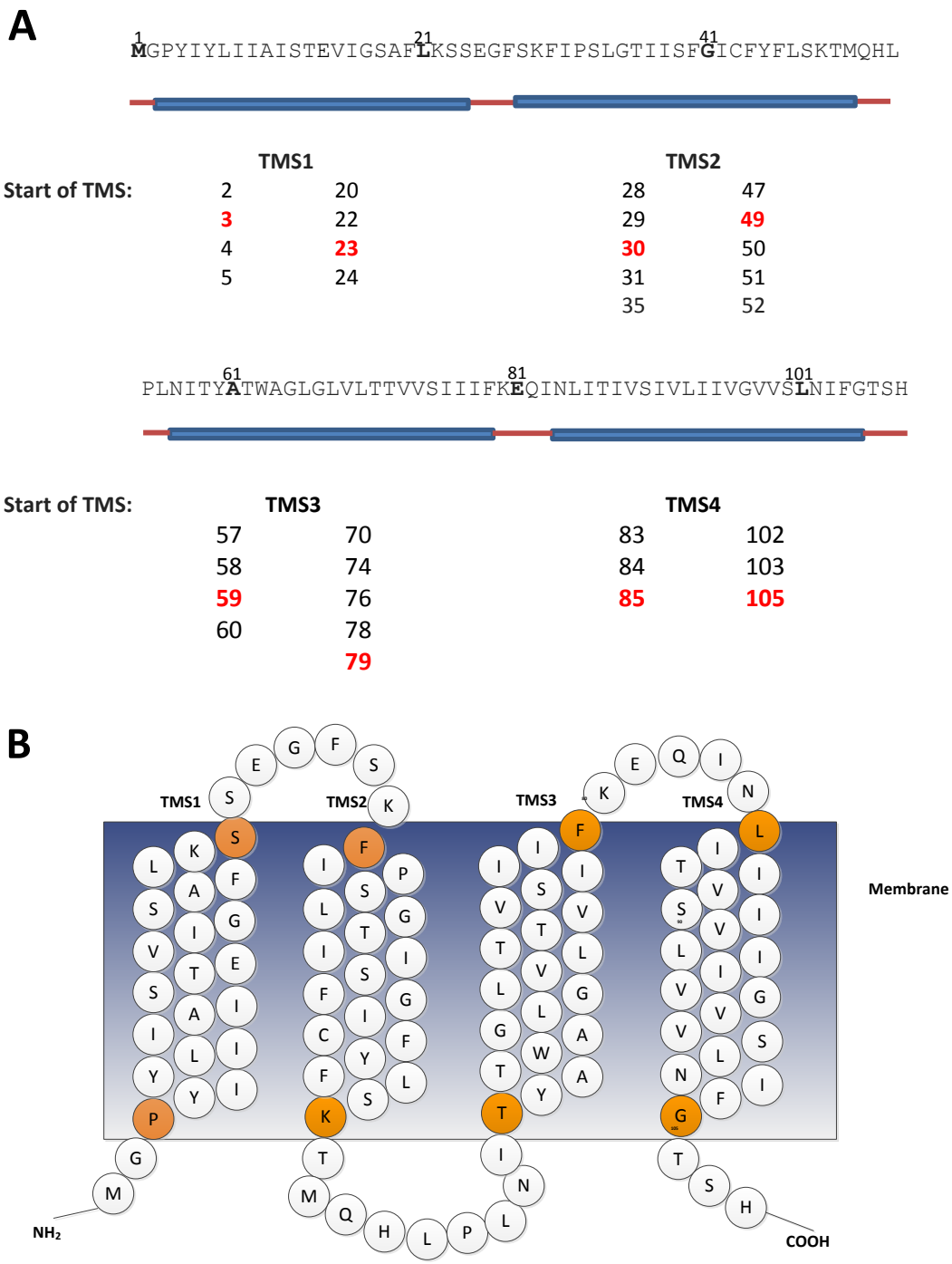

Figure 3.15: Final prediction of the secondary structure of QacC

(A) Residues with each TMS, as predicted in Section 3.4, were compiled and the possible beginning and end of each TMS noted, with the most commonly predicted residue highlighted in red. The QacC sequence used as the input sequence is given above the predicted TMS, with a graphical representation of the possible TMSs (blue bars) and loops (orange line) shown below. (B) The final 3D model as predicted by computer modelling with the residues shaded in orange indicating those which form the boundary of each TMS.

alignment generated between the target sequence, i.e. QacC, and other members of the specific family, in this case the SMR family. Results generated for both methods did not show much variation in the assignment of residues to various TMS, with the largest variation occurring in allocation of residues to TMS2 and TMS3 (Figure 3.15). Final analysis generated a possible secondary structure of QacC with the allocation of Pro3 to Ser23 making up TMS1, Phe30 to Lys49 making up TMS2, Tyr59 to Phe79 TMS3 and finally TMS4 being made up of Leu85 to Gly105 (Figure 3.15). This model and allocation of residues to various TMS is in close agreement with the model proposed by the Paulsen et al., (1995) QacC genetic fusion study, as discussed in Section 1.6.1. To further refine this model, results from Cys-scanning mutagenesis (Chapter 5) in conjunction with those found in this Chapter were used to reassess the possible position of each residue within QacC and, if needed, to redefine each TMS and loops of the QacC multidrug transporter (see Section 5.6).

With no available solved structures of QacC and only four submitted PDB structures of the *E. coli* EmrE (Chen et al., 2007; Fleishman et al., 2006), the only member of the SMR family to have a structure(s) solved, 3D modelling of QacC was initially carried out following the *de novo*/*ab initio* approach (Section 3.3). Using the QUARK algorithm and the CABS-fold server a number of models of QacC were generated (Section 3.4) of which the best two predicted structures were assessed for quality (Section 3.5). The 3D model of QacC as predicted by the QUARK algorithm was shown to satisfy all three validation criteria (Table 3.2) which included the stereochemical quality and assessment of the amino acid side-chain environment. The Ramachandran plot showed that 95.7% residues were in the most favoured region and the Verify3D plot scored a high 82.82%, indicating that the environment profile of the model was good. Finally, the Errat2 quality factors, which measure the distribution of non-bonded interatomic distances within a protein structure (Colovos and Yeates, 1993), were shown to be high at 87.88%. All three analyses reported for the QUARK algorithm predicted QacC structure represent “good consistency” indicating a high quality model.

Although the assessment of the CABS-fold server predicted QacC model (Figure 3.7) was not shown to be as high in quality as the QUARK predicted model, it was nonetheless assessed to be of good quality (Table 3.2, Section 3.2). Passing both the PROCHECK and Errat2 validation criteria, the CABS-fold predicted QacC model had a Verify3D score which fell below the fully acceptable score. With a final score of 72.22%, the predicted model was identified as "worth-pursuing", meaning that refinement of the model may result in a final structure which is more representative of QacC.

Although SMR proteins like QacC have been predicted to function as a dimer (Chen et al., 2007; Poulsen et al., 2009; Tate et al., 2003), the *de novo*/*ab initio* approach to modelling relies only on the amino acid sequence of the protein; as such, the final results of the predicted models are show in monomeric form. Attempts to predict a dimeric model by submitting a repeat of the QacC input sequence failed to generate any results using both the QUARK algorithm and the CABS-fold web server, as the maximum number of residues that can be submitted was 200 and 120, respectively.

Comparative 3D modelling of QacC was undertaken using both the MODELLER program (Sanchez and Sali, 2000) and the Phyre2 server (Kelley and Sternberg, 2009) (Section 3.4). Given that all solved structures identified as homologues of QacC were those of the *E. coli* EmrE, the number of templates available for QacC comparative modelling was limited. As previously discussed in Section 1.6.1, analysis of the sequence identity between QacC and EmrE is 42%, which was above the 40% specified for a good target-template alignment (Sanchez and Sali, 2000). However, the final 3D model of QacC, as generated by the MODELLER program, was largely incomplete. This can be seen by the lack of residues being assigned to specific regions, such as helices, within the protein (Figure 3.8, Section 3.4). However, this may be due to the low resolution of the four structures available for EmrE, with the best being 3.8 Å. At such low resolutions, only the basic contours of the protein chain can be deduced, thus the use of these as templates is limiting and explains the lack of detail in the final QacC model. Nonetheless, although limiting in the assignment of each residue within the protein to a helix or loop, the

arrangement and number of possible TMS was similar to the solved structures of EmrE, as shown in Figure 3.13, with TMS1 appearing to be positioned between TMS2 and TMS3, and TMS4 in chain A predicted to lie next to TMS4 of chain B. Evaluation of this model carried out by PROCHECK, Verify3D and Errat2 (Table 3.2, Section 3.5), revealed that it did not meet any of the three validation criteria. The Ramachandran plot identified only 63.8% residues with the ϕ/ψ conformation in the "most favoured" regions, which is well below the 90% criteria that is indicative of a good model (Hooft et al., 1997). Low results from Verify3D and Errat2 further confirmed the low quality of this model, which due to the limitations of templates themselves was not unexpected.

The final 3D model of QacC was predicted by the Phyre2 web server (Kelley and Sternberg, 2009) which generated a number of models ranked by a confidence level and percentage of identical residues. The highest ranking predictions of the QacC model were based on two EmrE templates, with the highest ranking model based on an EmrE template (PDB ID code – 1S7B) which has since been retracted (Chang et al., 2006). The fact that this model continues to be used as a possible template for 3D modelling by Phyre2 is a cautionary sign, as this template has been removed from the actual PDB server, and as such should be excluded from prediction programs. Once the highest ranking model was excluded, the second model (Figure 3.9, Section 3.4) also predicted using EmrE as a template (PDB ID code – 2I68), was assessed for quality. Results revealed that this model also did not meet any of the validation criteria as assessed by PROCHECK, Verify3D and Errat2 (Table 3.2, Section 3.5). Although showing it to be of higher quality than the MODELLER predicted model, it is difficult to compare these two models, as one predicts the monomeric form and the other a possible dimeric form of QacC. An attempt to model a dimeric form of QacC using the Phyre2 web-server failed to generate a possible structure (data not included). Nonetheless, the placement of the TMS in the monomeric model is similar to the predicted MODELLER structure, which was expected as both models were based on the EmrE structure.

A final assessment of the QacC 3D models, as predicted by the MODELLER program and the Phyre2 server, was carried out using the TM-align algorithm

(Zhang and Skolnick, 2005) (Section 3.6). This was employed to identify the best structural alignment (superposition) between protein pairs and to determine if they shared the same fold. In addition to the TM-score, which indicates structural similarity, the RMSD was also calculated for each alignment by the TM-align algorithm. Of the two models assessed, the highest TM-scores and the lowest RMSD value were allocated to the MODELLER predicted 3D QacC model, indicting a comparable fold and structural similarity. Figure 3.13 shows both structures superimposed, where overlaps signify homology between the models. The highly conserved Glu14 residue, found in TMS1 (originally Glu13 in native QacC, Section 1.6.2), is highlighted. As can be seen in Figure 3.13, this residue is positioned to face the inside of the structure, which in the dimeric form appears to face itself, possibly forming a substrate-binding pocket. Although showing little overlap, the position of each TMS also appears to be aligned between each model, as shown in both the MODELLER predicted QacC model and the one generated by Phyre2 (Figure 3.13 and 3.14). Scoring less in the TM-score and obtaining a higher RMSD value, the Phyre2 generated QacC model is still considered to have a similar fold to the EmrE template and share structural similarity, although to a lesser degree than the MODELLER predicted model. Interestingly, the Glu14 residue, in the superposition, appears to be perfectly aligned in both models. Given that the Phyre2 server predicted the monomeric form of QacC only, using a dimeric structure of EmrE as a template, the superposition is highlighted for one chain only. However both the TM-score and RMSD are calculated against the dimeric form, thus 2 TM-scores were given. The lower TM-score obtained for this model and the higher RMSD value may result from the lack of alignment seen with TMS3 and TMS4 between the QacC and EmrE models. Nevertheless, given the TM-scores, both models appear to share the same fold as the template, as all are over 0.5. The RMSD value given for both models is harder to evaluate as there is no defined cut-off value between what constitutes high similarity and what is not at all similar. The smaller the value obtained, the more similarity, however at 2.97 and 3.19, although implying a level of similarity, further adjustment and revaluation of the models is needed to bring these down to a more significant amount.

A final point of consideration was also given to the two purification tags added to QacC, the 1D4 and Hisx6 tag, and their possible effect on the models predicted in this study. To examine this, models of the secondary structure of QacC were also generated using the QacC sequence which included both tags (Section 3.2). Results generated by the TMHMM, TOPPRED and Jpred3 showed that the additional tags did not affect the final secondary structure of QacC, with each prediction showing QacC to be composed of 4 putative TMS, with the allocation of residues to each segment shown to be comparable to the models predicted without the tags. The 3D models were not evaluated, as these were based on regions of homology and are not resolved for the beginning and end of the protein by most prediction algorithms.

Limitations pertaining to the number and variation of templates available for comparative modelling make it difficult to obtain a model that can adequately represent the QacC structure. The models obtained, however, both from the *de novo*/*ab initio* and comparative modelling; mark the first step in gaining a better understanding of QacC and its possible 3D structure. The models obtained in this chapter will continue to be refined, both by further modelling and by experimental design.

3.8 Conclusion

The models of QacC presented in this chapter provide a possible framework upon which functional and structural considerations can be proposed and tested further. These models will continue to be refined by the experimental approaches undertaken in this project.

CHAPTER 4 – OPTIMISATION OF INDUCTION AND PURIFICATION OF QACC

4.1 Introduction

Membrane transport proteins play a vital role in living cells and are essential for cell signalling, substrate transport and metabolism. Although representing about 20-30% of all protein sequences, they continue to be underrepresented in the PDB (Kozma et al., 2013; Tusnady et al., 2004, 2005). Of the current solved protein structures, MPs make up approximately 2% of the submissions (http://www.rcsb.org/pdb/home). In fact, according to the Stephen White MPSTRUC database, which lists all known 3D structures of MPs (http://blanco.biomol.uci.edu/mpstruc/), the number of unique MP structures available as of June 2015, numbered 541, compared to April 2013, when 393 unique structures were available (Duarte et al., 2013), a stark contrast to soluble proteins which are solved by the thousands each year.

The importance of membrane transporters in living cells and their potential to be targeted by inhibitors, has led to increased efforts to evaluate their structural and functional properties. Nonetheless, the inherent nature of MPs, namely their structural and physiochemical properties can restrict and complicate experimental design. Analysis of MPs is often problematic due to their inherent instability and low natural abundance. Thus, one of the first hurdles to overcome when working with MPs is to obtain sufficient protein yields for purification and further analysis. To obtain high yields of MPs, host organisms such as *E. coli* are frequently utilised for heterologous protein production. Factors contributing to its success as a model host bacterium include its well characterised genetics, ability to grow rapidly and at high density and the availability of a large number of cloning vectors as well as a variety of mutant host strains (Baneyx, 1999; Sorensen and Mortensen, 2005). Nonetheless, even when utilising model organisms such as *E. coli*, a number of challenges are encountered when expressing MPs. Problems arising with MP expression include toxicity following induction and overexpression, inclusion body

formation, inactivity of MPs through mis-folding upon detergent solubilisation, aggregation and protein instability (Bernaudat et al., 2011; Martinez Molina et al., 2008; Niegowski et al., 2006; Sorensen and Mortensen, 2005).

Induction and overexpression of MPs frequently results in cell toxicity, which has been attributed to the accumulation of integral proteins in the cellular membrane leading to its destabilisation. Overexpression of MPs can also overburden protein sorting which may consequently prevent the biogenesis of endogenous proteins, reducing cell viability (Wagner et al., 2007). In addition to these, a vital process for proper MP function is their successful insertion into the lipid bilayer. This is believed to occur by the interaction of translating ribosomes, the bacterial signal recognition particle system, Sec translocase (SecYEG) and the YidC insertase (Dalbey et al., 2011; Gubellini et al., 2011; Raine et al., 2003). Certain MPs, however, do not require the above for proper insertion and folding, as they can assemble into pure lipid membranes without the signal recognition particle pathway and SecYEG (Gubellini et al., 2011). Nonetheless, failure to properly insert into the membrane can result from inefficient export or targeting of preproteins to the membrane associated translocation apparatus, can lead to inclusion body formation and proteolytic degradation of the overexpressed protein (Baneyx and Mujacic, 2004).

Another fundamental component for correct MP insertion is the composition of the lipid bilayer. Recent studies have shown that depletion of the phospholipid phosphatidylethanolamine (PE), or another net neutral lipid from the *E. coli* membrane, can lead to improper protein folding, as was demonstrated with the *E. coli* lactose permease LacY (Bogdanov et al., 2002; Vitrac et al., 2013; Zhang et al., 2003). Given that lipid composition appears to play a crucial role in insertion and stability of MPs, their efficient solubilisation, a vital step in structural analysis, presents a different subset of challenges. The role of detergents in MP solubilisation and purification is to mimic the lipid bilayer; detergents essentially envelop the extracted MP forming a micelle, providing scaffolding for the protein. Amphipathic and consisting of a polar head group and a hydrophobic chain, detergents used for solubilisation fall into four categories; ionic detergents, bile

acid salts, non-ionic detergents and zwitterionic detergents. Of these, non-ionic detergents, considered to be mild and relatively non-denaturing, are frequently used as they do not affect the proteins structural features, allowing it to be isolated in its native and functional form (Scott et al., 2013; Seddon et al., 2004). Detergent choice is also influenced by the type of work being carried out, as factors such as temperature and pH can impact solubilisation and protein stability (Seddon et al., 2004).

Although the above obstacles can occur when working with MPs, there are a number of successful strategies which have been developed to combat these. Recombinant protein expression in *E. coli* can be optimised by choosing a suitable expression system with a focus on the copy number of a plasmid, promoter, a strong ribosome binding site and an efficient transcription terminator (Hannig and Makrides, 1998; Makrides, 1996; Sorensen and Mortensen, 2005). A number of expression systems have been developed for heterologous protein expression, of these the pBAD expression system, based on the *araBAD* promoter (P_{BAD}), is designed for tight control of protein expression. P_{BAD} activity can be modulated by the addition of arabinose, which when coupled with a strain deficient in arabinose transport and degradation, such as the *E. coli* TOP10 strain, can lead to increased expression for proteins under the control of this promoter (Guzman et al., 1995; Sorensen and Mortensen, 2005). In contrast to pBAD, the pBluescript II SK phagemid expression vectors are designed to simplify cloning, sequencing and protein expression via the T7 promoter. They contain an extensive polylinker, T7 and T3 RNA polymerase promoter sequences and the ability to check for cloning efficiency by blue/white colony colour selection. Both of these systems have been successfully utilised to overexpress a number of MPs such as the QacA multidrug transporter (Hassan et al., 2009), and thus were chosen for the expression of QacC in this study.

This chapter describes the expression and purification of the QacC multidrug transporter. As *qacC* is from *S. aureus* and *E. coli* is to be used as a host system for heterologous protein expression, the *qacC* gene was codon optimised for *E. coli* (Section 4.2.1). In this study various QacC constructs were developed, which

included the addition of a number of affinity tags designed to facilitate protein purification and aid future crystallisation studies. For Cys-scanning mutagenesis a Cys-less QacC derivative was constructed and utilised as a template for subsequent mutagenesis (Section 5.2). Two different QacC expression conditions were optimised in this study. Expression of QacC from *E. coli* was optimised for functional assessment and resistance profiling, which utilised non-purified QacC protein (Section 4.3.1 and 4.5), such as transport assays (Section 6.2) and MIC analysis (Section 6.3), as well as for fluorescence-polarisation assays (Section 6.4.2), which required high yields of purified QacC protein (Section 4.3.2 and 4.6).

4.2 Preliminary QacC studies

4.2.1 The *qacC* construct

In an effort to maximise heterologous QacC expression, purification and characterisation of QacC was carried out in *E. coli* TOP10 and BL21 (DE3) strains. Given that *S. aureus* is an AT-rich Gram-positive bacterium, employing a different codon usage than *E. coli* which may limit the expression of *S. aureus* genes due to the lack of appropriate tRNA molecules, a synthetic *qacC* gene was constructed. Codons optimal for expression in *E. coli* were employed to maximise efficient translation of *qacC* in *E. coli* cells (Figure 4.1) and construction of the synthetic *qacC* gene was accomplished using the Stemmer method (Dunham et al., 2001; Stemmer et al., 1995). Briefly, this method utilises a series of PCR reactions and oligomers designed in 40-mer lengths, which are then used to construct and amplify the desired gene. The gel purified PCR product was digested with *Nco*I and *Xba*I and ligated into a modified pBAD myc-HIS B (LifeTechnologies) expression vector, in which the *myc* epitope was replaced with the 1D4 tag and a polyhistidine tag (Hisx6) (Figure 4.2). The Hisx6 tag was included to enable purification of QacC proteins using nickel affinity chromatography (Section 2.7.4), whilst the addition of the 1D4 tag was designed to facilitate future co-crystallisation experiments (Dunham et al., 2001). Incorporation of short epitope tags, such as the anti-1D4 Fab (fragment antibody binding) fragment, into proteins facilitates not only

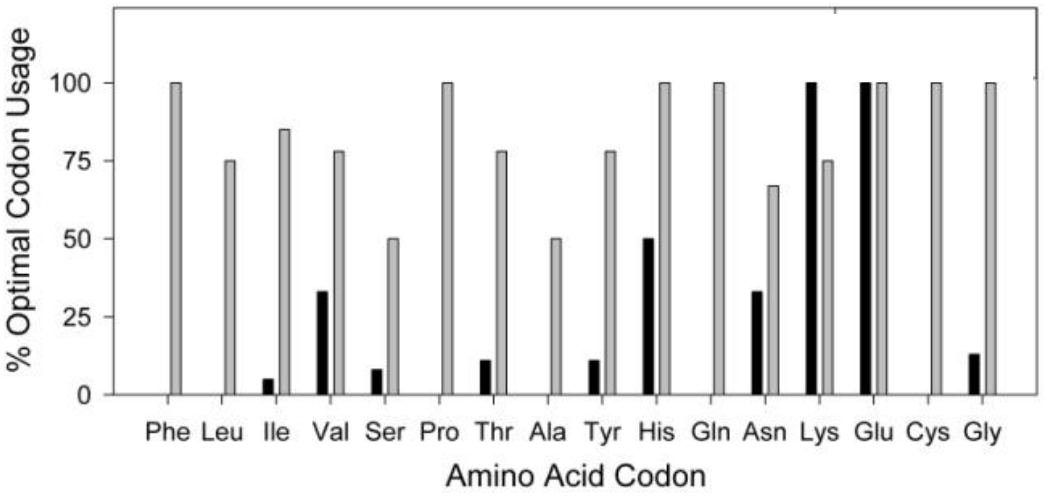

Figure 4.1: Codon optimisation of the *qacC* gene

A comparison of the improved codon usage versus the original codons in *qacC* for expression in *E. coli*, with black bars showing the native *qacC* codon usage and grey bars the synthetic *qacC* codon usage. Amino acids not codon optimised are not included in the above graph.

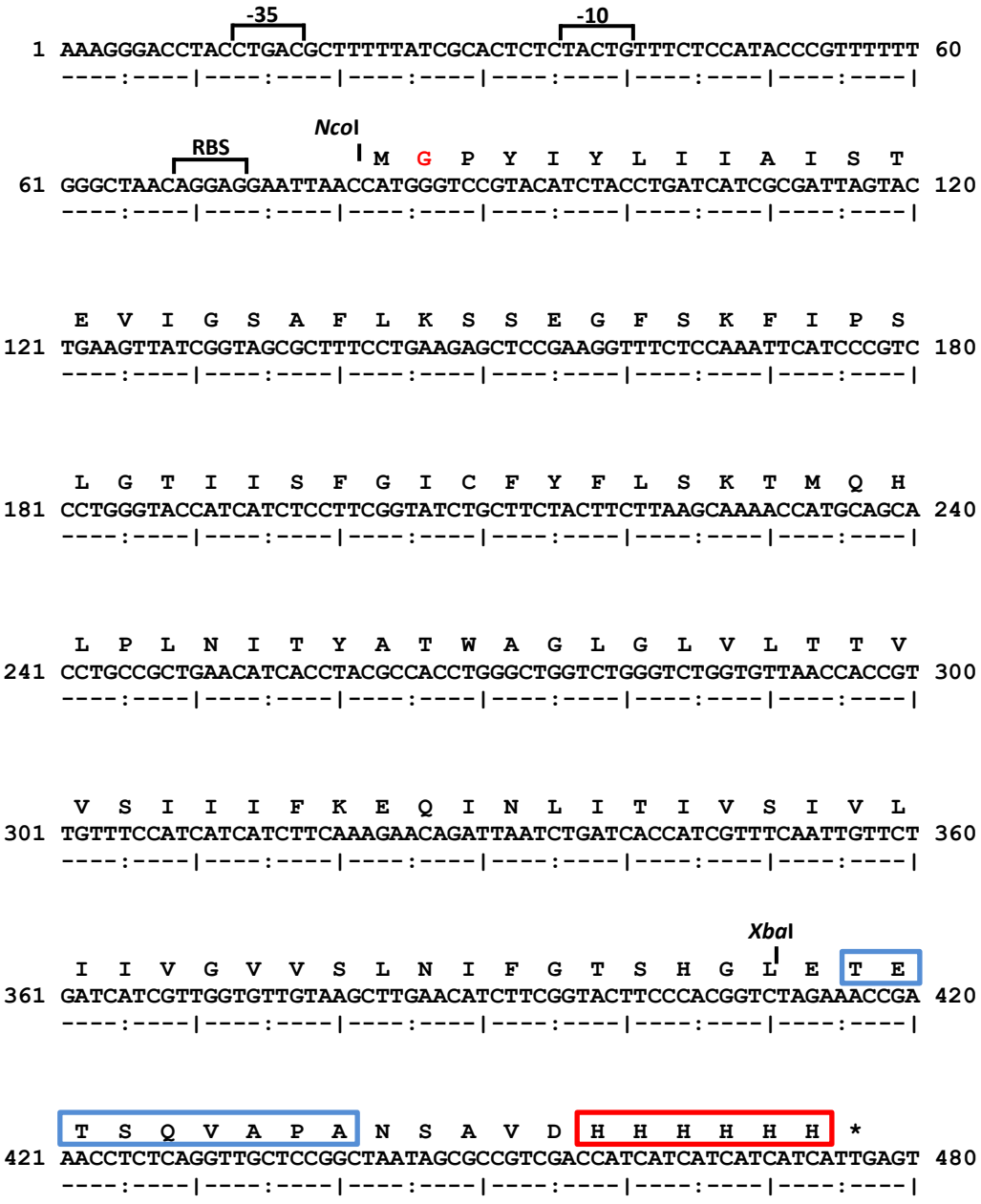

Figure 4.2: Sequence of *qacC* cloned into the pBAD based expression vector

The nucleotide and amino acid sequence of codon optimised *qacC* cloned into the pBAD vector. Locations of restriction sites utilised for the insertion of *qacC* into pBAD are indicated above the nucleotide sequence, whilst the 1D4 (boxed in blue) and Hisx6 (boxed in red) tags used for purification are boxed. The pBAD ribosome binding site (RBS) and the promoter elements (-10 and -35) are underlined. The start of the QacC amino acid sequence is indicated by the first methionine (M), whilst the terminator sequence is denoted with a period. An additional Gly residue, added to facilitate cloning (Section 4.2), is indicated in red.

purification but also the detection, characterisation and localisation of heterologously expressed proteins used in structure-function studies (Molday and Molday, 2014).

To aid mutagenesis studies, restriction sites were introduced every 40 to 50 bp in the synthetic *qacC* gene. An additional residue (Gly) was also added at the N-terminus in the codon optimised QacC construct when a restriction site to facilitate easier cloning was incorporated. This Gly residue was inserted between residues one (Met) and two (Pro) (Figure 4.2), increasing all residue numbers by one e.g., the highly conserved Glu residue within TMS1 went from being Glu13 to Glu14 (Dunham et al., 2001).

For comparison of expression and function, the *qacC* gene was also cloned into the pBluescript II SK cloning vector. This was carried out by amplifying the QacC_Cys43Thr fragment from the pBADQacC_C43T plasmid (Table 2.3) using the forward primer C43T_For and the reverse primer C43T_Rev (Table 2.4) (Section 2.6.1). Once amplified, the *qacC* gene was digested with *XhoI* and *Eco*RI, into the pBluescript II SK plasmid downstream of the T7 promoter (Sections 2.6.4 and 2.6.5). This construct carried both the 1D4 and Hisx6 tags, of which the Hisx6 tag was used for protein detection and purification.

To decrease the possibility that the addition of the two tags, requiring an additional 20 amino acid residues, to a protein of only 108 amino acids did not negatively influence the function of QacC, a new oligonucleotide, C43T_RevHis6 (Table 2.4), was designed to eliminate the 1D4 tag, resulting in a QacC construct carrying only the Hisx6. This primer, along with the sequencing M13 forward primer (Table 2.4) was used to generate a new *qacC* construct, which was cloned into the *Nco*I and *Eag*I sites of the pBluescript II SK cloning vector (Sections 2.6.1, 2.6.4 and 2.6.5).

4.3 Optimisation of protein expression from the pBAD expression system

4.3.1 Conditions for phenotypic assays

Initially the *qacC* gene was cloned behind the P_{BAD} in the pBAD/*Myc*-His B plasmid. Genes encoding both the positive and negative regulator of P_{BAD}, *araC* are also included in this expression system, allowing for tight regulation of expression which can be induced by the addition of L-arabinose or repressed by glucose. Efficient expression is obtained by expressing constructs in *ara* mutant *E. coli* strains, such as the *E. coli* TOP10 strain (Table 2.2), which does not metabolise L-arabinose allowing for its intracellular concentration to remain elevated acting as a long lived inducer (Guzman et al., 1995).

Initial expression of *qacC* from the pBAD-based vector was carried out in *E. coli* TOP10 cells, following induction with various concentrations of L-arabinose ranging from 0.02% to 0.002%. Different induction times and growth conditions were also examined (Table 4.1). In order to determine if the overexpressed QacC was functionally active, fluorimetric transport assays were undertaken (Section 2.11.1), conducted as previously described (Hassan et al., 2008; Xu et al., 2006).

Induction conditions used in preliminary QacC expression studies used 0.02% L-arabinose added at OD_{600}=0.55 followed by further growth at 37°C for two hours. Induction however, resulted in cell toxicity which, as discussed in Section 4.1, appears to be a problem arising with MP overexpression. Growth curves revealed that QacC expression from the pBAD-based vector after induction with L-arabinose, at any concentration or time of induction, proved to have a toxic effect on the *E. coli* Top10 cells, as they ceased growing almost immediately upon induction (Figure 4.3). To determine whether the toxicity exhibited by the addition of the arabinose rendered the cells metabolically inactive, cells were assayed for their ability to efflux ethidium out of the cell (Section 2.11.1), with results revealing that although the cells were no longer growing they were still able to efflux ethidium at a high level. Thus, although not detrimental to function, induction conditions aiming

Table 4.1: Examinations of induction and expression conditions for pBAD-constructs in *E. coli* TOP10 cells

Media[a]	Induction	Conc. (%)	OD_{600}[b]	Temp (°C)	Duration of induction	Optimal conditions used for expression
LB	None	-	-	37	-	No expression detected
LB	L-arabinose	0.02	0.40	37	1 hour	Highest expression level obtained using 0.02% L-arabinose, at OD_{600}=0.80, followed by 1 hour growth
		0.02	0.55	37	1 hour	
		0.002	0.55	37	2 hours	
		0.02	0.55	37	2 hours	
		0.02	0.80	37	30 minutes	
		0.02	0.80	37	1 hour	
		0.02	1.00	37	30 minutes	
		0.02	1.00	37	1 hour	
		0.02	1.15	37	30 minutes	
		0.02	1.00	37	1 hour	
TB	L-arabinose	0.02	0.55	37	2 hours	Highest expression level obtained using 0.02% L-arabinose, at OD_{600}= 0.95 to 1.15, followed by 30 minutes growth
			0.95	37	30 minutes	
			1.15	37	30 minutes	

a: Luria-Bertani (LB), Terrific Broth (TB) – supplemented with 100 µg/ml ampicillin
b: OD_{600} at time of induction

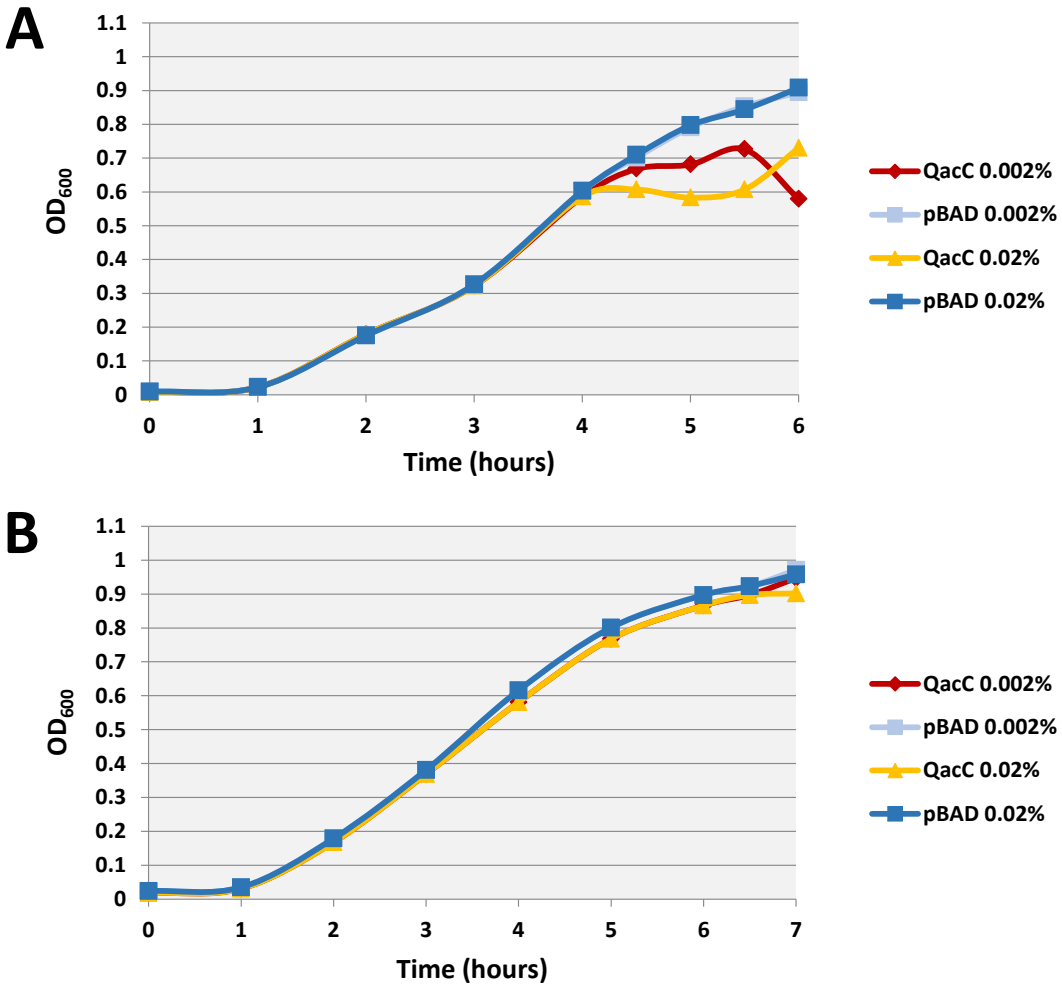

Figure 4.3: Growth curve of *E. coli* TOP10 cells expressing pBAD-constructs following induction with 0.02% and 0.002% L-arabinose

Cultures of *E. coli* TOP10 cells carrying the pBADQacC_C43T clone (Table 2.3) and vector alone were grown overnight in LBAmp at 37°C. Cells were diluted 1:20 in fresh media, grown to OD_{600}=0.55 or 0.95 and induced with 0.002% or 0.02% L-arabinose (Section 2.7.1). (A) Growth of cells following induction at OD_{600}=0.55 with 0.002% and 0.02% L-arabinose. (B) Growth rate of cells following late induction at OD_{600}=0.95 with 0.002% and 0.02% L-arabinose.

towards maximising QacC protein expression were explored. Thus the timing of induction was tested at OD_{600}=0.8, 1.0 and 1.15, with 0.02% L-arabinose and the duration of induction tested at 30 minutes and one hour at 37°C. In addition to 37°C, incubation at 25°C overnight was also tested, but did not result in improved protein expression (data not shown). Results indicated that optimal expression of QacC occurred with induction at OD_{600}=0.8 with 0.02% L-arabinose followed by growth for one hour at 37°C (Figure 4.4). Given these results, all 108 single and nine double pBAD-based QacC constructs (Table 2.3) were expressed following these induction conditions.

4.3.2 Conditions for fluorescence-polarisation binding assay

To obtain high yields of purified QacC protein needed for fluorescence-polarisation binding assays (Section 6.4.2), expression of QacC from the pBADQacC_C43T vector (Table 2.3) in *E. coli* TOP10 cells was carried out in LBAmp, as well as TBAmp. TB, an enriched medium containing glycerol, was developed for optimal growth of *E. coli*, as such it was used in this study in an attempt to increase final yields of QacC expression. A comparison of growth and expression of QacC in these media, following induction at OD_{600}=0.8 with 0.02% L-arabinose and subsequent growth for one hour at 37°C revealed that the amount of QacC expressed was higher from the cultures growth in TBAmp, as determined by examination of the membrane fraction (Section 2.7.2) and quantification by a Lowry assay (Section 2.9.4) (Table 4.2).

4.4 Optimisation of protein expression from the pBluescript II SK system

As with the pBAD-based QacC system, induction and expression of QacC from the pBluescript II SK vector was examined and optimal conditions determined. Preliminary overexpression trials were focussed on *E. coli* DH5α cells, as studies on QacA (Section 1.3.2.3) have revealed that when cloned into the pBluescript II SK-

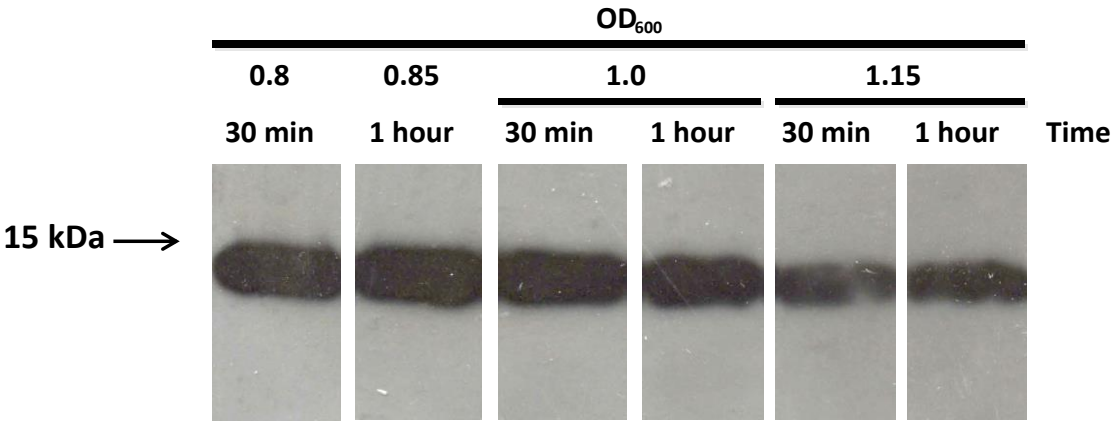

Figure 4.4: Relative expression of QacC in *E. coli* TOP10 cells

Cultures of *E. coli* TOP10 cells carrying pBADQacC_C43T (Table 2.3) were grown to OD_{600}=0.8, 1.0 and 1.15 in LBAmp and induced for 30 minutes or one hour with 0.02% L-arabinose (Section 2.7.1). Cells were collected and membrane fractions isolated (Section 2.7.3). Membrane fractions were solubilised with 1% (w/v) DDM, and 20 µg of each sample was subjected to SDS-PAGE (Section 2.8.1), transferred to a PVDF membrane (Section 2.9.2) and probed with an anti-Hisx6 antibody (Section 2.9.3). A comparison of conditions revealed that induction at OD_{600}=0.8 followed by growth for one hour, resulted in higher expression levels of QacC. Locations of the QacC protein bands (at 15 kDa) are indicated on the left-hand side of the Western blot.

Table 4.2: Conditions examined for induction of *qacC* transcription in *E. coli* TOP10 cells

Media[a]	Starting volume	OD_{600} at time of induction	Duration of induction	Membrane fractions (mg)	Purified proteins (mg)
LB	200 ml	0.55	2 hours	1.2 – 1.5	0.01
		0.8	30 minutes	1.6	0.01
		0.85	1 hour	2.0	0.01
		1	30 minutes	2.4	0.01
		0.95	1 hour	2.6	0.01
		1.15	30 minutes	3.0	0.01
		1.15	1 hour	3.1	0.02
TB	200 ml	1	30 minutes	1.7	0.03
		0.8	1 hour	1.7	0.03

a: Media abbreviations: Luria-Bertani broth (LB), terrific broth (TB)-supplemented with 100 µg/ml of ampicillin (Section 2.3)

based plasmid, QacA was constitutively expressed at low-levels via a "leaky" mechanism in a number of *E. coli* strains, including the DH5α strain (Hassan et al., 2008; Hassan et al., 2009). To investigate this, *E. coli* DH5α cells carrying pBSQacC_C43T (Table 2.3) were grown to OD_{600}=1.0, and whole cell lysate (Section 2.7.2) samples of QacC were subjected to Western blot analysis (Section 2.9.2) using the anti-Hisx6 tag antibodies (Section 2.9.3). Results revealed that unlike QacA, QacC was not expressed to a measurable level from the pBluescript II SK plasmid in *E. coli* DH5α cells (data not shown).

To determine the optimal induction conditions for the expression of QacC from the 15 pBluescript II SK-based QacC mutants (Table 2.3), the pBSQacC_C43T plasmid was transformed into *E. coli* BL21 (DE3) cells (Table 2.2) (Section 2.4.2). Derived from the *E. coli* B strain, this strain was constructed for high-level expression of MPs. It carries an IPTG-inducible *lac* UV5 promoter that controls the T7 polymerase from the lysogenic λ prophage DE3, which allows for expression of proteins to be adapted and optimised for maximum expression but minimise the toxic effect these might have on the cell. It also carries genetic markers that help recombinant proteins to be overexpressed and accumulate to high levels without being degraded (Miroux and Walker, 1996; Studier and Moffatt, 1986). From the conditions tested (Table 4.3), optimal induction conditions for QacC were obtained by the addition of 0.1 mM IPTG at OD_{600}=0.55 followed by growth for one hour at 37°C. Fluorimetric transport assays demonstrated that following this induction, QacC was able to mediate ethidium efflux from the *E. coli* BL21 (DE3) cells. Although inducible, leaky expression in *E. coli* BL21 (DE3) was also investigated and found that unlike in *E. coli* DH5α cells, expression was observed for QacC in the absence of induction (Figure 4.5). Nonetheless, although detectable on a Western blot, uninduced *E. coli* BL21 (DE3) cells carrying the pBluescript II SK-based QacC constructs did not mediate ethidium efflux with the same efficiency as induced cells (data not shown). Therefore, all of the 15 pBluescript II SK-based QacC constructs in *E. coli* BL21 (DE3) cells were induced with the aforementioned conditions.

Table 4.3: Condition examined for induction of *qacC* transcription from the pBluescript II SK vector in *E. coli* with IPTG

***E. coli* strain**	**Media[a]**	**Induction**	**IPTG[b] (mM)**	**OD_{600}**	**Temp (°C)**	**Duration (hours)**	**Optimal conditions used for expression**
DH5α	LB	None	NA	-	37	-	No leaky expression [c]
BL21 (DE3)	LB	None	NA	-	37	-	Leaky expression observed [c]
BL21 (DE3)							Optimal expression following induction at OD_{600}=0.55 with 0.1 mM IPTG, followed by growth for 1 hour at 37°C
	LB	IPTG	0.1	0.55	37	1, 2 and 3	
			1	0.55	37	1, 2 and 3	

a: Luria-Bertani broth (LB) – supplemented with 100 µg/ml of ampicillin (Section 2.3)
b: Isopropyl β-D-1-thiogalactopyranoside (IPTG)
c: Cells grown till OD600=0.6, expression analysed correspondingly with induced samples at 1, 2 and 3 hours after induced cells were sampled (comparison of induced vs uninduced cells)

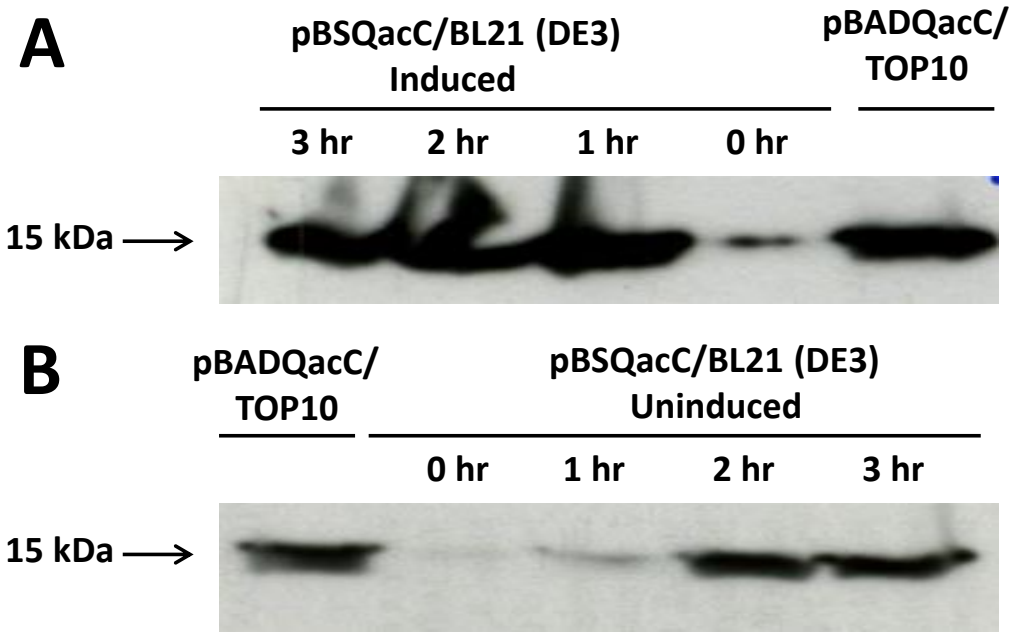

Figure 4.5: Relative expression of QacC in *E. coli* BL21 (DE3) cells

(A) Cultures of *E. coli* BL21 (DE3) cells carrying pBSQqacC_C43T (Table 2.3) were grown to OD_{600}=0.55 in LBAmp and induced for three hours with 0.1 mM IPTG (Section 2.7.1) with samples taken at zero, one, two and three hours following induction. (B) Leaky expression was also tested, with cells grown at 37°C and samples collected at the same time points as for the induced samples. Cells were pelleted, lysed and 20 µg of each sample was subjected to SDS-PAGE (Section 2.8.1), transferred to a PVDF membrane (Section 2.9.2) and probed with an anti-Hisx6 antibody (Section 2.9.3). Locations of the QacC protein bands (at 15 kDa) are indicated on the left-hand side of the Western blot. The QacCpBAD/TOP10 samples were included for comparison to show relative amounts of expression.

4.5 Optimisation of QacC protein purification

For small-scale investigation of QacC purification, membrane fractions of *E. coli* cells carrying either the pBAD or pBluescript II SK-based QacC proteins (Section 2.7.4) were suspended in membrane buffer (Table 2.1) to a total protein concentration of approximately 4-5 mg/ml. Solubilisation of membrane fractions was carried out in 1% DDM, as this concentration has been shown to be optimal for other membrane transporters belonging to the SMR family (Ninio and Schuldiner, 2003; Soskine et al., 2004; Weinglass et al., 2005). After the insoluble material was removed by centrifugation, the soluble material was incubated in binding buffer and 0.1 volumes of Pro-Bond™ Ni-resin for one hour at 4°C. Following binding, the Ni-resin was pelleted and washed with bead wash buffer supplemented with 20 mM imidazole (pH 8.0) to remove non-specifically bound proteins. The QacC protein was eluted from the beads at 37°C for 30 minutes in bead wash buffer supplemented with 200 mM imidazole (pH 8.0) (Section 2.7.4). Samples were resolved on a 15% SDS-PAGE gel (Section 2.8.1) and stained with Coomassie Brilliant Blue R-250 (Section 2.9.1). The results of the stained gel revealed that the above protocol produced low quantities of the QacC protein, as the band corresponding to QacC was very faint (data not shown). In addition to the band corresponding to the purified QacC other non-specific protein bands were also seen. To increase the quantity of purified QacC protein eluted off the Ni-resin and to reduce non-specific binding, the binding buffer was supplemented with 20 mM imidazole (pH 8.0) whilst the bead wash buffer was supplemented with 50 mM imidazole (pH 8.0). Addition of imidazole to both buffers decreased the binding of non-specific proteins; however, the concentration of the purified QacC protein still remained low.

Three factors were subsequently investigated in respect to purification of the QacC protein: 1) duration of binding to the Ni-resin; 2) duration of incubation in the elution buffer and; 3) concentration of imidazole in the elution buffer. To test these, binding to the Ni-resin was undertaken for one hour or overnight at 4°C, and elution time tested overnight at 4°C or 30 minutes at 37°C, with the bead wash buffer supplemented with 200 or 400 mM imidazole (pH 8.0). Samples were

compared using a Lowry assay (Section 2.9.4) and SDS-PAGE gels stained with Coomassie Brilliant Blue R-250 (Section 2.8.1 and 2.9.1) (Figure 4.6). These indicated that overnight binding of QacC to Ni-resin at 4°C followed by a 30 minute elution step at 37°C with 400 mM imidazole (pH 8.0) gave the highest yield of purified QacC protein. To verify that all purification steps examined resulted in purified QacC, Western blotting (Section 2.9.3) was also carried out (Figure 4.6B). Detection by Western blotting was required in order to evaluate expression levels of QacC as Coomassie staining lacked sensitivity (Figure 4.6A). This confirmed that in all conditions QacC was purified with 400 mM imidazole (pH 8.0). The purification steps mentioned above were used to purify all QacC constructs, from both the pBAD- and pBluescript SK II-based vectors.

4.6 Purification of QacC for fluorescence-polarisation binding assays

For assays requiring higher concentrations of the QacC protein, such as those used to assess protein:drug binding (Section 6.4.2), QacC purification performed on larger volumes of *E. coli* cells habouring pBADQacC_C43T (Table 2.3). Initial purification trials were based on cultures, which due to the toxicity to *E. coli* cells following the induction of QacC, were left to grow until late exponential phase prior to addition of L-arabinose (0.02%) (Section 2.7.1). The membrane fractions of the cells expressing QacC protein were isolated (Section 2.7.3) and solubilised in 20 mM Tris-HCl (pH 8.0) containing 200 mM NaCl and 1 % DDM. Solubilised MPs were bound to Ni-resin overnight at 4°C and washed four times with bead wash buffer supplemented with 50 mM imidazole (pH 8.0). QacC protein was then eluted with one bed volume of bead wash buffer supplemented with 400 mM imidazole (pH 8.0) (Figure 4.7). Expression and purification of the QacC protein using the method described typically resulted in the purification of 0.2 mg/ml of QacC protein. Purified QacC protein was dialysed twice against 2 litres of dialysis buffer (Table 2.1) (Section 2.7.4) to remove the imidazole, which could interfere with downstream applications such as the assessment of substrate binding. To increase the final concentration of purified QacC, samples were concentrated using centrifugal filter units (Section 2.7.4).

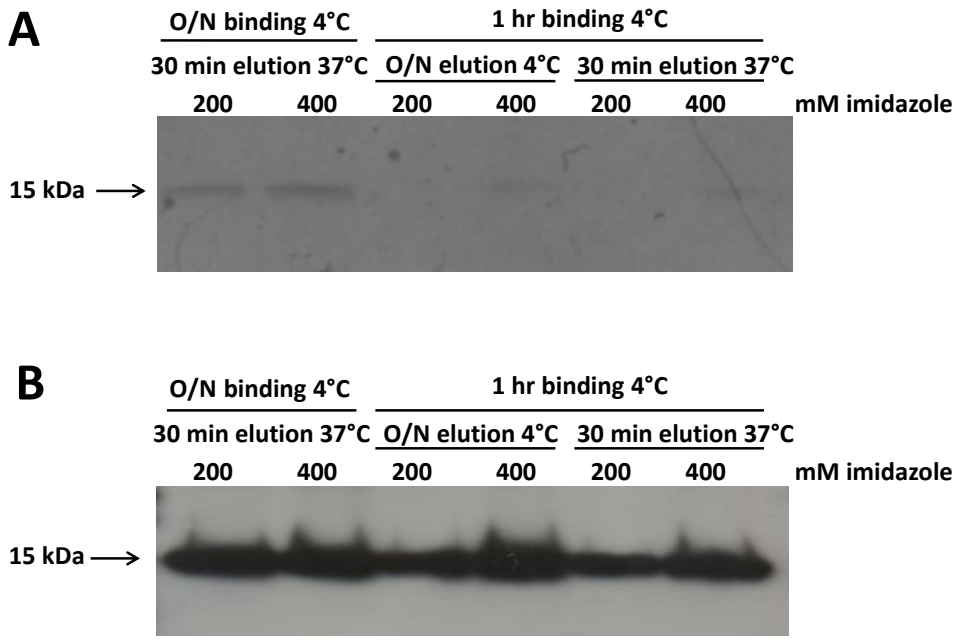

Figure 4.6: Purification of QacC from pBAD in *E. coli* TOP10 cells

Cultures of *E. coli* TOP10 cells carrying pBADQacC_C43T (Table 2.3), were grown to OD_{600}=0.8 in LBAmp and induced for one hour with 0.02% L-arabinose (Section 2.7.1). Cells were collected and membrane fractions isolated (Section 2.7.3). The membrane fractions were solubilised with 1% (w/v) DDM and optimal binding and elution of QacC/1D4/Hisx6 tested as described in Section 4.5. Solubilised QacC/1D4/Hisx6 protein was bound to Ni-resin overnight or for one hour at 4°C and eluted for 30 minutes at 37°C or overnight at 4°C with bead wash buffer containing either 200 mM or 400 mM imidazole (pH 8.0). Equal volumes of samples were subjected to SDS-PAGE (Section 2.8.1). (A) Coomassie Brilliant Blue R-250 stained gel (Section 2.9.1) (B) gel was transferred to a PVDF membrane (Section 2.9.2) and probed with an anti-Hisx6 antibody (Section 2.9.3). Locations of the QacC protein bands (at 15 kDa) are indicated on the left-hand side of the Western blot.

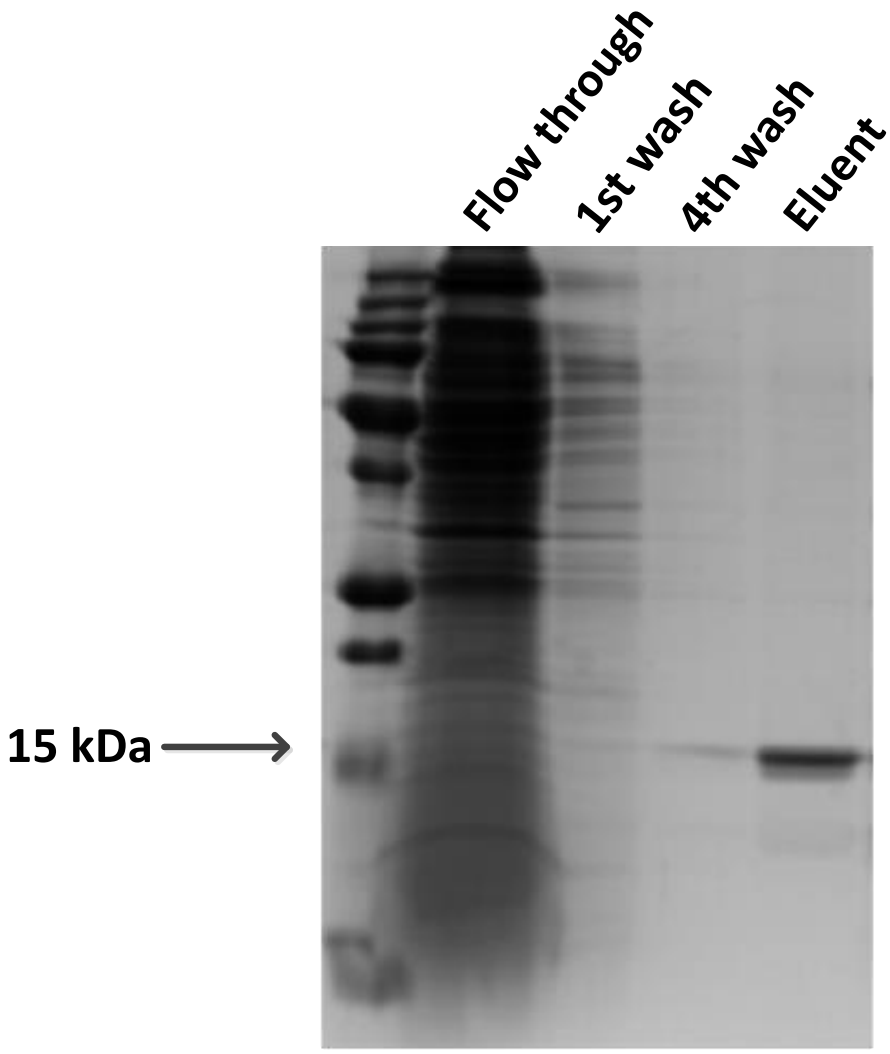

Figure 4.7: Large scale Ni-affinity column purification of QacC

The QacC-Hisx6 protein was expressed and purified from *E. coli* Top10 cells carrying pBADQacC_C43T (Table 2.3), as outlined in Section 2.7.1, 2.7.3 and 2.7.4. Samples of the column flow through, wash (1 and 4) and the eluent were subjected to SDS-PAGE (Section 2.8.1) which was stained with Coomassie Brilliant Blue R-250 (Sections 2.9.1). Locations of the QacC protein bands (at 15 kDa) are indicated on the left-hand side of the gel.

4.7 Discussion

Construction of the synthetic *qacC* was initially carried out by our collaborators with the intent to conduct crystallographic and dynamic structural studies, in an effort to understand the basic mechanism of substrate binding and to evaluate the conformational changes to QacC required for transport (Dunham et al., 2001). To this aim a number of different QacC constructs were created, varying in the types of epitope tags included for purification and detection. The Hisx6 tag was added to enable purification by nickel affinity columns, as was carried out in this study, whilst the 1D4, in addition to allowing for purification was added to facilitate antibody-mediated co-crystallisation experiments.

The aims of the studies presented in this thesis were to assess QacC function and topology. More specifically, Cys-scanning mutagenesis, discussed further in Section 5.1, was used to evaluate the structure and function of the QacC membrane transporter. To this end, each residue was substituted with Cys and the protein structure and function were assessed. Construction of each QacC-Cys mutant was based on the pBAD QacC/1D4/Hisx6 construct and its expression (Section 5.2.3), position relative to the membrane (Section 5.3), ability to export ethidium out of the cell (Section 6.2) and confer resistance (Section 6.3) was analysed. Expression of QacC was optimised for functional assessment such as efflux assays and analysis of resistance, in addition to the evaluation of topology by fluorescein-5-maleimide labelling. It was also optimised for functional studies such as binding assays (Section 6.4.2) which required higher concentrations of purified QacC.

Although difficulties in working with the pBAD expression vector were evident from the beginning; this included toxicity to the *E. coli* TOP10 cells almost directly upon induction with L-arabinose, construction of QacC mutants continued in this vector. It was not until an additional 40 QacC mutants were constructed in this study using the pBAD vector, that it was discovered that of the 22 mutants constructed by our collaborators in the USA, only six were viable. Given the number of QacC mutants generated in this vector at that time and the resources used to

carry this out, it was determined that all remaining QacC mutants were to be constructed using the pBAD vector.

Due to the difficulties experienced in working with the pBAD expression system, 15 QacC mutants were also generated using the pBluescript II SK vector and results compared with those of the pBAD QacC constructs. In contrast to the pBAD-based system, induction of QacC in the pBluescript II SK cloning vector with 0.1 mM IPTG did not result in *E. coli* cell toxicity (Section 4.5), allowing for more straightforward induction and expression of QacC than with the pBAD-based system. A host analysis study has shown that *E. coli* BL21 can achieve higher growth rates when compared to the *E. coli* K-12, possibly owing to a differentially regulated glucose transport. In addition to growth rates, the BL21 strain has been identified as the ideal candidate of choice for processes which require higher rates of biomass and a lower accumulation of by-products (Marisch et al., 2013). However, although possibly easier to work with, a change to the pBluescript II SK vector meant the recloning of over 40 mutants into a new system and their complete reassessment, which would have resulted in lengthy delays, thus assessment of QacC and its variants in the pBAD based system was continued.

Given that the original pBAD based QacC/1D4/Hisx6 construct carried two purification tags and that the 1D4 tag was most likely not to be used, a QacC construct was generated to carry the Hisx6 tag only (Section 4.2.1). This was carried out to exclude the possibility that the addition of two tags may inadvertently have an effect on the function of a small protein like QacC. Functional analysis of the pBluescript II SK QacC/1D4/Hisx6 and QacC/Hisx6 constructs confirmed that both were phenotypically identical (Section 6.3.2). Based on these results, all further assessment was carried out on QacC mutant constructs carrying both epitope tags.

Focussing on the pBAD-based system of QacC expression a range of expression conditions were evaluated to maximise QacC expression. These included concentration and timing of L-arabinose induction and duration of induction. Of these both the timing and duration of induction with L-arabinose were shown to have the greatest impact on the expression levels of QacC and its variants. Maximal expression levels were obtained following induction with 0.02% L-arabinose when

E. coli TOP10 cell density reached OD_{600}=0.8 with further growth of one hour at 37°C (Section 4.3). Induction time and duration of induction were also seen as crucial factors for optimal pBluescript II SK-based QacC expression, although temperature also contributed to increased levels of QacC expression (Section 4.3). For high yields of QacC, needed for structural evaluation and binding assays of QacC, a change in media from LB to TB, allowed for an increase in the expression levels of QacC and its variants in the pBAD-based system. This is due to the nature of TB, known as an enriched medium, which is specifically designed to help recombinant *E. coli* strains to maintain an extended growth phase (Tartof and Hobbs, 1987). Thus, growth in TB allowed for a longer exponential growth phase which resulted in higher expression levels of QacC.

Yields of the QacC protein following the Ni-affinity column purification method using *E. coli* TOP10 cells carrying the QacC/1D4/Hisx6 constructs, averaged at 0.1 mg/litre of culture, with purity averaging at approximately 95% as assessed by Coomassie staining (Section 2.9.1). In order to concentrate QacC for binding studies (Section 6.4.2) samples were subjected to centrifugal filtration (Section 4.5) using units capable of effectively concentrating DDM micelles of 60-70 kDa (Prive, 2007; Strop and Brunger, 2005). Previous studies on the use of detergents on MPs have identified DDM to be the optimal detergent for solubilisation and retention of function (Elbaz et al 2004, Hassan et al 2009, Putman et al 1999). The addition of 1% DDM is in accordance with assessment of other SMRs like EmrE (Soskine et al., 2004; Steiner-Mordoch et al., 2008), TBsmr (Charalambous et al., 2008) and Hsmr (Ninio and Schuldiner, 2003), and was found to meet solubilisation requirements for QacC in this study.

Although Ni-affinity column purification resulted in samples of QacC that were approximately 95% pure, the problem of non-specific binding continued to plague the purification process. Such non-specific binding can result from the presence of surface-exposed histidine residues of unspecific proteins, eluting with the target. Although the affinity of the target protein is much higher, more stringent conditions needed to be employed to remove such contaminants. In this study, imidazole (pH 8.0) was added as a competitive agent to increase the purity of QacC.

By the addition of 20 mM imidazole (pH 8.0) to the binding buffer and 50 mM imidazole (pH 8.0) to the wash buffer, much of the non-specific binding was eliminated (Section 3.6). Given that high concentrations of imidazole were used throughout the purification process, with the final concentration of the elution buffer containing 400 mM imidazole (pH 8.0), dialysis was carried out to remove imidazole, to ensure that it did not interfere with downstream assessment of QacC.

Though both expression and purification of QacC were optimised, yields of QacC remained low. Even when cloned and expressed from the pBluescript II SK vector using *E. coli* BL21 (DE3) cells, where immediate cessation of bacterial growth was not observed following induction, the final amount obtained of QacC was still very low. This may be as a result of the toxicity which is associated with the expression of many MPs, which when expressed even at low levels can be toxic to the cell when associated with the membrane. Toxicity may result from the over accumulation of integral proteins to the cell membrane, destabilising it and ultimately leading to cell death. The transcriptional and translocation machinery may also become over loaded, and export functions which are vital to the cell may be interrupted, once again leading to cell death (Grisshammer and Tate, 1995). These factors, amongst others, may ultimately be contributing to the low expression of QacC as seen in this study.

Lastly, the purification process itself may also have contributed to the final low yields of QacC, with problems associated in both the binding and elution of QacC from the Pro-Bond^{TM} Ni-resin. An inaccessibility of the QacC Hisx6 tag may have prevented some of the solubilised QacC from being able to bind to the Ni-resin, whilst, although a higher concentration of imidazole was needed to improve elution (Section 4.6), QacC may still not have been completely eluted off the beads.

The above results illustrate the challenges associated with MP expression. These and other limitations pertaining to MP production are continually being analysed, with methodologies being developed aimed at overcoming the bottlenecks associated with their expression. A number of strategies employed in this study, such as the use of both the pBAD and pBluescript II SK expression systems and various *E. coli* strains, were initially analysed and optimised for the

overexpression of the QacA multidrug efflux pump (Hassan et al., 2009). However, although shown to be successful for the production of QacA, utilising similar methodologies failed to produce high yields of the QacC MP. This illustrates that although successful for the expression of some, expression systems need to be evaluated and tailored to the protein being assessed. However, studies have now shown that the conventional approach of repeated trial-and-error, which see optimisation of external variants such as promoters, tags, culture pH, aeration and temperature, may not be sufficient to successfully promote the expression of MP to obtain high yields of protein (Bill et al., 2011). Instead, a more targeted approach, evaluating and tailoring the components involved in MP biogenesis and inactivating or deleting proteases that interfere with MP expression, increasing the levels of chaperones associated with MP production, may lead to improved expression levels (Bill et al., 2011; Wagner et al., 2006). Alternatively, the use of cell-free protein synthesis systems, which continue to be refined in order to overcome the general limitations associated with conventional cellular expression systems (Carlson et al., 2012; Junge et al., 2011), may also be considered. Advantages include the rapid production of high protein yields using cost-effective strategies, as well as active monitoring, sampling and possible optimisation of the protein synthesis process, which is made possible due to the lack of the cell wall (Carlson et al., 2012). However, even this system remains to be optimised especially when used for the expression of MPs, which continue to be problematic due to their hydrophobic nature (Liguori et al., 2008).

Reassessment of both the overexpression and purification of QacC may be needed, especially for future studies needing large quantities of QacC. Continued optimisation should be focussed on a different expression vector or host *E. coli* cells which may allow for greater control and expression levels of QacC. Alternatively, a cell-free system may be analysed and possibly optimised for QacC expression. Purification may also need to be further explored, possibly utilising the 1D4 tag as an alternative to the Hisx6 tag to increase the final yield of purified QacC.

4.8 Conclusion

This study established a method for the expression and purification of the QacC protein from both the pBAD and pBluescript II SK-based expression systems. The pBAD system employed induction using L-arabinose and was driven via the P_{BAD} promoter whereas the pBluescript II SK system was induced via IPTG and used the T7 RNA polymerase and T7 promoter system. Although sufficient amounts of purified QacC needed for binding assays were difficult to obtain, the expression and purification of the pBAD-based expression system was optimised to produce high yields of purified QacC and mutant QacC proteins, which could be used for functional and structural assessment in this study.

CHAPTER 5 – CYS-SCANNING MUTAGENESIS OF QACC

5.1 Introduction

Difficulties in obtaining crystal structures of membrane transport proteins, as discussed in Chapter 3, has led to the development of computer modelling, which serves as an alternative method for protein structural analysis. However, crystal structures and models generated by computer prediction algorithms produce not only a static structure but a structure which is more often than not, not in its native environment. Therefore, these predictions need to be biochemically assessed to determine if the native state of the protein, influenced by its surrounding environment, is represented by the model. One of the methods developed to analyse transmembrane protein topology is Cys-scanning mutagenesis. This method is based on the generation of mutants in which residues are selectively changed to a Cys residue and then challenged with various sulphydryl reagents to evaluate exposure and accessibility (Mordoch et al., 1999). Their reactivity with sulphydryl-specific reagents makes them a good target for site-specific chemical modification, whilst their rarity in proteins (Pe'er et al., 2004) makes it simple to generate protein derivatives which contain only one or a few Cys residues, making it easy to probe and evaluate protein topology.

Cys-scanning mutagenesis has been used to analyse both the structure and function of a wide range of MPs, including the *E. coli* lactose permease LacY (Abramson et al., 2003a; Frillingos et al., 1998; Sahin-Tóth et al., 1994). Analysis of various LacY mutants revealed a number of amino acids identified in substrate translocation as well as those involved in proton translocation. The results obtained in these studies were then used to generate a model of LacY. This model has since been shown to be highly consistent with the solved 3D crystal structure of LacY (Abramson et al., 2003b), highlighting the importance and relevance of Cys-scanning mutagenesis in structure and function analysis of MPs. Other membrane transporters successfully evaluated by site-directed mutagenesis include the AcrB multidrug efflux pump (Husain and Nikaido, 2010; Murakami et al., 2004), the *S. aureus* QacA (Xu et al., 2006) and TetA(K) MPs (Hassan et al., 2006b), as well as

eukaryotic transporters such as P-glycoprotein (Liu and Sharom, 1996; Loo et al., 2008; Loo and Clarke, 1995, 1999; Zolnerciks et al., 2007) and the glucose transporter Glut1 (Hruz and Mueckler, 1999; Mueckler and Makepeace, 2002, 2005). In addition to evaluating protein topology, this method, in conjunction with various membrane impermeable thiol-specific reagents, also allows for the identification and analysis of residues which delineate the substrate translocation pathway and/or play a role in substrate binding, as well as those which may constitute a part of a water filled channel. It can also be used in combination with chemical cross-linking for structural assessment, such as TM packaging within the membrane (Liu et al., 2010; Zhang et al., 2002), domain arrangement (Zolnerciks et al., 2007) as well to analyse the role of essential residues involved in the alternating access mechanism, discussed throughout Chapter 1 (Zhou et al., 2012).

In this chapter Cys-scanning mutagenesis was used to evaluate the structure of the QacC membrane transporter. A functional QacC protein was constructed devoid of its one native Cys residue and used as the template for Cys-scanning mutagenesis. Replacement of individual amino acids with a single Cys residue was carried out sequentially until all QacC residues were mutated; resulting in 103 single and 2 double Cys mutants. Accessibility of each introduced Cys residue was investigated on purified protein with the membrane-impermeable sulphydryl-specific reagent fluorescein-5-maleimide. Whole cell labelling of selected QacC-Cys mutants was also carried out with fluorescein-5-maleimide as well as with eosin-5-maleimide, a bulkier thiol-specific reagent, to assess QacC membrane orientation.

A number of fluorescein-5-maleimide reactive residues found in TMS1 were assessed by solvent accessibility studies in order to determine their role in substrate binding and/or translocation. Lastly, as positively-charged amino acids such as Lys and Arg define the “positive-inside” rule (Section 1.5.1) (Heijne, 1986; von Heijne, 1992), the four QacC Lys residues were evaluated for their position within QacC. Given the distribution of these amino acids in QacC, with Lys49 being the only positively-charged residue predicted to lie on the periplasmic face of the membrane, Lys49 was assessed for its role as a structural determinator.

5.2 Site-directed mutagenesis of QacC

5.2.1 Construction of a QacC Cys43Thr derivative

The native QacC protein contains only a single Cys residue (Cys42); the role of which was previously investigated by mutagenesis (Paulsen et al., 1995) where it was found that substitution of Cys to Thr retained WT QacC expression and function. In the codon-optimised version of QacC used throughout this study, as a result of the introduction of an additional Gly residue (Gly2) as discussed in Section 4.2.1, the position of Cys42 changes to Cys43. In this background, Cys43 was replaced by Thr using site-directed mutagenesis (Section 2.6.2), producing pBADQacC_C43T. To confirm that this change did not alter expression or function of QacC, a Western blot of the C-terminally Hisx6 tagged QacC and Cys43Thr, expressed from *E. coli* cells carrying pBADQacC and pBADQacC_C43T was conducted (Sections 2.9.2 and 2.9.3). To assess functionality, a fluorimetric ethidium transport assay using *E. coli* TOP10 cells expressing both proteins was carried out (Section 2.11.1). As seen in Figure 5.1 and 5.2, both expression and function of Cys43Thr is consistent with that of WT QacC. Therefore, this version of QacC was used as a template for all further site-directed mutagenesis and all results were compared to this mutant; as such it was subsequently designated as the WT QacC.

5.2.2 Construction of a bank of QacC Cys-substituted mutants

The construction of Cys-substituted QacC mutants was carried out using the codon optimised pBAD-based Cys-less QacC (pBADQacC_C43T) as a template (Sections 4.3 and 2.6.2). Initial attempts at site-directed mutagenesis included the incorporation of a silent restriction site to the pair of complementary mutagenic primers (Table 2.4), which were then used to aid in the screening procedure to identify the mutant. Although mutants could be identified that had the restriction site incorporated and the mutation to Cys, numerous mutants appeared to have multiple primers incorporated into the DNA, which were only identified after sequencing (Section 2.5.4). As these insertions carried the silent restriction site,

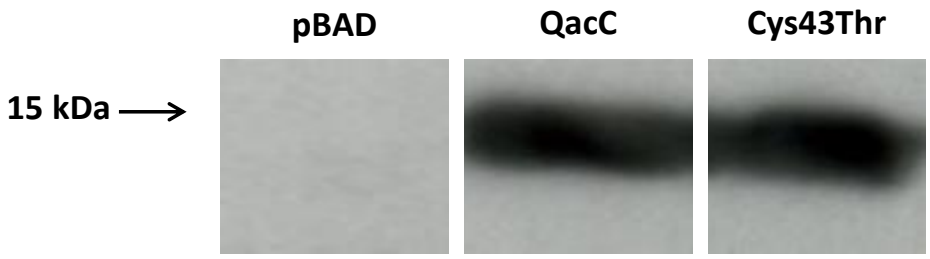

Figure 5.1: Expression of QacC and the QacC mutant Cys43Thr from pBAD in *E. coli* TOP10 cells

Cultures of *E. coli* TOP10 cells carrying pBADQacC and pBADQacC_C43T (Table 2.3) were grown to OD_{600}=0.55 in LBAmp and induced for two hours with 0.02% L-arabinose (Section 2.7.1); cells were collected and membrane fractions isolated (Section 2.7.3). Membrane fractions were solubilised with 1% (w/v) DDM, and 20 µg of each sample was subjected to SDS-PAGE (Section 2.8.1), transferred to a PVDF membrane (Section 2.9.2) and probed with an anti-HISx6 antibody (Section 2.9.3). Locations of the QacC and Cys43Thr protein bands (at 15 kDa) are indicated on the left-hand side of the Western blot.

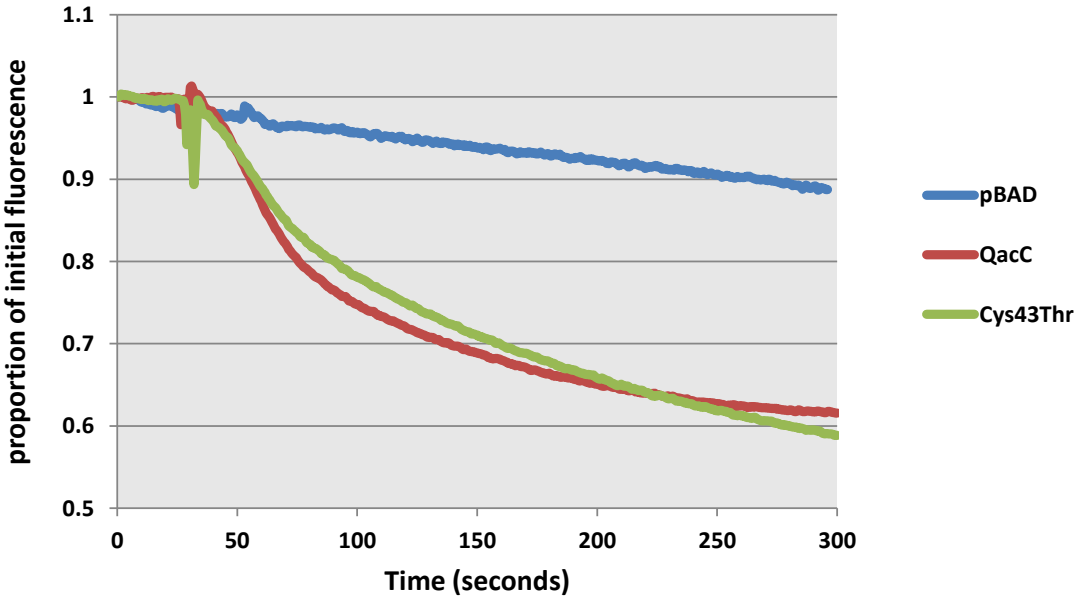

Figure 5.2: Fluorimetric assay of ethidium export by QacC and QacC Cys43Thr

Fluorimetric assays were carried out on *E. coli* TOP10 cells carrying pBADQacC and pBADQacC_C43T (Table 2.3). Cultures were grown to OD_{600}=0.55 in LBAmp and induced for two hours with 0.02% L-arabinose (Section 2.7.1). Cells were collected and loaded with 15 µM ethidium bromide (Section 2.11.1) and ethidium efflux energised by the addition of 160 mM sodium formate at time zero (Section 2.11.1). Colours represent: blue-*E. coli* TOP10 cells carrying an empty pBAD vector, red-WT QacC, green-the QacC mutant Cys43Thr, subsequently designated WT QacC.

screening by digestion failed to detect them, as the resulting digestions generated a number of small fragments, less than 20 bps, which were too small to be detected on a 1% agarose gel. Optimisation of the melting and annealing temperature for each primer did not eliminate these insertions, nor did the adjustment of the concentration of primers added. Evaluation of the final concentration of Mg^{2+}, DNA polymerase and the addition of DMSO were all also explored, however, these too did not eliminate the unwanted inserts. Thus the incorporation of a silent restriction site for screening was abandoned and instead PCR screening was used on all newly generated mutants (Section 2.6.3). Using the pBAD sequencing primers (Table 2.4), each clone was screened by PCR and resulting fragment viewed on a 1% agarose gel. The size of each band was then compared to a positive control, pBADQacC_C43T which was included in each PCR. Although greatly improving the screening process of all newly constructed mutants, this method still failed to identify short primer insertions. To address this issue, the agarose gel concentration was increased from 1 to 2% in order to increase band separation of smaller fragments (Figure 5.3); however, short inserts up to 20 base pairs still could not be easily identified, and could only be verified by DNA sequencing (Section 2.5.4). Therefore, certain mutants required numerous individual clones to be sequenced until one was found which was free of additional insertions.

To evaluate the function and secondary structure of the QacC protein, each amino acid from position three to 105 in the Cys-less QacC protein was thus replaced one-by-one with a Cys residue. In addition to the 103 single Cys replacements, nine double mutants (Table 2.3) were also generated to assess QacC function and structure (Sections 5.5 and 7.3).

5.2.3 Evaluation of expression of QacC Cys-mutants

To evaluate the effect of the mutation on protein expression, *E. coli* TOP10 cells expressing each of the 103 pBAD-based QacC mutants were grown following conditions outlined in Section 4.3.1. A Western blot using the anti-Hisx6 tag antibodies (Section 2.9.2 and 2.9.3) was then carried out on isolated membrane

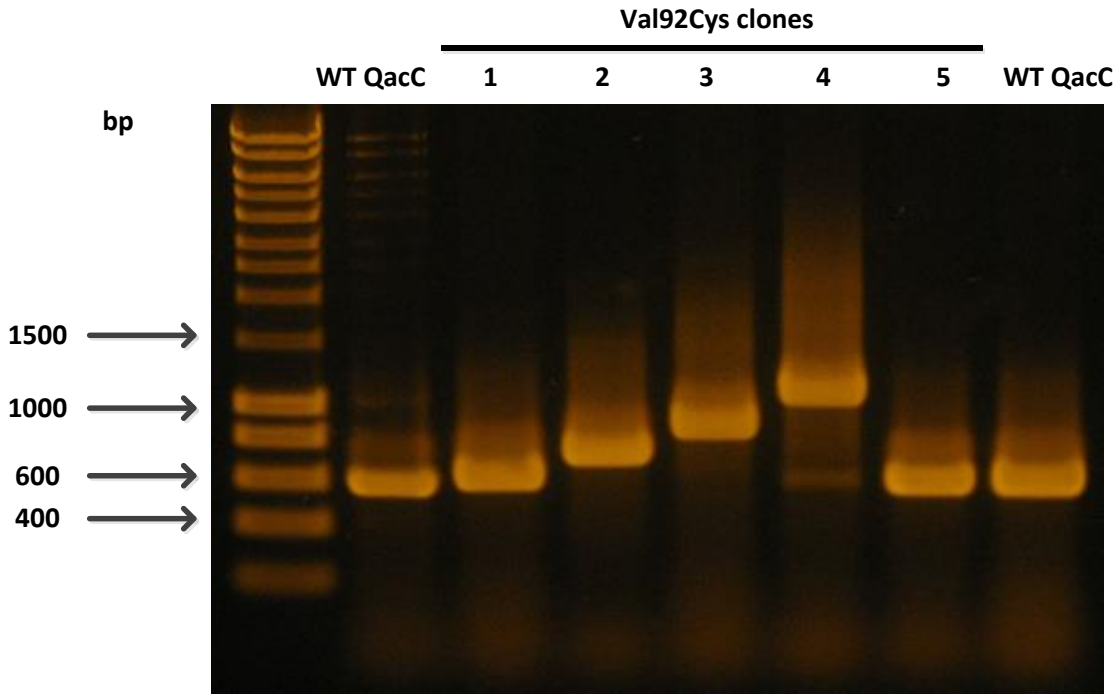

Figure 5.3: Colony PCR screening to identify a QacC Val92Cys mutant

E. coli TOP10 cells were transformed (Section 2.4) with *Dpn*I-treated PCR amplified pBADQacC_V92C mutant plasmid. Single colonies were assessed by colony PCR screening (Section 2.6.3) using the pBAD forward and reverse sequencing primers (Table 2.4). Samples were subjected to electrophoresis on a 2% (w/v) agarose gel (Section 2.5.2) and compared to pBADQacC_C43T. Of the five colonies used in the PCR screening of this mutant, number 5 appears to be comparable in size to the two WT QacC controls, and as such was DNA sequenced and the mutation verified. The positions of the MW markers (HyperLadder™ 1kb) are shown on the left hand side.

fractions (Section 2.7.3), and revealed that all pBAD-based mutants were expressed, with no significant deleterious effect seen in any QacC-Cys mutant. Some variation was detected with three mutants showing significantly lower expression rates, and 14 others showing reduced rates when compared to WT QacC (Figure 5.4).

5.3 Solvent accessibility analysis of the QacC mutants using thiol-reactive reagents

5.3.1 Fluorescein-5-maleimide labelling assay of QacC mutants

The level of reactivity of fluorescein-5-maleimide with each Cys-substituted QacC mutant was assessed to determine the solvent accessibility of every amino acid position in QacC (Section 2.14.1). Initial attempts to label the QacC-Cys mutants met with a number of complications which included; uneven labelling of samples, difficulties binding the labelled proteins to the ProBond^{TM} Ni-resin and effectively eluting the bound proteins off the beads. Labelling with fluorescein-5-malemide was initially carried out using 50 µl of a protein membrane fraction and 0.5 mM fluorescein-5-malemide. Although the final protein concentration of membrane fractions was similar in all mutants (averaging at 4 to 5 mg/ml) this represented the total yield of all proteins within the membrane fraction, including QacC. However, the concentration of each QacC mutant protein within the membrane fraction varied from mutant to mutant. Attempts to standardise the amount of QacC labelled were successful for QacC-Cys mutants whose expression was equivalent to WT. However, difficulties were encountered when the levels of expression varied markedly and labelling equal amounts proved challenging. Membrane fractions used in fluorescein-5-maleimide labelling of QacC-Cys mutants were generally prepared from 200 ml of starting culture (Section 2.7.1), resulting in 4 to 5 mg/ml of total protein per sample. Attempts to purify QacC from these membrane fractions prior to labelling failed to generate sufficient concentrations of QacC for labelling to be carried out. This was not surprising given the difficulties encountered with QacC purification, even with larger starting volumes,

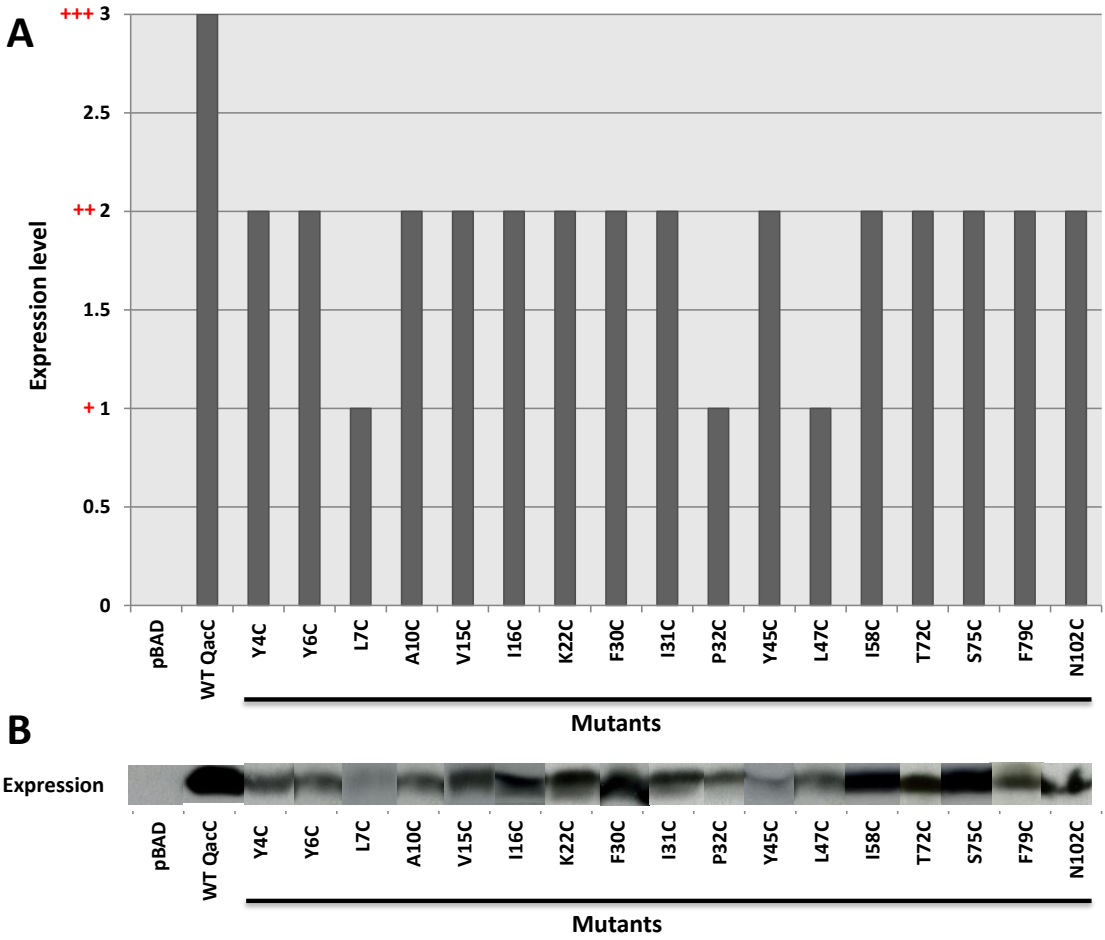

Figure 5.4: Expression levels of pBAD-based constructs as expressed from *E. coli* TOP10 cells

Cultures of *E. coli* TOP10 cells carrying pBAD, pBADQacC_C43T and all other QacC-Cys mutant plasmids derived thereof (Table 2.3) were grown to OD_{600}=0.55 in LBAmp and induced for two hours with 0.02% L-arabinose (Section 2.7.1). Cells were collected and membrane fractions isolated (Section 2.7.3). Membrane fractions were solubilised with 1% (w/v) DDM, and 20 µg of each sample was subjected to SDS-PAGE (Section 2.8.1), transferred to a PVDF membrane (Section 2.9.2) and probed with an anti-HISx6 antibody (Section 2.9.3). The relative level of expression was scored using the WT QacC as the control for maximum protein expression. Expression levels were assessed and put in a range of 26-50% as +, 51 to 75% as ++ and 76 to 100% as +++. (A) Only those which showed reduced levels of expression are shown in the graph, all other QacC Cys mutants were expressed at WT QacC levels. (B) Comparable Western blot of the QacC Cys mutants presented in the graph, with the WT QacC and empty pBAD vector included as positive and negative controls, respectively.

as discussed in Chapter 4. As a result, samples were resolved on a test SDS-PAGE gel, and their expression levels and labelling evaluated. Based on these results, samples were then standardised and resolved on a second SDS-PAGE gel.

As discussed in Section 4.6, purification of QacC was not without challenges. As purification was also required for fluorescein-5-maleimide labelling of the QacC-Cys mutants, optimisation of the various steps ensuring optimal yields of labelled QacC was carried out. Early attempts to purify labelled QacC-Cys mutants revealed that not all of the labelled samples were binding to the ProBond^{TM} Ni-resin and those that were, were subsequently not being fully eluted off the beads. To increase binding efficiency, the duration of binding was tested (initially one to two hours at 4°C) as was the concentration of imidazole added to the sample buffer (200 mM to 400 mM) (Table 2.1) and the duration of elution (30 minutes at RT to two hours at RT and overnight at 4°C) (Figure 5.5). As can be seen from Figure 5.5, higher protein yields were obtained with samples which were eluted off the Ni-resin following the addition of sample buffer supplemented with 200 to 400 mM imidazole (pH 8.0) with an incubation time of up to two hours at RT. The duration of binding did not appear to have a noticeable effect on increasing the binding efficiency as all samples showed unbound protein in the quenching/binding buffer when resolved on by SDS–PAGE (results not shown). Binding was tested by incubating the beads with the protein preparation overnight at 4°C; however, unbound protein was still detected in the binding buffer. Given these results, binding was subsequently undertaken for two hours at 4°C, for maximum results, and elution carried out at RT for 30 minutes, in sample buffer supplemented with 400 mM imidazole (pH 8.0). An illustration of the above optimised conditions, as presented on a sub-set of QacC-Cys mutants is presented in Figure 5.6.

All 103 pBAD-based QacC-Cys mutants were labelled with fluorescein-5-maleimide as were the two double mutants used to assess the Lys residue at position 49 (Section 5.5). Glu25Cys was identified early on to be highly reactive with fluorescein-5-maleimide and was thus used to represent 100% labelling, whilst WT QacC, due to a lack of a Cys residue, was used throughout the study as a negative control.

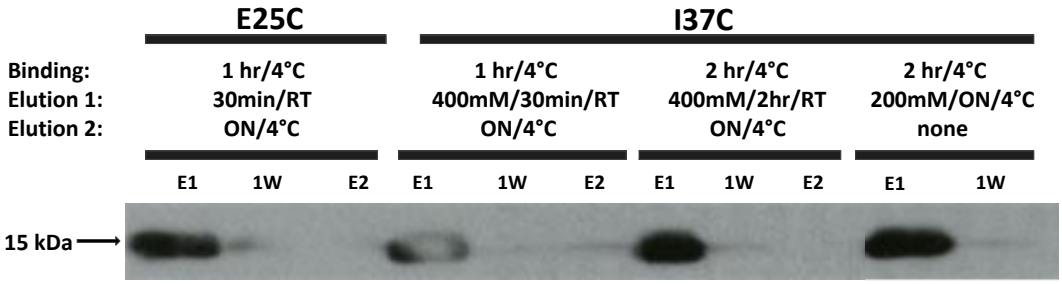

Figure 5.5 Examination of purification conditions for QacC Glu25Cys and Ile37Cys mutants labelled with fluorescein-5-maleimide

Western blot detecting fluorescein-5-maleimide labelled QacC mutant Glu25Cys and Ile37Cys proteins, purified as described in Section 5.3.1. Samples were subjected to SDS-PAGE (Section 2.8.1), transferred to a PVDF membrane (Section 2.9.2) and probed with an anti-HISx6 antibody (Section 2.9.3). Glu25Cys was used as a positive control as it is highly reactive with fluorescein-5-maleimide and is expressed and purified at high levels, whilst Ile37Cys was shown to have low reactivity with fluorescein-5-maleimide, and although expressed at WT levels, low levels were obtained following its purification. Conditions examined included both binding and elution. Samples were tested of eluents (E1 and E2), first wash (1W) and from the binding buffer prior to first wash. Locations of the QacC protein bands (at 15 kDa) are indicated on the left-hand side of the Western blot.

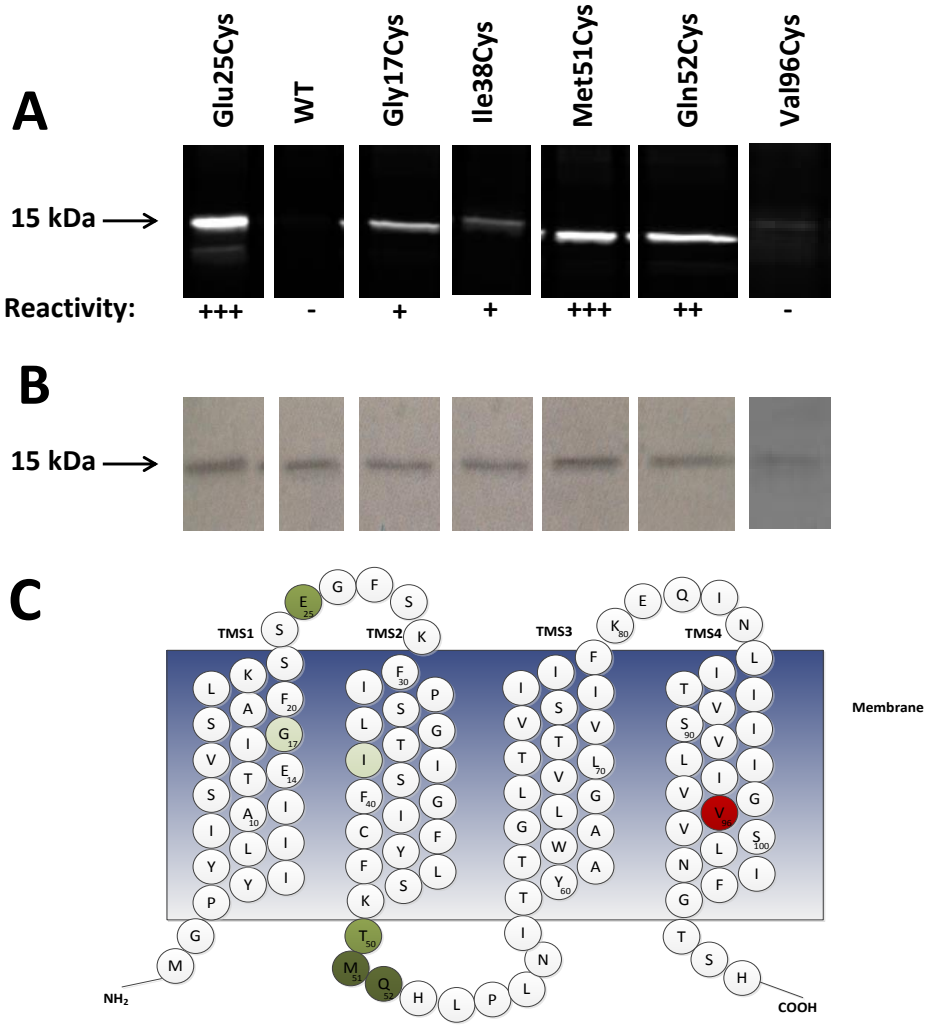

Figure 5.6: Effect of fluorescein-5-maleimide labelling on a sub-set of QacC-Cys mutants

An illustration of the results obtained from fluorescein-5-maleimide binding assays carried out on QacC Cys mutants, as described in Section 2.14.1. Labelling was performed in disrupted membranes (Section 2.7.3) on an equal amount of total MPs. Following labelling, samples were purified (Section 2.7.4) and subjected to SDS-PAGE (Section 2.8.1) (A) Reactivity was detected using a Bio-Rad Gel Doc EZ imager (Bio-Rad) and scored by comparison with Glu25Cys, which exhibited high levels of reactivity with fluorescein-5-maleimide and WT QacC, which lacking a Cys residue was used as a negative control. Reactivity levels were assigned visually and put in a range of 0–25% as -, 26-50% as +, 51–75% as ++, 76–100% as +++. (B) To confirm loading of equal amount of protein, gels were stained by Coomassie Brilliant Blue R–250 (Section 2.9.1). Locations of the QacC protein bands (at 15 kDa) are indicated on the left-hand side of the SDS-PAGE gel and Western blot. (C) Putative locations of assessed residues as shown on a 2D model of QacC. Reactivity of each residue, as determined by labelling, is given as dark green - +++ (highly reactive), green - ++ (medium reactivity), light green - + (minimal reactivity) and red - - (non-reactive).

Whole cell fluorescein-5-maleimide labelling (Section 2.14.2) was carried out on four QacC-Cys mutants to assess the orientation of the QacC monomers within the membrane. Results obtained for these would help assess the theory that SMR proteins function as dimers, with their monomers arranged in an antiparallel orientation, as observed for EmrE (Lloris-Garcera et al., 2012; Nara et al., 2007). The selection of the four mutants chosen for assessment was based on their strong reactivity with fluorescein-5-maleimide in disrupted membranes, suggesting their location within QacC is in a loop region or possibly at the membrane boundary of the protein. As such, each one of these four QacC-Cys mutants represents an inserted Cys residue within each of the four QacC loops: Glu25Cys, loop 2; His53Cys, loop 3; Ile76Cys, loop 4; and Ile103Cys, loop 5, as shown in the final structural model of QacC, Figure 5.19. Labelling of these mutants, even those purported to be located in a cytoplasmic loop, may suggest that monomeric QacC integrates into the membrane in two orientations (N-in/C-out and N-out/C-in). As can be seen (Figure 5.7), although varying in the degree of reactivity, all four QacC-Cys mutants were labelled with fluorescein-5-maleimide, irrespective of their location in the protein, indicating that QacC monomers appear to be inserted into the membrane in two orientations. Given evidence that QacC also functions as dimer, as discussed in Chapter 7, this might suggest that QacC forms a dimer with monomers taking on an anti-parallel organisation.

5.3.2 Eosin-5-maleimide labelling of QacC mutants

To further investigate QacC topology and the possible orientation of each monomer within the membrane, labelling of a subset of QacC-Cys mutants was undertaken with eosin-5-malemide. With a MW of 742.9 kDa, it is considered to be a more bulky reagent than fluorescein-5-maleimide (427.4 kDa). Following the same methodology used to label QacC-Cys mutants with fluorescein-5-maleimde in whole cells (2.14.2), the reactivity of eosin-5-maleimide was examined with the QacC mutants, Glu25Cys, His53Cys, Ile76Cys and Ile103Cys, utilising WT QacC as a negative control (Figure 5.8). Comparatively, the fluorescence of the

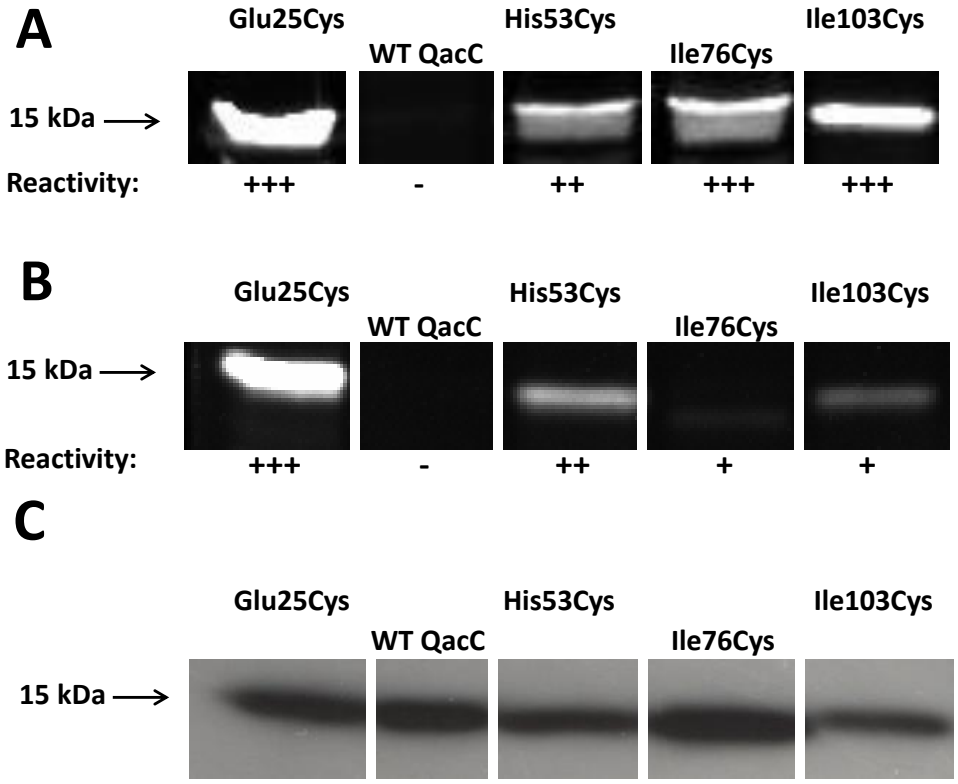

Figure 5.7: Whole cell fluorescein-5-maleimide labelling with selected QacC mutants

Fluorescein-5-maleimide reactivity of QacC mutants containing Cys substitutions at Glu25, His53, Ile76 and Ile103. (A) Labelling in disrupted cells (B) whole cells (Sections 2.14.1 and 2.14.2). Labelled QacC-Cys mutant proteins were purified (Section 2.7.4) and subjected to SDS-PAGE (Section 2.8.1) before being visualised using the Bio-Rad Gel Doc EZ imager (Bio-Rad), and scored based on their reactivity levels. These were assigned visually and put in a range of 0-25% as -, 26-50% as +, 51–75% as ++, 76–100% as +++. (C) Expression levels of QacC and QacC-Cys mutant proteins were determined by Western blotting using an anti-Hisx6 antibody (Sections 2.9.2 and 2.9.3). Locations of the QacC protein bands (at 15 kDa) are indicated on the left-hand side of the SDS-PAGE gels and Western blot.

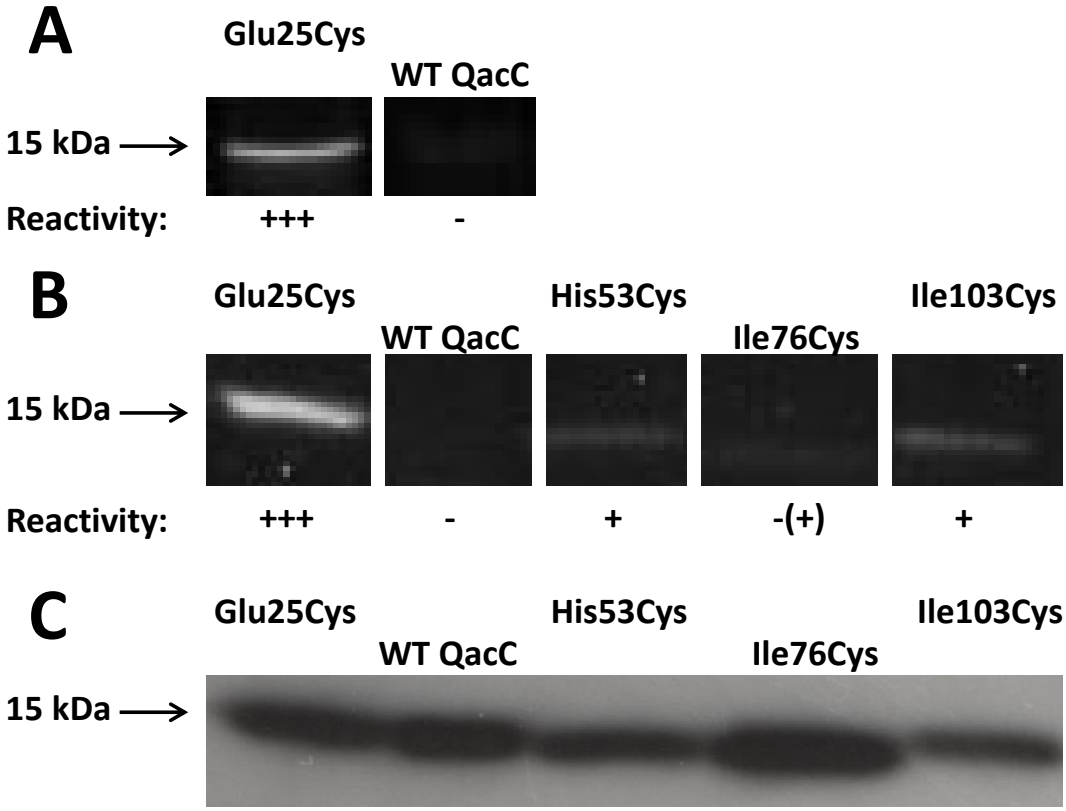

Figure 5.8: Whole cell eosin-5-maleimide labelling with selected QacC mutants

Eosin-5-maleimide reactivity of QacC-Cys mutants containing Cys substitutions at GLu25, His53, Ile76 and Ile103. (A) Labelling in disrupted cells (B) whole cells (Sections 2.14.1 and 2.14.2). Labelled QacC-Cys mutant proteins were purified (Section 2.7.4) and subjected to SDS-PAGE (Section 2.8.1) before being visualised using the Bio-Rad Gel Doc EZ imager (Bio-Rad), and scored based on their reactivity levels. Reactivity levels were compared to Glu25Cys and WT QacC, these were assigned visually and put in a range of 0-25% as -, 26-50% as +, 51–75% as ++, 76–100% as +++. (C) Expression levels of QacC and QacC-Cys mutant proteins were determined by Western blotting using an anti-Hisx6 antibody (Sections 2.9.2 and 2.9.3). Locations of the QacC protein bands (at 15 kDa) are indicated on the left-hand side of the SDS-PAGE gels and Western blot.

QacC Glu25Cys mutant, which is highly reactive with fluorescein-5-maleimide (Figure 5.7), was much lower with eosin-5-maleimide, as seen in reactions obtained from membrane fractions of Glu25Cys, which label more efficiently than that of whole cells. Attempts to increase reactivity by adding higher concentrations of eosin-5-maleimide and extending the duration of labelling did not result in an improved fluorescence signal. Nonetheless, although a reduced fluorescence signal was observed, the results mirror those obtained in Figure 5.7, with Glu25Cys, His53Cys and Ile103Cys all showing reactivity with eosin-5-maleimide. This could indicate that QacC monomers adopt alternative orientations within the membrane. Reactivity of the QacC Ile76Cys mutant to eosin-5-maleimide was very low, similar to observations with fluorescein-5-maleimide. Therefore, this residue may lie on the border of the membrane, and when in its native environment, the reactive side-chain may be partially buried within the membrane, resulting in a quenched reaction.

5.4 Solvent accessibility of residues within TMS1

The accessibility of fluorescein-5-maleimide is limited to binding Cys residues located in a non-hydrophobic environment (McLachlin and Dunn, 1996). This includes residues located in loop regions of a protein or those that form substrate translocation pathways which are fluid filled (Frillingos et al., 1998). Thus, to evaluate the substrate binding and translocation pathway in QacC the level of reactivity of fluorescein-5-maleimide with a subset of single Cys-substituted QacC TMS1 mutants, clustering Glu14, were analysed (Section 2.14.1). This region was selected as Glu14, located within TMS1, has been identified in SMR proteins to be an essential residue for substrate binding and translocation (Grinius and Goldberg, 1994; Yerushalmi and Schuldiner, 2000). Reactivity of these QacC-Cys mutants was assessed on membrane fractions (Section 2.7.3), which were initially pre-incubated with the QacC substrate ethidium (Section 2.14.1), as it was hypothesised that the addition of this substrate would block labelling of amino acids which make up the ethidium substrate-binding site and translocation pathway. Although shown to be a strong substrate of QacC, as determined by fluorimetric transport assays (Section

6.2) and resistance profiling (Section 6.3), pre-incubation with ethidium at various concentrations did not affect the fluorescein-5-maleimide reactivity levels observed for Glu14Cys. Therefore a second QacC substrate was tested, pyronin Y. Although shown to be ineffective with ethidium, initial trials on the QacC Glu14Cys mutant revealed that the addition of 150 µM pyronin Y reduced the binding of fluorescein-5-maleimide (Figure 5.9). In order to determine the concentration of pyronin Y needed to completely block labelling, the QacC Glu14Cys mutant was pre-incubated with 500 µM, 1 mM and 1.5 mM pyronin Y. Results revealed that blocking of fluorescein-5-maleimide began at 150 µM and increased to near total blocking at 1.5 mM (Figure 5.10). Based on these results pre-incubation was carried out with 1.5 mM pyronin Y to achieve maximum blocking to fluorescein-5-maleimide.

Accessibility studies coupled with pyronin Y pre-incubation were carried out on QacC-Cys mutants predicted to be in TMS1 and which were found to react with fluorescein-5-maleimide (Figure 5.11) and postulated to line the substrate binding and translocation pathway (Figure 5.12). Of the QacC-Cys mutants examined, which included Ala10Cys, Glu14Cys, Gly17Cys, Ser18Cys, Ala19Cys and Leu21Cys, pre-incubation with pyronin Y consistently reduced the reactivity of only one tested mutant, Glu14Cys. Faint quenching of fluorescein-5-maleimide reactivity was also observed in the QacC Gly17Cys and Ala19Cys mutants, however, variation in the level of quenching was observed between experiments, making it difficult to evaluate the final results.

5.5 Analysis of the QacC Lys residues – topological determinants?

As positively-charged residues Arg and Lys have been identified to be more prevalent in cytosolic connecting loops (Section 1.5.1) and are considered topogenic determinants of MPs (Boyd and Beckwith, 1989; von Heijne, 1986) they were further evaluated in this study. The QacC membrane transporter contains four positively-charged residues, all of which are Lys. To evaluate these and assess their putative position in QacC, each Lys residue was mutated to Cys (Section 4.3, Section 2.6.2) and accessibility to fluorescein-5-maleimide assessed in disrupted

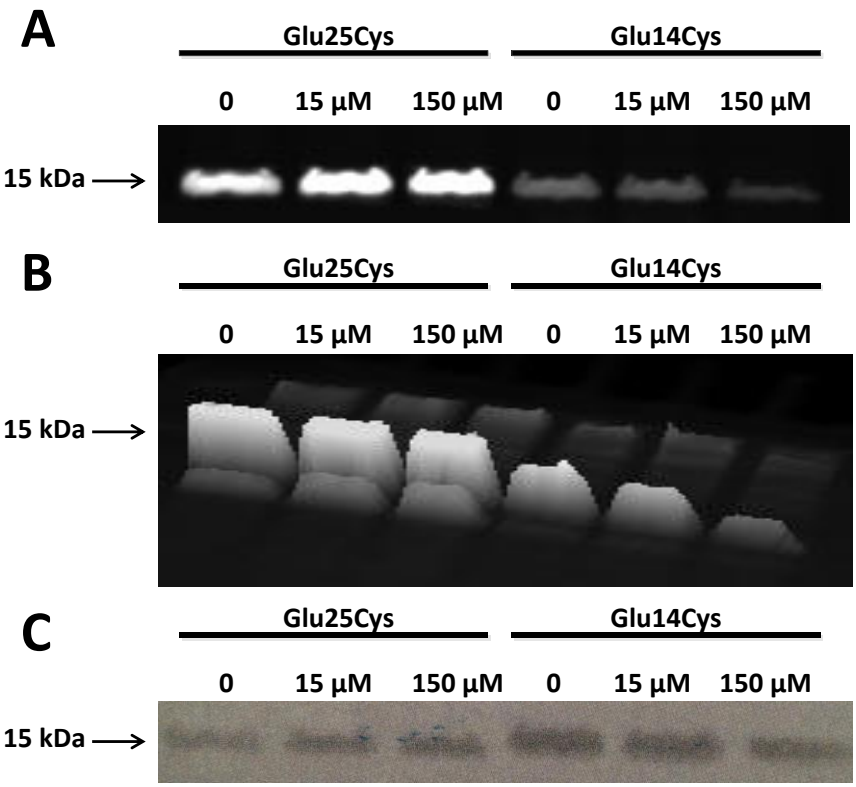

Figure 5.9: Accessibility of fluorescein-5-maleimide to Glu25Cys and Glu14Cys mutants following pre-incubation with pyronin Y

Fluorescein-5-maleimide reactivity of QacC-Cys mutants was carried out following pre-incubation with 0, 15 and 150 µM pyronin Y. Labelling was carried out in disrupted membranes essentially as described in Section 2.14.1. Fluorescein-5-maleimide treated QacC-Cys mutant proteins were purified (Section 2.7.4) and subjected to SDS-PAGE (Section 2.8.1) before being visualised using (A) the Bio-Rad Gel Doc EZ imager (Bio-Rad) and (B) as shown in profile view, with higher peaks indicting higher reactivity. Reactivity levels of both Glu25Cys and Glu14 Cys were compared to samples not treated with pyronin Y, representing 100% labelling of each sample. (C) Expression levels of labelled QacC-Cys mutant proteins were detected by Coomassie Brilliant Blue R–250 staining (Section 2.9.1). Locations of the QacC mutant protein bands are indicated on the side of the SDS-PAGE gels with the 15 kDa arrow.

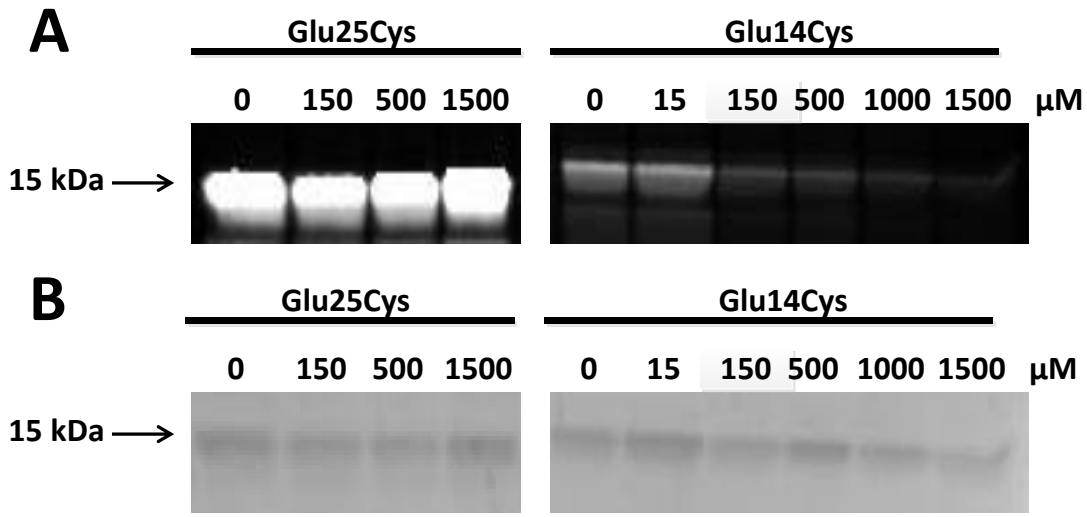

Figure 5.10: Accessibility of fluorescein-5-maleimide to Glu25Cys and Glu14Cys mutants following pre-incubation with pyronin Y

Fluorescein-5-maleimide reactivity of QacC-Cys mutants was carried out following pre-incubation with 0, 150, 500 and 1500 µM pyronin Y. Labelling was carried out in disrupted membranes essentially as described in Section 2.14.1. Fluorescein-5-maleimide treated QacC-Cys mutant proteins were purified (Section 2.7.4) and subjected to SDS-PAGE (Section 2.8.1) before being visualised using (A) the Bio-Rad Gel Doc EZ imager (Bio-Rad). Reactivity levels of both Glu25Cys and Glu14Cys were compared to samples not treated with pyronin Y, representing 100% labelling of each sample. (B) Expression levels of labelled QacC–Cys mutant proteins were detected by Coomassie Brilliant Blue R–250 staining (Section 2.9.1). Locations of the QacC protein bands (at 15 kDa) are indicated on the left-hand side of the SDS-PAGE gels.

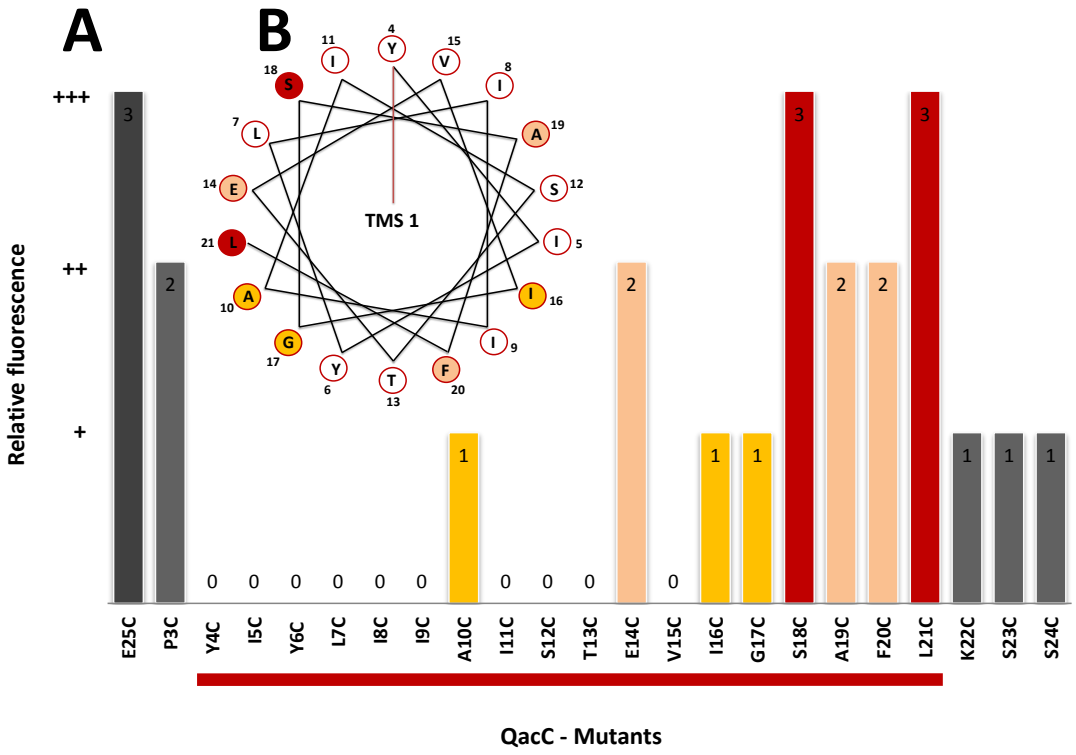

Figure 5.11: Accessibility of fluorescein-5-maleimide to QacC TMS1 Cys-substituted mutants

(A) Reaction conditions were essentially described in Section 2.14.1 whilst scoring of each Cys replacement was undertaken as set out in the legend to Figure 5.6. The proposed extents of TMS1 are indicated by the red bar beneath the graph. (B) Reactive residues, indicated with red (+++), orange (++) and yellow (+) are mapped onto an α-helix wheel, generated by Heliquest (Section 2.17.4) depicting TMS1 with the QacC membrane transporter.

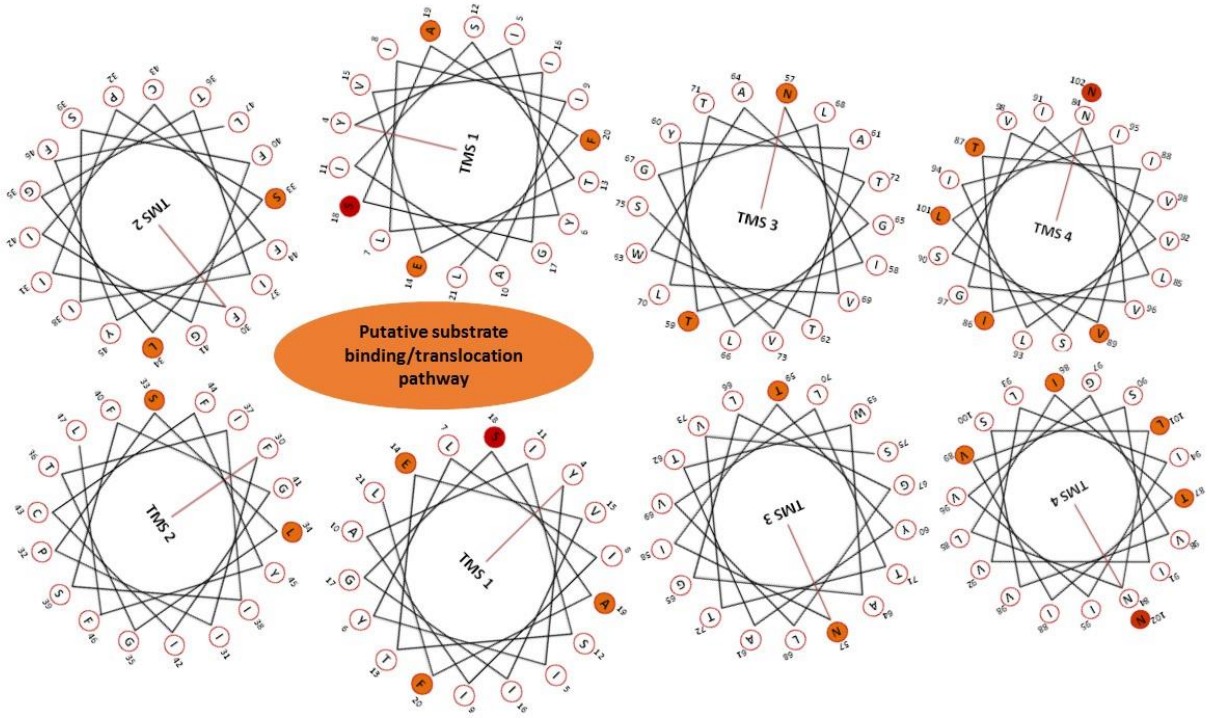

Figure 5.12: Proposed two-dimensional model of QacC in its dimeric form, based on fluorescein-5-maleimide labelling data

The secondary structure of the QacC multidrug transporter in its dimeric form, with the allocation of residues to each TMS based on results obtained from fluorescein-5-maleimide labelling (Section 5.3.1). The helical wheels, generated by Heliquest (http://heliquest.ipmc.cnrs.fr/) (Section 2.17.4), were amended and arranged to represent the putative location of each TMS within the QacC dimer. This arrangement, based on the solved structure of EmrE (Ubarretxena-Belandia et al., 2003), shows the putative location of each monomer within the dimer as well as the possible substrate binding/translocation pathway, denoted by an orange oval. Residues shown to be reactive with fluorescein-5-maleimide are shaded in orange and red, indicative of those which showed a reactivity of ++ and +++, as described in Figure 5.7.

membranes (Section 2.7.3 and 2.14.1) (Figure 5.13). As can be seen in the secondary structure model of QacC (Section 3.7) there is a three to one distribution of Lys residues within QacC, with Lys22, Lys29 and Lys80 positioned on the opposite side of the membrane to Lys49. Accessibility studies show that although all four Cys mutants were found to be reactive, the highest level of reactivity was exhibited by a Cys at position 29 (Figure 5.13).

To further analyse QacC Lys residues, and assess their possible role as topogenic determinants, six double mutants were generated, Lys22Cys_Lys29Cys, Lys22Cys_Lys49Cys, Lys22Cys_Lys80Cys, Lys29Cys_Lys49Cys, Lys29Cys_Lys80Cys and Lys49Cys_Lys80Cys (Section 2.6.2). Expression levels of these double mutants were determined by Western blot analysis (Section 2.9.2 and 2.9.3) and the capacity of the *E. coli* TOP10 cells expressing these mutants to efflux ethidium out of the cell was carried out by fluorimetric transport assays (Section 2.11.1) (Figure 5.14). Although none of the mutations seriously compromised the expression of QacC, functionality was shown to be reduced for all double QacC-Cys mutants, indicating that although a single Lys to Cys substitution was permissible, double mutations render QacC non-functional. Since fluorescein-5-maleimide labelling has indicted that QacC has only one highly reactive Lys residue, at position 49, that is not in a membrane environment (Figure 5.13), this residue was also mutated to Arg and Ala. A substitution of Lys49 to Arg was carried out to assess whether the positive charge is vital for protein function and topology, whereas the Lys49Ala mutation was created to determine whether the removal of a positively-charged residue from this position within QacC would have an effect on topology and could result in locking QacC into only one orientation. Using the QacC Lys49Ala mutant as a template, two double mutants, Glu25Cys_Lys49Ala and Gly105Cys_Lys49Ala, were generated via site-directed mutagenesis (Section 2.6.2), and their reactivity to fluorescein-5-maleimide assessed and compared with Glu25Cys and Gly105Cys (Figure 5.15). As discussed in Section 5.3.1. Glu25Cys was shown to be highly reactive with fluorescein-5- maleimide, such high reactivity was also shown for Gly105Cys, and as both are putatively allocated to loops found on opposite sides of the membrane, they represent good targets for the assessment of topology.

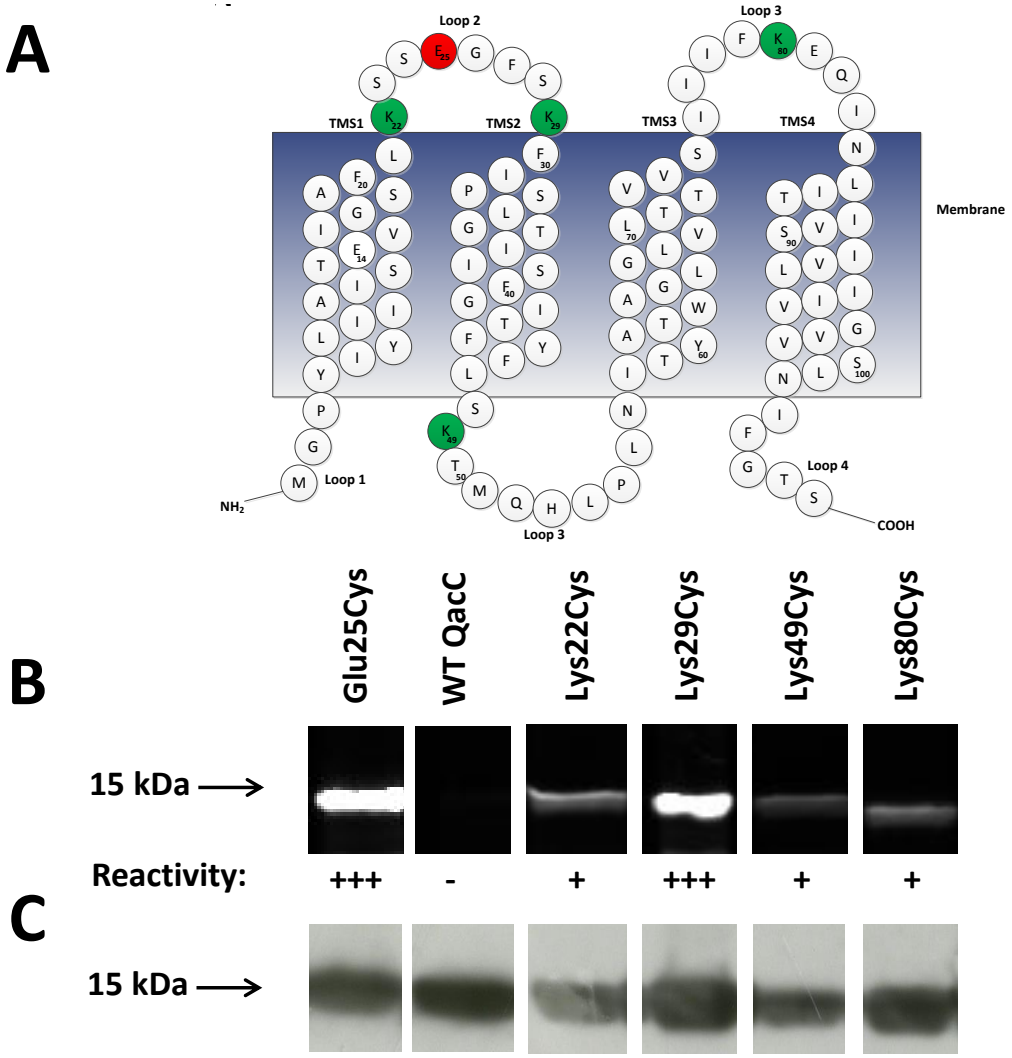

Figure 5.13: Accessibility of fluorescein-5-maleimide to Lys residues in QacC

(A) A model of the secondary structure of QacC showing the putative positions of the four Lys residues within QacC. Each of the four green highlighted Lys residues were substituted with Cys by site-directed mutagenesis (Section 2.6.2) and membrane fractions obtained from cells expressing each QacC mutant used in a fluorescein-5-maleimide binding assay. (B) Fluorescein-5-maleimide reactivity of QacC mutants containing Cys substitutions at Lys22, Lys29, Lys49 and Lys80. (A) Labelling was carried out in disrupted cells (Sections 2.14.1). Labelled QacC-Cys mutant proteins were purified (Section 2.7.4) and subjected to SDS-PAGE (Section 2.8.1) before being visualised using the Bio-Rad Gel Doc EZ imager (Bio-Rad), and scored based on their reactivity levels. Reactivity levels were compared to Glu25Cys and WT QacC as set out in Figure 5.6. Although lower levels of expression can be seen for Lys22Cys and Lys49Cys as determined by (C) a Western blot (Section 2.9.2 and 2.9.3), levels were sufficient for accessibility to be determined. Locations of the QacC protein bands (at 15 kDa) are indicated on the left-hand side of the SDS-PAGE gels and Western blot.

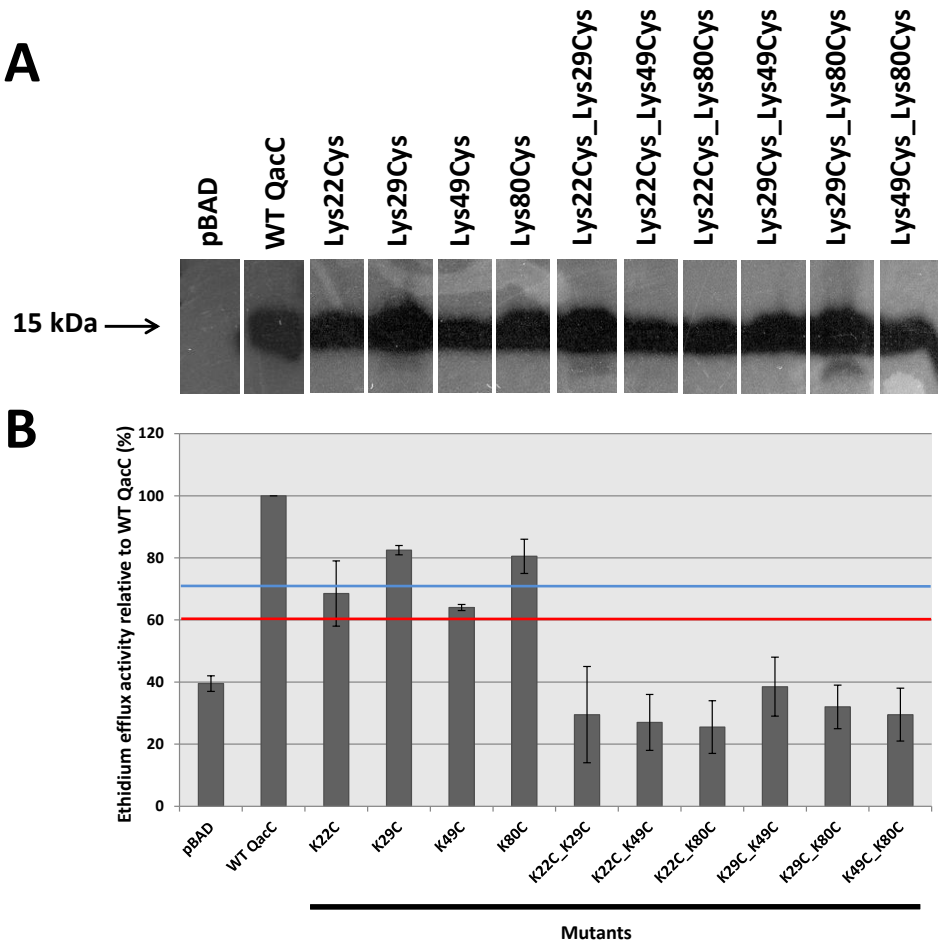

Figure 5.14: Protein expression drug efflux mediated by QacC Lys mutants

(A) Cultures of *E. coli* TOP10 cells carrying the pBAD-based QacC and Lys mutant plasmids (Table 2.3) were grown to OD_{600}=0.55 in LBAmp and induced for two hours with 0.02 % L-arabinose (Section 2.7.1). Cells were collected and membrane fractions isolated (Section 2.7.3). Membrane fractions were solubilised with 1% (w/v) DDM, and 20 µg of each sample was subjected to SDS-PAGE (Section 2.8.1), transferred to a PVDF membrane (Section 2.9.2). Locations of the QacC protein bands (at 15 kDa) are indicated on the left-hand side of the Western blot. (B) Fluorimetric assays were carried out on *E. coli* TOP10 cells as described in Section 2.11.1. Following expression cells were collected and loaded with 15 µM ethidium bromide and ethidium efflux energised by the addition of 160 mM sodium formate at time zero (Section 2.11.1). Reactions were observed over 100 seconds and efflux calculated as a percentage of WT QacC, which was indicative of 100% efflux. Mutants showing significantly impaired capacity for efflux (0-60% WT QacC) are shown below the red bar, whilst those with impaired capacity (61-70% WT QacC) are shown below the blue bar. Results above represent the average efflux of ethidium from at least three repeat experiments for each mutant, as labelled, with error bars representing the standard error of the mean for each mutant assessed.

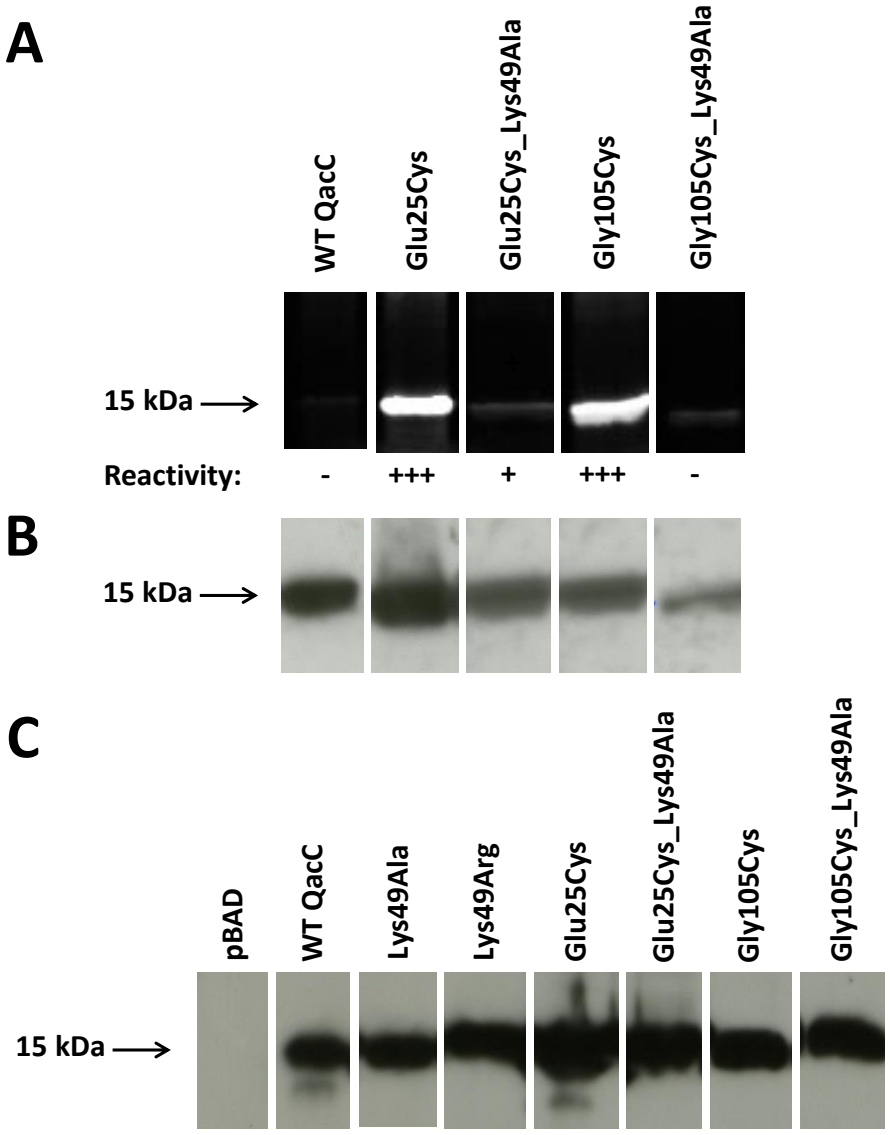

Figure 5.15: Fluorescein-5-maleimide reactivity and expression levels of QacC Lys49 mutants

Labelling was carried out in disrupted cells (Sections 2.14.1). Labelled QacC-Cys mutant proteins were purified (Section 2.7.4) and subjected to SDS-PAGE (Section 2.8.1) before being visualised using the (A) Bio-Rad Gel Doc EZ imager (Bio-Rad), and scored based on their reactivity levels. Reactivity levels were compared to Glu25Cys and WT QacC as set out in Figure 5.6. (B) Protein expression levels were detected by Western blot using the anti-Hisx6 antibody (Sections 2.9.2 and 2.9.3). (C) Protein expression levels were also determined on membrane fractions for the QacC Lys49Ala and Lys49Arg mutants by Western blot. Locations of the QacC protein bands (at 15 kDa) are indicated on the left-hand side of the SDS-PAGE gel and Western blots.

Results indicate a significant decrease in fluorescein-5-maleimide reactivity in both double mutants when compared to the single corresponding QacC-Cys derivatives. Analysis of expression, however, shows that all QacC-Lys49 mutants were expressed at levels on par with WT QacC (Figure 5.15).

Fluorimetric transport assays (Section 2.11.1) of the above QacC mutants showed that efflux was reduced in both of the double QacC Lys49 mutants, as well as for Lys49Cys and Lys49Ala (Figure 5.16). As can be seen in Figure 5.16, a large drop in the capacity to export ethidium was observed for the Glu25Cys_Lys49Ala mutant. Given that the substitution of Glu25 to Cys resulted in a mutation which appeared to increase ethidium transport levels beyond the full capacity of WT QacC (140% of WT QacC), the addition of the Lys49Ala mutation resulted in a 60% decrease in transport capacity (80% of WT QacC), possibly implying that the second mutation impacted the ability of this protein to mediate ethidium efflux. This was also observed in the results obtained for the Lys49Ala_Gly105Cys mutant, as the ability of this protein to efflux ethidium out of the cell was also significantly impaired.

To further evaluate the functional impact that the incorporation of mutations at position 49 had, MIC analyses (Section 2.10) were carried out. Initial attempts using the microtiter plate method (Section 2.10.1) indicated that levels of resistance as conferred by each mutant were significantly lower compared to WT QacC (Figure 5.17). Discrepancies between MIC and ethidium efflux assay results were also observed, with the QacC Lys49Arg mutant showing comparable transport to WT QacC (95% of WT QacC), but significantly lower MIC results for ethidium (2 fold resistance compared to 8 fold for WT QacC). Such a discrepancy was shown to be even greater for the QacC Glu25Cys_Lys49Ala double mutant, which, despite facilitating moderate rates of ethidium efflux (80% of WT QacC), conferred no resistance to the expressing *E. coli* TOP10 cells.

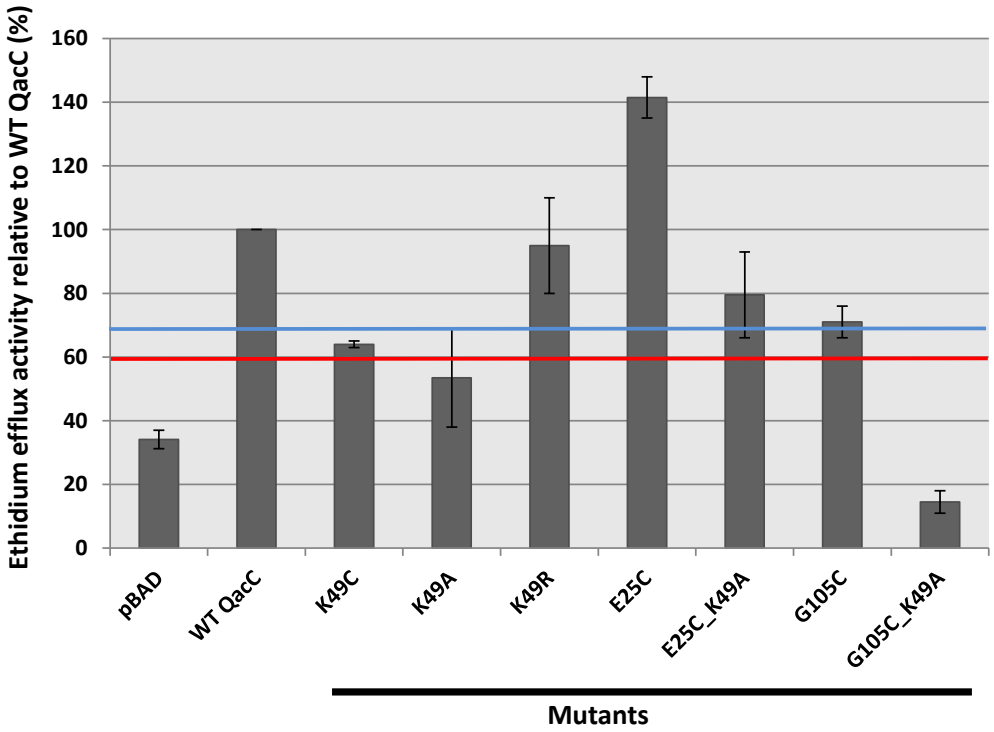

Figure 5.16: Transport of ethidium mediated by QacC Lys49 mutants

Fluorimetric assays were carried out on *E. coli* TOP10 cells carrying pBADQacC, pBADQacC_C43T and Lys mutant plasmids (Table 2.3). Cultures were grown to OD_{600}=0.55 in LBAmp and induced for two hours with 0.02 % L-arabinose (Section 2.7.1). Cells were collected and loaded with 15 µM ethidium bromide (Section 2.11.1) and ethidium efflux energised by the addition of 160 mM sodium formate at time zero (Section 2.11.1). Reactions were observed over 100 seconds and efflux calculated as a percentage of WT QacC, which was indicative of 100% efflux. Mutants showing significantly impaired capacity for efflux (0-60% WT QacC) are shown below the red bar, whilst those with impaired capacity (61-70% WT QacC) are shown below the blue bar. Results above represent the average efflux of ethidium from at least three repeat experiments for each mutant, as labelled, with error bars representing the standard error of the mean for each mutant assessed.

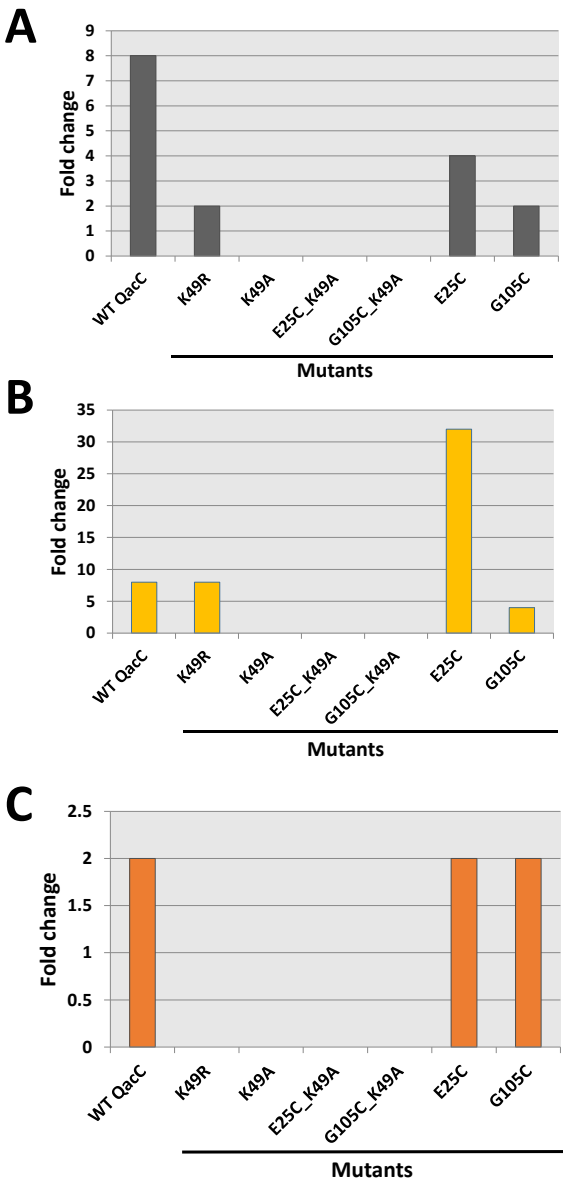

Figure 5.17: Fold change in resistance for *E. coli* TOP10 cells expressing QacC and QacC mutant proteins against ethidium, pyronin Y and acriflavine

The relative fold change of MICs of *E. coli* TOP10 QacC and QacC mutant expressing cells compared to cells carrying the empty pBAD vector. Cells were induced with 0.02% L–arabinose, diluted to OD_{600}=0.06 and added to 96 well microtiter plates containing serial dilutions of (A) ethidium, (B) pyronin Y and (C) acriflavine (Section 2.10.1). Results were obtained following overnight growth at 37°C and are based on tests carried out in triplicate.

5.6 Discussion

Cys-scanning mutagenesis has been used in numerous studies to analyse both the structure and function of multidrug transporters such as EmrE (Mordoch et al., 1999), QacA (Xu et al., 2006), TetA (Hassan et al., 2006b) and AcrB (Husain and Nikaido, 2010; Murakami et al., 2004), to name a few. In this study, a Cys-less QacC derivative was used as a template for the sequential replacement of each residue with Cys, resulting in 103 individual QacC-Cys mutants. Accessibility of each Cys residue was then examined using the maleimide derivative, fluorescein-5-maleimide, that selectively labels solvent-exposed Cys residues. Reactive residues were putatively assigned to exposed loops or to the substrate binding and/or translocation pathway, while unbound Cys residues were assigned to the membrane region. The relative position of each residue was then mapped as either forming the TMS, when shown to be non-accessible, or being part of a connecting loop when accessible. The results obtained in this study were then compared with the predicted secondary structural models of QacC as generated in Section 3.2. In addition to this, the secondary structure of QacC as predicted by NMR analysis (Poget et al., 2010) was included in order to generate a consensus 2D model of QacC (Figure 5.18). As can be seen from the bars given beneath each graph, each denoting a possible TMS, results obtained from the binding studies carried out in this study, are in good agreement with structural predictions carried out in Chapter 3. As can be seen from Figure 5.18, although small variations pertaining to the allocation of each residue to a TMS or loop are evident, especially in areas which define the boundary of each TMS, agreement between models can be seen in all four TMS. Although showing similarity for TMS1 and TMS2, NMR results differ considerably for residue allocation to TMS3 and TMS4. However, TMS4 was said to be under-predicted by the secondary-shifts, as residues 84-94 could not be assigned by the methodology employed in the NMR study (Poget et al., 2010).

Although assigned to TMSs, a number of residues were shown to have significant levels of reactivity (++) with fluorescein-5-maleimide; these were located in TMS1, 2 and 4. It is postulated that these reactive residues may be located in the

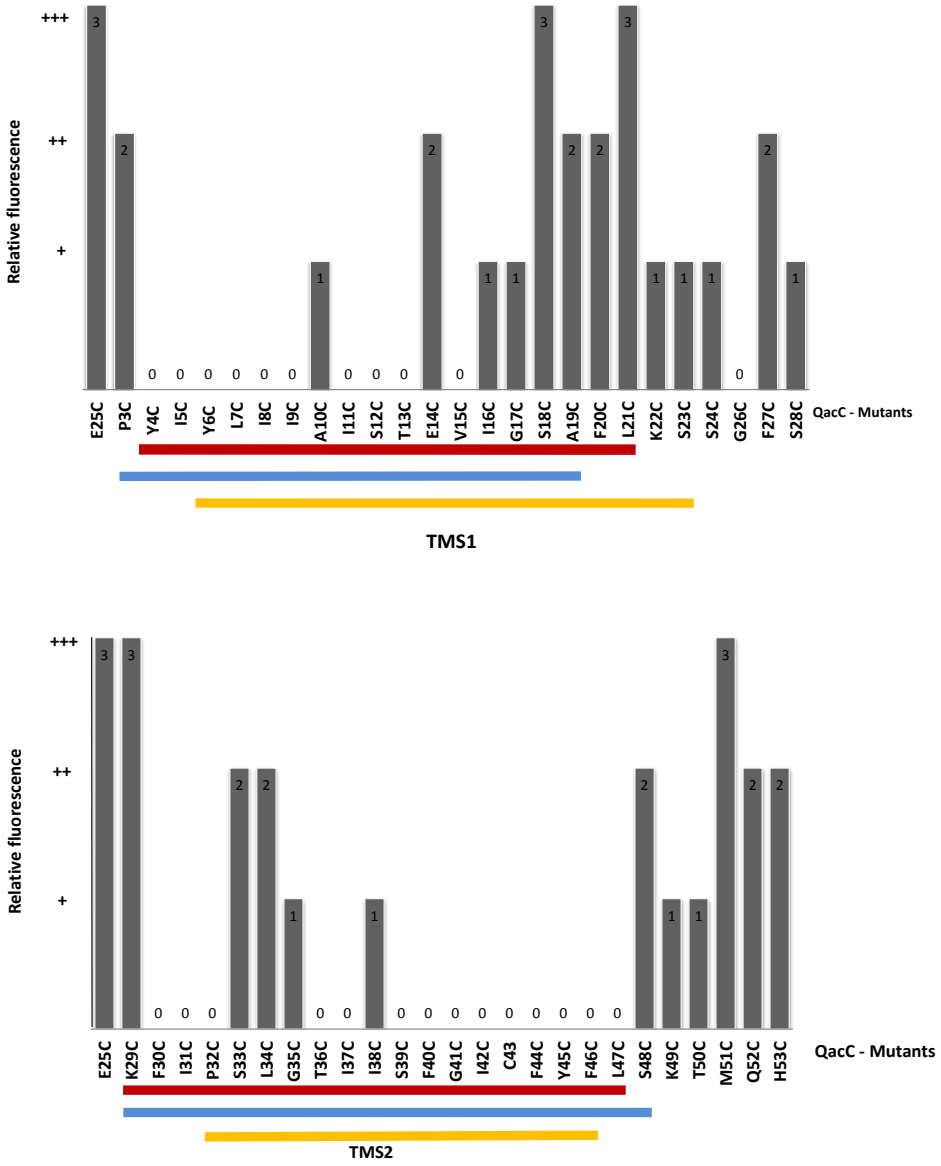

Figure 5.18a: Summary of results of fluorescein-5-maleimide binding to Cys replacements in QacC (TMS1 to TMS2)

Labelling was performed as essentially described in Section 2.14.1 and scored against Glu25Cys, which represented 100% labelling (explained in legend to Figure 5.6). Each bar depics one TMS and parts of the loops surrounding it, with bars added below representing the boundries of each segment as determined in this study (red), by TMHMM prediction (blue) and from results obtained in NMR studies (Poget et., al. 2010) (yellow).

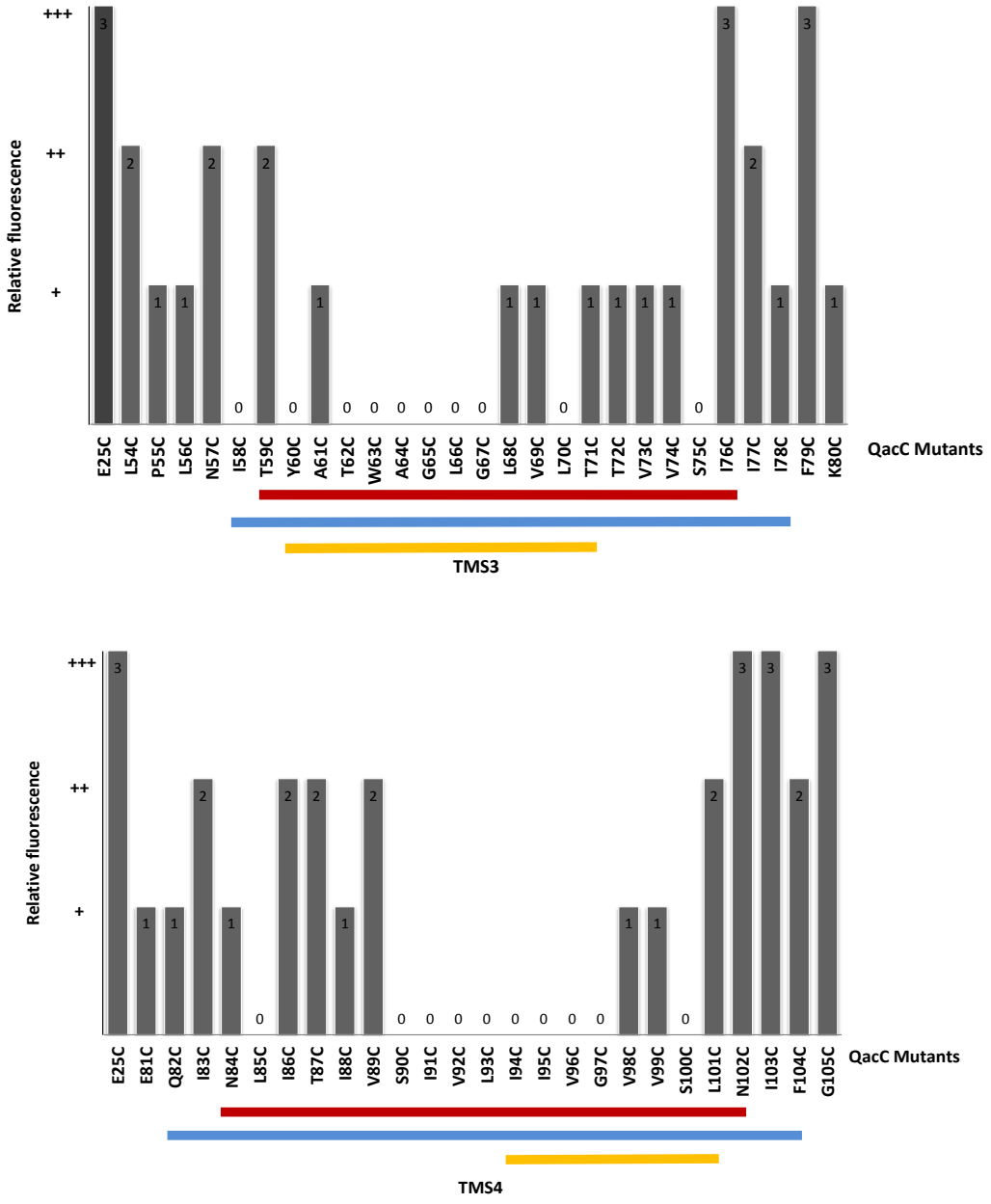

Figure 5.18b: Summary of results of fluorescein-5-maleimide binding to Cys replacements in QacC (TMS3 to TMS4)

Labelling was performed as essentially described in Section 2.14.1 and scored against Glu25Cys, which represented 100% labelling (explained in legend to Figure 5.6). Each bar depicts one TMS and parts of the loops surrounding it, with bars added below representing the boundries of each segment as determined in this study (red), by TMHMM prediction (blue) and from results obtained in NMR studies (Poget et., al. 2010) (yellow).

vicinity of the substrate-binding site and/or translocation pathway. One such mutant, Glu14Cys, supports the findings that Glu14, which has been shown to be highly conserved within the SMR family, plays a role in substrate binding and/or translocation in other SMR proteins like EmrE (Gutman et al., 2003). Substrate-mediated protection from thiol labelling is also an efficient means of determining the location of substrate binding sites. Although Glu14Cys has been shown to be reactive with fluorescein-5-maleimide (Figure 5.18), pre-binding with pyronin Y (Section 5.4), a substrate of QacC, reduced its reactivity levels (Figure 5.9), suggesting that this residue does indeed play a role in substrate binding. In addition to pre-binding with pyronin Y, attempts to utilise ethidium as the substrate of choice were also investigated (Section 5.4), however, although shown to be a good QacC substrate (Section 6.3), ethidium failed to reduce reactivity levels to those observed for Glu14Cys when pre-incubated with pyronin Y. It is uncertain as to why the addition of pyronin Y differed from the results obtained with ethidium, however, contributing factors may pertain to the size and structure of pyronin Y as compared to ethidium. With a molecular mass of 302.80 g/mol pyronin Y is less bulky than ethidium, which has a molecular mass of 394.29 g/mol and an additional phenyl ring. It is therefore possible to postulate that a change to the putative substrate binding/ translocation pathway, which Glu14 has shown to be a part of (Section 1.4.3.2), brought on by the Cys mutation (Glu14Cys), may have resulted in changes which are more permissible for pyronin Y access than for ethidium.

In addition to Glu14Cys, assessment of other reactive QacC residues within TMS1, which are thought to cluster around Glu14 were also analysed by pre-binding with pyronin Y. Faint quenching of the reactivity was detected for the Gly17Cys and Ala19Cys mutants, however, variation to quenching was observed between trials, making it difficult to make a final conclusion pertaining to the involvement of these residues in substrate binding. The role of these residues and others which have been shown to be part of a conserved motif within TMS1 of SMR proteins (Grinius et al., 1992), was further explored in Chapter 6.

Studies investigating the oligomerisation of SMR proteins such as EmrE, have revealed that the orientation of the monomers within the dimeric structure, which

is thought to be the functional unit of EmrE, take on an antiparallel orientation, and as such have been termed “dual-topology” MPs (Rapp et al., 2006; Ubarretxena-Belandia et al., 2003) (Chapter 1). To evaluate the orientation of the QacC monomers, fluorescein-5-maleimide reactivity of selected QacC derivatives with single Cys substitutions were determined in whole cells (Section 5.3.1). Of the four QacC-Cys mutants analysed, which included residues putatively assigned to loops 2, 3, 4 and 5 (Figure 5.7), all except Ile76Cys were shown to be reactive with the membrane-impermeable sulphydryl-reagent, supporting an antiparallel topology in the membrane. However, this reactivity was lower than when determined for the same derivatives in disrupted membranes. Nonetheless, this may be explained by the fact that in whole cells the final amount of fluorescein-5-maleimide available for labelling may be lower due to the presence of the outer membrane, which could limit its access into the periplasmic space.

The orientation of QacC monomers was further tested using eosin-5-maleimide (Section 5.3.2), which, like fluorescein-5-maleimide, is membrane impermeable but is a bulkier sulphydryl-reagent, and as such was used as a more stringent indicator of periplasmically exposed residues. Following the procedure outlined for whole cell labelling with fluorescein-5-maleimide (Section 2.14.2), eosin-5-maleimide was used to assess reactivity in the four loops representing QacC-Cys mutants listed above (Figure 5.8). Although less fluorescent, results support the findings obtained with fluorescein-5-maleimide, indicating the QacC monomers can have opposite orientations, with inverted topologies. In the case of MPs, evidence has been provided that loops rich in positively-charged residues tend to orient towards the cytoplasm (Heijne, 1986; von Heijne, 1992) (Section 1.5.1). Nonetheless it has been proposed that dual topology proteins have a weak positive-inside bias, showing only a small difference in the number of positively-charged residues between two sides (Rapp et al., 2007). The four positively-charged residues within QacC, all Lys, were evaluated in this study and their role in structure and function assessed (Section 5.5). Fluorescein-5-maleimide labelling putatively identified these residues to reside in loop regions within QacC, with a three Lys residues on one side of the membrane and one on the other (Figure 5.13). Evaluation of single QacC Lys to Cys

substitutions revealed that protein expression and the ability of *E. coli* TOP10 cells expressing these mutants to efflux ethidium out of the cell was not significantly reduced. In contrast, double Lys to Cys mutants, although expressed at levels comparable to WT QacC, displayed a significantly reduced capacity for ethidium efflux (Figure 5.14).

As fluorescein-5-maleimide labelling results identified Lys49 to be the only positively charged residue to lie on the periplasmic side of the membrane, in loop 3, the charge of the residue at this position was assessed by the substitution of Lys to Arg. Shown to be expressed at WT QacC levels (Figure 5.15), *E. coli* TOP10 cells expressing the QacC Lys49Arg mutant were also able to effectively extrude ethidium (Figure 5.16), verifying that this replacement did not affect expression nor function. Finally, further evaluation of the Lys49 residue, employed two QacC double mutants, Glu25Cys_Lys49Ala and Gly105Cys_Lys49Ala, in fluorescein-5-maleimide labelling, revealing that the removal of this positively-charged residue impacted both structure and function. As seen in Figure 5.15, although both highly reactive, when paired with a Lys49Ala substitution, binding levels of both Glu25Cys and Gly105Cys were significantly reduced, indicating a possible alteration to QacC topology. Possible structural modification(s) caused by these double mutations may also explain impaired function, as shown in both fluorimetric assays and MICs of the *E. coli* TOP10 cells expressing these mutants (Figures 5.16 and 5.18), further illustrating the importance of this positively-charged residue within QacC. However, to extend this study, and to further evaluate the possible role of a positively-charged residue at position 49, double mutants such as Glu25Cys_Lys49Arg and Gly105Cys_Lys49Arg would need to be generated and assessed for structure and function following the same conditions as those carried out for the above stated Cys and Ala double mutants. Without these it is difficult to ascertain whether the results obtained were due to a double mutation or due to the removal of the positively-charged residue.

5.7 Conclusion

This study evaluated the secondary structure of QacC using results obtained from a large Cys-scanning mutagenesis study and fluorescein-5-maleimide labelling. These were then compared to the 2D model of QacC as predicted by computer modelling carried out in Chapter 3 and with results obtained by NMR analysis of QacC (Poget, Harris et al. 2010) (Figure 5.18). Compiling all of the results, the QacC secondary structure is predicted to consist of four α-helices with residues Tyr4 to Leu21 making up TMS1, Phe30 to Leu47 TMS2, Ile58 to Ser75 TMS3 and lastly Asn84 to Asn102 TMS4 (Figure 5.19). Cys-scanning mutagenesis identified residues within TMS1 which may take part in substrate binding. Analysis of Glu14 in this study, shown to be strongly conserved in members of the SMR family, supported studies which have shown this residue to be at the centre of substrate binding. Whilst investigations into the positively-charged residues revealed that although individual substitutions of Lys to Cys were permissible, double Lys mutations completely disrupted QacC function. Further analysis into Lys49, being the only positively-charged residue postulated to lie on one side of the protein, revealed that this positive charge was vital for optimal QacC function (Figure 5.16 and 5.17). Finally, whole cell labelling with fluorescein-5-maleimide and eosin-5-maleimide on loop indicating QacC Cys–mutants revealed that QacC monomers insert into the membrane in two orientations, indicating that like EmrE, in a dimer QacC monomer take on an antiparallel orientation within the membrane.

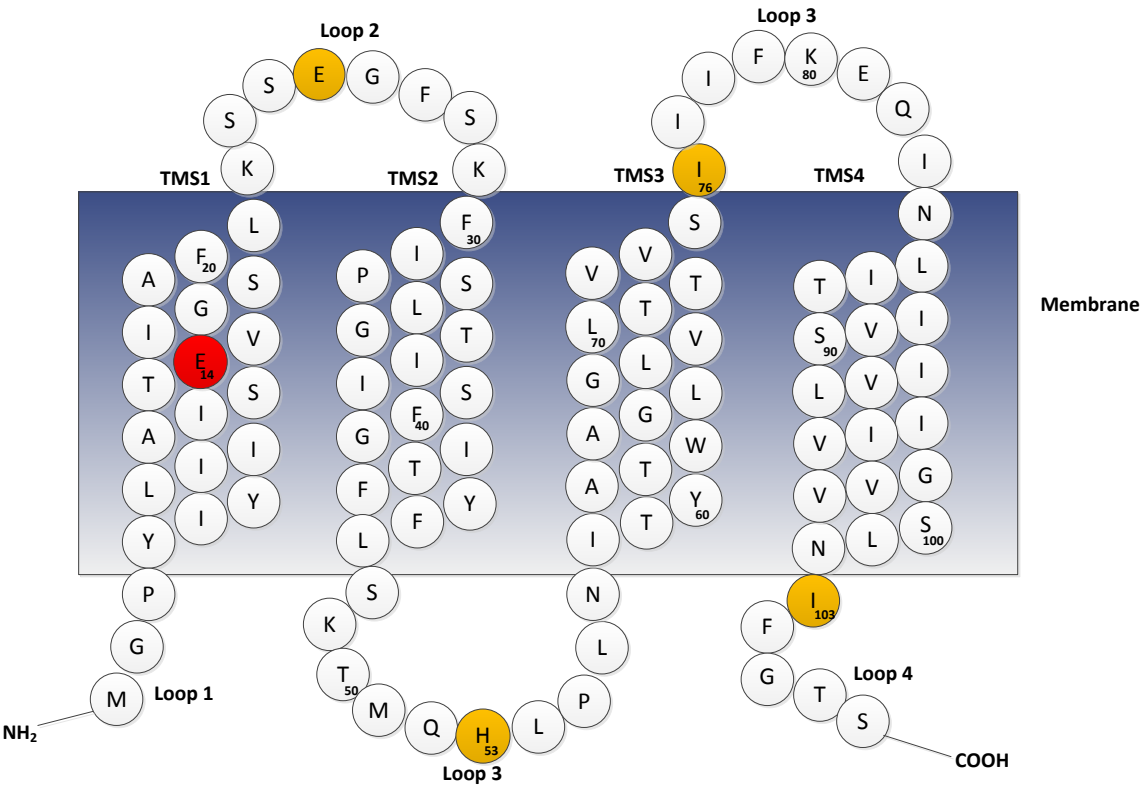

Figure 5.19: Schematic structure of QacC

Topology model for QacC depicting the possible localisation of each amino acid within the protein, as derived from results obtained from Cys-scanning mutagenesis (Section 5.2) and fluoresceine-5-malemide labelling (Section 5.3.1). The highly conserved negative residue, Glu14 (Section 1.4.3.2), putatively located within TMS1 is shaded in red, whilst residues Glu25Cys, His53Cys, Ile76Cys and Ile103Cys, assessed by whole cell labelling (Section 5.3.1), are shaded in orange.

CHAPTER 6 – ANALYSIS OF THE FUNCTION OF QACC

6.1 Introduction

Cys-scanning mutagenesis, in addition to being utilised for structural analysis, creates a set of amino acid replacement mutants that can be used to effectively evaluate and assess functional mechanisms of membrane transporters, demonstrated in studies on *E. coli* LacY (Kaback et al., 2011), AcrB (Murakami et al., 2004) and MdfA (Adler and Bibi, 2004, 2005), to name a few. Using the QacC-Cys mutants generated in the pBAD and pBluescript SK II vectors, as described in Chapter 4, this chapter assesses the functional role of individual QacC residues in resistance to cationic compounds. Fluorimetric ethidium transport assays were carried out on 103 *E. coli* TOP10 cells harbouring the 103 pBAD-based QacC-Cys mutants. Transport assays, in combination with maleimide modification were also performed on TMS1 Cys-mutants thought to play a role in substrate binding and/or translocation, as identified in Section 5.4.

In addition to active transport, the resistance profile of pBAD and pBluescript SK II-based QacC-Cys mutants were determined in either *E. coli* TOP10 or *E. coli* BL21 (DE3) cells against three previously identified QacC substrates; ethidium, acriflavine and pyronin Y. Finally, fluorimetric analysis was utilised to determine the kinetic parameters K_m and V_{max} of QacC for transport of ethidium and pyronin Y, whilst fluorescence polarisation-based binding assays were undertaken using rhodamine 6G as a substrate, and the dissociation constant (K_D) determined.

6.2 Transport assays

6.2.1 Ethidium efflux mediated by QacC mutants

To analyse the transport capabilities of QacC mutant proteins, the capacity of cells expressing each QacC derivative to efflux ethidium out of the cell was examined using a fluorescence-based transport assay (Section 2.11.1). The method was essentially based on that established previously for the evaluation of the QacA multidrug transporter (Xu et al., 2006). Fluorimetric transport assays were

conducted in *E. coli* TOP10 cells following the induction of protein expression with 0.02% L-arabinose for two hours prior to the addition of CCCP and subsequent loading with ethidium (Section 2.11.1). In order to optimise the assay the following conditions were evaluated; temperature, concentration of ethidium to be loaded and the energising compound to be used to initiate transport. Of the five different temperatures assessed for efflux (27, 32, 37, 42 and 47°C) assays conducted at 37°C to 42°C were shown to be optimal for efficient ethidium efflux, whilst the optimal concentration of ethidium for cell loading, initially trialled at 15 and 30 µM, was determined to be 15 µM. Finally the most efficient energiser added to initiate transport was shown to be sodium formate added a concentration of 160 mM (Table 6.1).

Initial attempts at measuring ethidium efflux in *E. coli* TOP10 cells revealed that levels of ethidium diffusion from cells carrying an empty pBAD vector, when compared to WT QacC and measured over 300 seconds, were relatively high (50% of WT QacC). Such diffusion was not dependant on the addition of sodium formate, as it occurred steadily from the moment fluorescence was read (data not shown). To minimise this, the duration that transport was measured was also optimised, and as such transport of ethidium was measured over 100 seconds only, allowing for the differentiation of passive diffusion versus active efflux. This differentiation was evident by the sudden drop in fluorescence following the initiation of transport seen in cells expressing QacC, as compared to a steady decrease in fluorescence over time as seen in cells carrying the pBAD vector only.

Ethidium transport capacities of each QacC-Cys mutant were compared to WT QacC, which was measured over 100 seconds and represented 100% efflux. Of the 103 pBAD-based QacC-Cys mutants analysed (Figure 6.1) 15 mutants were shown to have a significantly reduced capacity for ethidium efflux (0-60% of WT), whilst a further seven were shown to have impaired efflux (61–70% of WT). Of these, five QacC-Cys mutants, Tyr4Cys, Leu7Cys, Pro32Cys, Leu47Cys and Is58Cys, were also shown to have lower than WT QacC protein expression levels (Figure 5.4, Section 5.2.3), which may have had an impact on their ability to extrude ethidium.

Table 6.1: Assessment of conditions used for ethidium transport assays in *E. coli* TOP10 cells expressing QacC derivatives

Condition tested	Scope of conditions tested	Optimal condition
Temperature (°C)	27	Evaluation of temperature revealed optimal ethidium efflux by cells incubated at temperatures ranging from 37°C and 42°C
	32	
	37	
	42	
	47	
Concentration of ethidium loading (µM)	15	Optimal concentration of ethidium to be loaded was found to be 15 µM
	30	
Sodium formate (mM)	160	Initiation of ethidium efflux was shown to be most efficient following the addition of 160 mM sodium formate
Glucose (mM)	10	
	25	

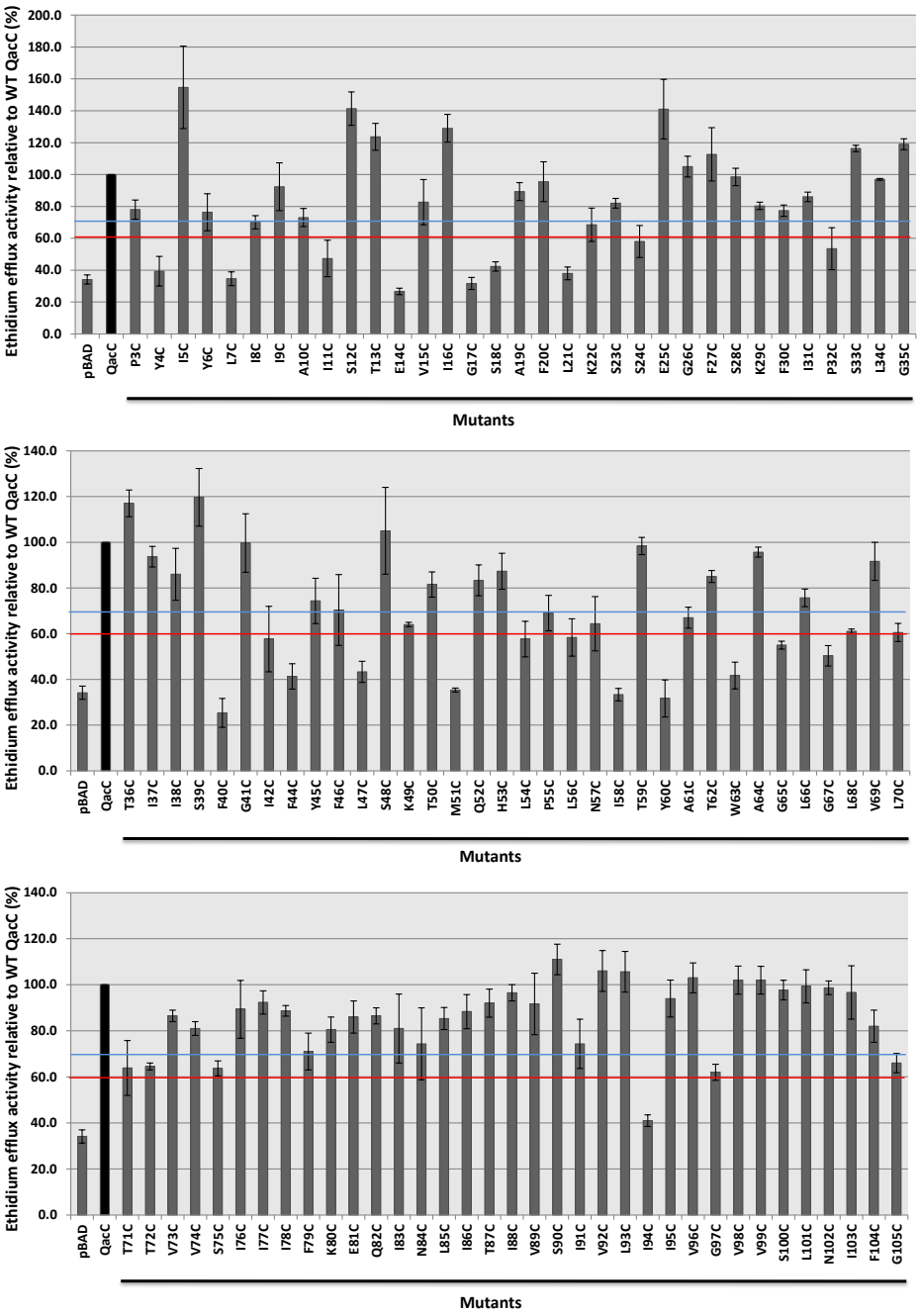

Figure 6.1: Summary of results for the fluorimetric assays of ethidium export in *E. coli* TOP10 cells

Efflux of ethidium (15 µM) from cells harbouring QacC and QacC-mutant plasmids (Table 2.3). Transport assays were carried out as described in Section 2.11.1. Efflux of ethidium was calculated compared to WT QacC (black bar), which was indicative of 100% export. Mutants showing significantly impaired capacity for efflux (0-60% WT QacC) are shown below the red bar, whilst those with impaired capacity (61-70% WT QacC) are shown below the blue bar. Results obtained from cells carrying the empty pBAD vector are included as a negative control. Error bars are standard deviations of three biological replicates.

The remaining 81 QacC-Cys mutants were found to export ethidium at levels more or less comparable to WT QacC.

6.2.2 Assessing if maleimide blocks ethidium efflux

NEM, a membrane-permeant thiol-reactive reagent that covalently bonds to sulphydryl groups of Cys residues (Chouchani et al., 2011; Lundbald, 2015), was used to further assess residues within TMS1 which may play a role in the transport pathway. Residues were chosen based on their reactivity with fluoresceine-5-maleimide (Section 5.4), putative position in TMS1 and proximity to Glu14, which as discussed in Section 1.4.3.2, has been determined to be part of the substrate and proton binding domain in SMR proteins (Gutman et al., 2003; Yerushalmi and Schuldiner, 2000b). QacC mutants Glu14Cys, Gly17Cys, Ser18Cys and Lys21Cys, although shown to also be reactive with fluoresceine-5-maleimdie, were not evaluated in this study as functional assays revealed that these mutations lead to significant impairment in the ability to transport ethidium (Section 6.2.1). Using fluorimetric assays in combination with NEM, as has been carried out in previous studies on QacC (Paulsen et al., 1995) and QacA (Hassan et al., 2008), the effect of NEM on *qacC*-encoded ethidium efflux was assessed (Section 2.11.1).

As can be seen in Figure 6.2, the addition of 5 mM NEM did not inhibit ethidium efflux in *E. coli* TOP10 cells harbouring WT QacC. This was expected as WT QacC does not possess a Cys residue, and as such the addition of NEM should not affect transport. Of the four mutants assessed, Ala10Cys and Ala19Cys were shown to have the rate of ethidium efflux modulated by the pre-incubation with NEM. The rate of ethidium transport mediated by the Ala10Cys QacC mutant decreased slightly upon the addition of NEM, whilst the opposite was observed for the QacC-Cys mutant Ala19Cys, which exhibited an increased rate of efflux following pre-incubation. The remaining two QacC-Cys mutants, Ile16Cys and Phe20Cys, did not display significant alterations in transport.

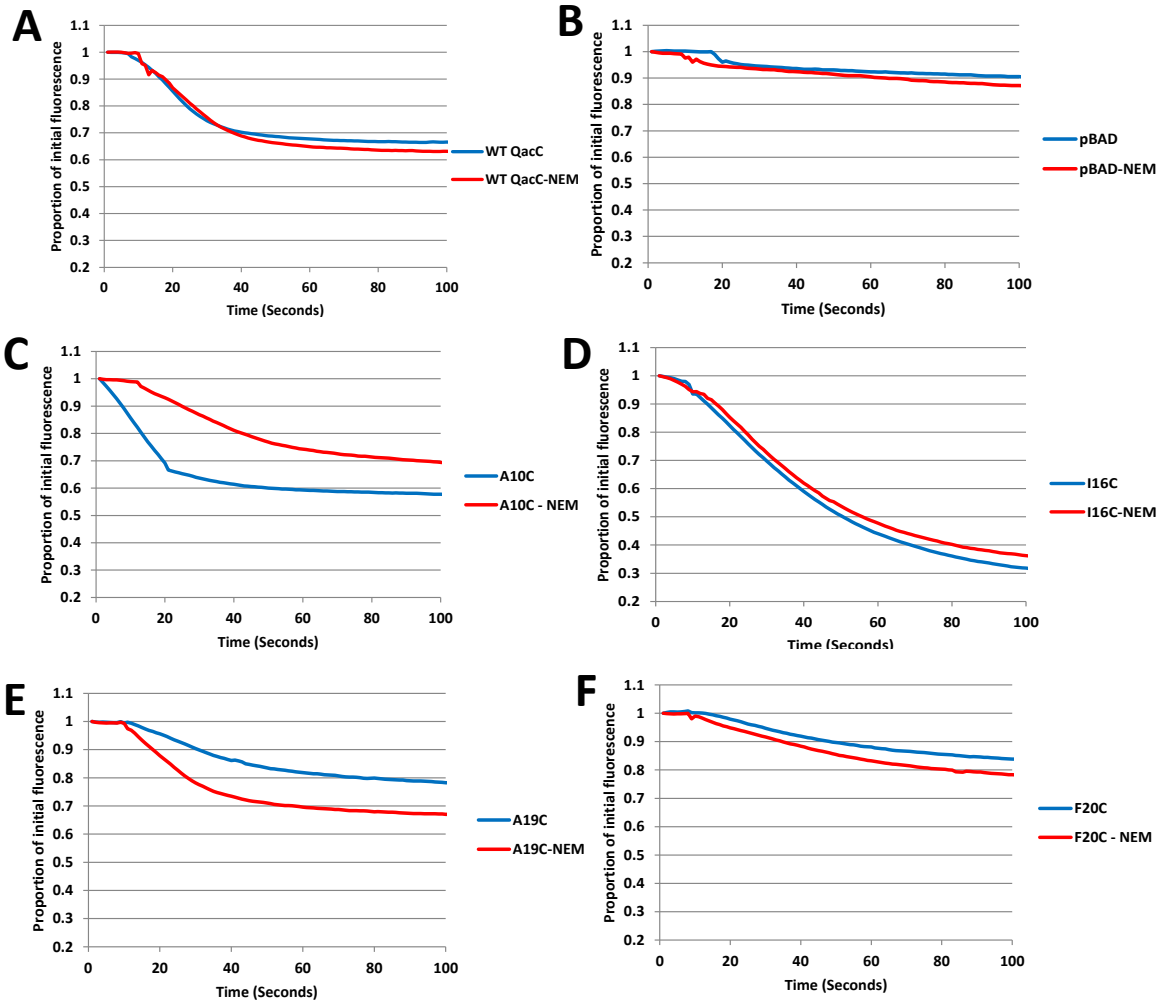

Figure 6.2: Ethidium transport mediated by QacC TMS1 derivatives following NEM binding

E. coli TOP10 cells expressing pBAD-based QacC mutants were prepared essentially as outlined in Section 2.11.1. To determine the effect of NEM binding on the capacity of cells to efflux ethidium, cells were pre-incubated with 5 mM NEM prior to being loaded with 15 µM ethidium in the presence of 10 mM CCCP. Transport was energised by the addition of sodium formate at time zero, at a final concentration of 160 mM. Control cells loaded with ethidium only were also included. Each graph (A to F) illustrates results of cells pre-treated with NEM (red) and those loaded with ethidium only (blue). All efflux experiments were performed in at least triplicate, with a representative shown above.

6.2.3 Pyronin Y efflux mediated by QacC mutants

Pyronin Y is a common substrate of SMR-family proteins (De Rossi et al., 1998; Srinivasan et al., 2009; Yerushalmi et al., 1995), however, it had not previously been identified as a substrate of the QacC exporter. Given the MIC results obtained in Section 6.3, which showed high QacC-mediated resistance against pyronin Y, assessment of active efflux was carried out in order to further evaluate this substrate. As no previous efflux assays using this substrate have performed with QacC, the methodology used for ethidium fluorimetric transport assays were used and optimised for pyronin Y. Unlike ethidium, when intracellular pyronin Y accumulates it binds to RNA, leading to quenching of fluorescence which is indicated by a sharp rise in fluorescence following active efflux (Mitchell et al., 1999; Schumacher et al., 2007). Initial trials carried out in order to determine the final concentration of pyronin Y to be loaded into *E. coli* TOP10 cells revealed that due to the high fluorescence of this compound and the limitations of the equipment used to measure fluorescence many of the concentrations tested, including those below its MIC (Section 6.3), did not allow for efficient efflux values to be calculated. Following numerous attempts, it was discovered that loading *E. coli* TOP10 cells expressing QacC, standardised at OD_{600}=0.4, with pyronin Y at a concentration of 33 nM resulted in optimal efflux, as can be seen in Figure 6.3. These results demonstrated that pyronin Y is a QacC substrate, and that these initial trials serve as a proof of concept, serving to illustrate an alternative or a possible additional method with which to functionally assess QacC.

6.3 Resistance profiling of QacC derivatives

6.3.1 Determination of the resistance profile of QacC-Cys mutants using microtiter plate assays

Assessment of resistance conferred to *E. coli* cells expressing QacC mutants was carried out using three common SMR substrates; ethidium, pyronin Y and acriflavine. Of the preliminary trials to determine the best technique to evaluate

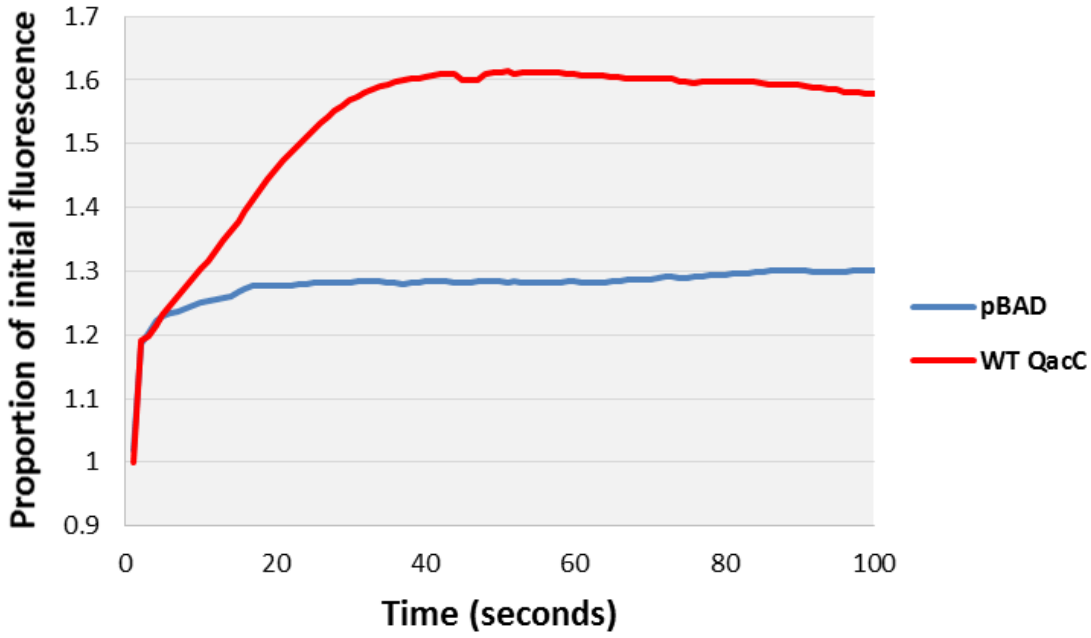

Figure 6.3: Transport of pyronin Y mediated by QacC

Fluorimetric assays were carried out on *E. coli* TOP10 cells carrying pBAD and pBADQacC_C43T (Table 2.3). Cultures were grown to OD_{600}=0.55 in LBAmp and induced for two hours with 0.02 % L-arabinose (Section 2.7.1). Cells were collected and loaded with 33 nM pyronin Y (Section 2.12) and pyronin Y efflux energised by the addition of 160 mM sodium formate at time zero (Section 2.12). Colours represent: blue - *E. coli* TOP10 cells carrying an empty pBAD vector, red - WT QacC. Representative results of duplicate transport reactions are shown.

resistance, microtiter MIC assays using serial two-fold dilutions of each substrate (Section 2.10.1) were shown to be accurate in determining the resistance levels of WT QacC to the compounds assessed. Thus, microtiter MIC assays were chosen as the primary method to evaluate all QacC mutants. Induction of QacC expression in *E. coli* TOP10 cells with 0.02% L-arabinose for one hour was determined ideal for MIC analysis (Section 2.10.1).

To establish the substrate profile to be used for the analysis of all QacC mutants, a number of substrates were initially evaluated, of these, ethidium, pyronin Y and acriflavine were chosen for all subsequent analysis as these performed optimally throughout all trials. Assessment of all 103 pBAD-based QacC mutants (Figure 6.4) revealed that in total, 15 QacC-Cys mutants were sensitive to all of the compounds analysed. Of these, only Pro32Cys was shown to have low levels of protein expression (Figure 5.4), and significantly impaired ethidium efflux (Figure 6.1). Interestingly, three of the QacC-Cys mutants which displayed no resistance in *E. coli* TOP10 cells, Pro55Cys, Phe79Cys and Glu81Cys were shown to efflux ethidium out of the cell at levels comparable to WT QacC (Figure 6.1). In addition to these, 10 QacC-Cys mutants conferred no resistance against two substrates and reduced resistance levels to third substrate. A comparison of these results with those obtained by fluorimetric analysis (Figure 6.1), revealed that those QacC-Cys mutants showing impaired ethidium efflux also displayed low resistance levels against ethidium. Of the 17 QacC mutants showing low levels of expression (Figure 5.4, Section 5.2.3) 10 were also shown to have low QacC-mediated resistance to all compounds, making it difficult to fully assess the impact the mutation had on resistance. In addition to these, a cluster of QacC mutants, proposed to be in loop 3 and the beginning of TMS3, displayed a reduced capacity for resistance against all three substrates, with the QacC Tyr50Cys mutant being the only one which demonstrated WT like resistance in that cluster and Tyr59Cys intermediate-levels of resistance to all compounds.

A comparison of levels of resistance to the three different substrates revealed that QacC-mediated resistance against acriflavine is low, reaching 31.25 μg/ml, as compared to 15.6 μg/ml obtained for *E. coli* TOP10 cells carrying an empty

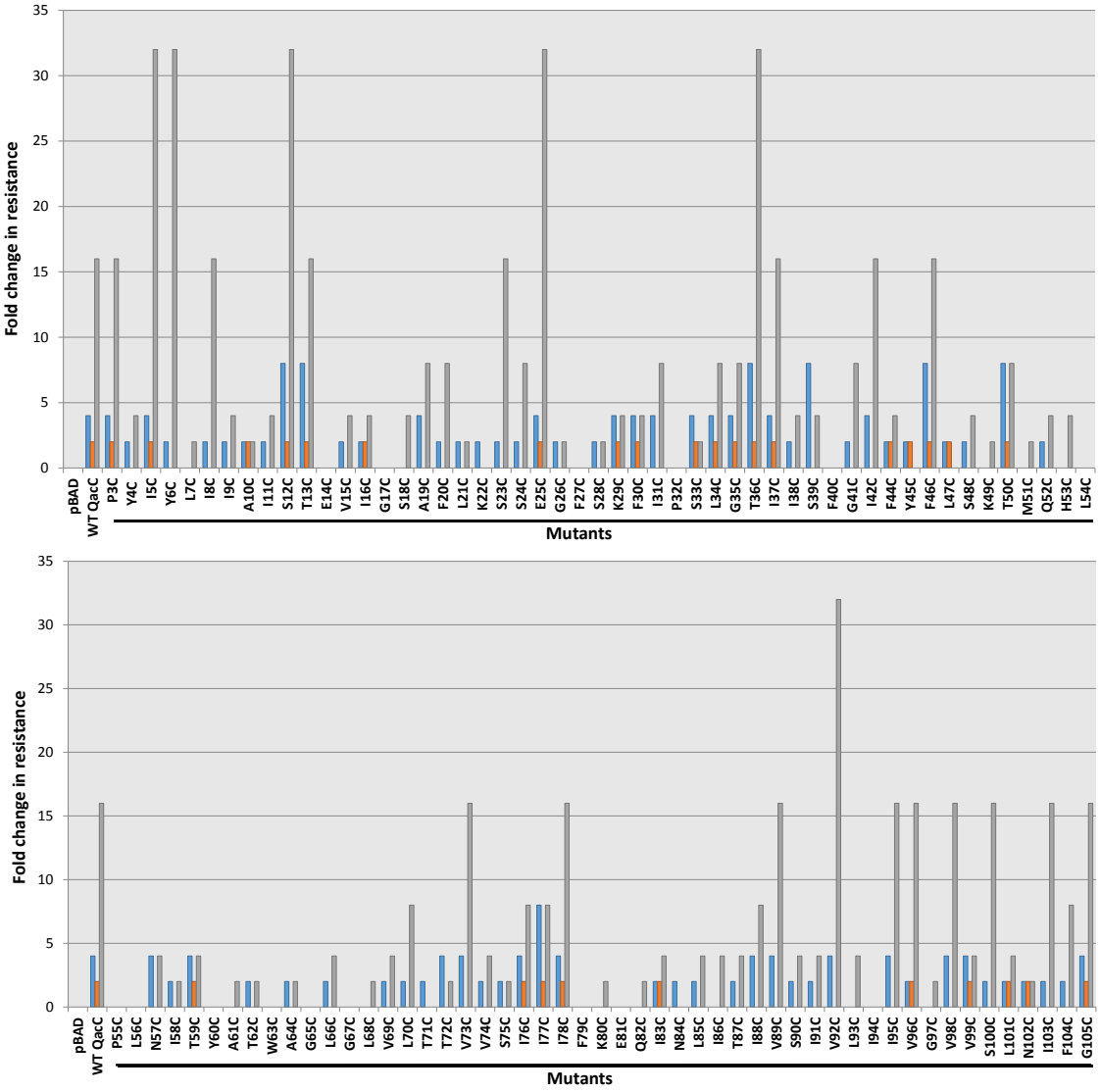

Figure 6.4: MIC_{90} results shown as fold change in resistance to ethidium, acriflavine and pyronin Y for *E. coli* TOP10 cells expressing QacC-Cys mutants

E. coli TOP 10 cells carrying pBAD based QacC plasmids (Table 2.3) were grown and induced with 0.02% L-arabinose, as set out in Section 2.10.1. Cells were standardised to OD=0.6, diluted 1:100 and added to serially diluted compounds. Absorbances were determined after overnight incubation at 37°C. MIC concentrations for ethidium (blue bars), acriflavine (orange bars) and pyronin Y (grey bars) were compared with results obtained for cells carrying the empty pBAD vector only and are presented as fold increases. MIC analyses of cells expressing WT QacC showed a four-fold increase in resistance against ethidium (800 µg/ml), two-fold increase against acriflavine (31.25 µg/ml) and a 16-fold increase against pyronin Y (32.25 µg/ml). Data includes results for ethidium (blue bars), acriflavine (orange bars) and pyronin Y (grey bars). Experiments were carried out in triplicate with the average presented above.

pBAD vector (Figure 6.4). In contrast, both ethidium and pyronin Y were shown to be good substrates for QacC, as MIC analysis demonstrated that QacC was able to confer resistance to ethidium and pyronin Y at concentrations up to 400 µg/ml and 65 µg/ml, respectively. Compared to results obtained for *E. coli* TOP10 cells carrying the empty pBAD vector, which reached 100 µg/ml for ethidium and 4 µg/ml for pyronin Y. The differences observed in resistance levels, when compared to the negative control, allowed for more accurate MIC determination of each QacC–Cys mutant.

MIC results revealed that 11 Cys substitutions led to an increase in QacC-mediated resistance against ethidium and/or pyronin Y (Figure 6.4). Of these, six mutants, Ile5Cys, Ser12Cys, Thr13Cys, Glu25Cys, Thr36Cys and Ser39Cys also displayed significantly higher levels of ethidium efflux (~120% of QacC WT) (Figure 6.1). This could not be attributed to an increase in protein expression as an examination of protein expression (Section 4) revealed that these mutants all produced protein levels comparable to the WT QacC. Thus it reflects an effect of the substitution on the activity of these mutants.

6.3.2 Plate dilution: pBAD-based vs pBluescript II SK-based QacC multidrug resistance

To further evaluate QacC-mediated resistance, the plate dilution method that has been used to analyse resistance conferred by EmrE (Mordoch et al., 1999) and MdfA (Adler and Bibi, 2005) was also used in this study. This method differs from the microtiter plate method by evaluating the growth of serial 10-fold dilutions of *E. coli* cells expressing the desired QacC-Cys mutant, on plates containing each substrate at a single concentration only (Section 2.10.3). As such, this allows for greater discrimination between samples as determined by the growth of cells at each dilution. Given the low levels of QacC mediated acriflavine resistance (Section 6.3.1), this method was utilised to discern smaller differences in the growth of *E. coli* TOP10 cells expressing each QacC-Cys mutant, at a given concentration.

Initial trials of this method were conducted using induced pBAD-based QacC in *E. coli* TOP10 cells, on ethidium (100 and 200 µg/ml) and acriflavine (25, 50 and 75

µg/ml) plates. These concentrations were based on results obtained from Section 6.3.1. However, results (Figure 6.5) did not reflect those obtained in Section 6.3.1, as growth on ethidium and acriflavine plates at all concentrations, did not significantly discriminate between cells expressing WT QacC or those carrying the empty pBAD vector.

To further evaluate this assay, analysis was undertaken using *E. coli* BL21 (DE3) cells (Section 4.2.1) with pBSQacC_C43T (Table 2.3). Following essentially the same methodology (Section 2.10.3), induced *E. coli* BL21 (DE3) cells carrying pBSQacC_C43T were grown and standardised to OD_{600}=0.6. A series of 10-fold dilutions was then prepared and spotted onto plates containing ethidium (100 and 200 µg/ml) and acriflavine (25 and 50 µg/ml). Although the ethidium concentrations were optimal for assessment of resistance, the concentrations of acriflavine were observed as being too high, as no growth occurred on either 25 or 50 µg/ml plates. A reduction to 12.5 µg/ml, however, resulted in adequate growth of cells by which conferred resistance could be assessed (Figure 6.6). Cells expressing QacC were able to grow on media containing ethidium and acriflavine at dilutions of 10^{-1}-10^{-7} and 10^{-1}-10^{-5} for ethidium at 100 and 200 µg/ml, respectively, and dilutions of 10^{-1}-10^{-5} for acriflavine at 12.5 µg/ml. In contrast, cells carrying the empty vector could not grow at any of the dilutions tested. To expand this analysis, as crystal violet has also been shown to be QacC substrate (Section 1.6.1) it was also analysed using this method, at a concentration of 15 µg/ml.

This method was also used to assess and compare the resistance levels of the QacC/Hisx6 and QacC/1D4/Hisx6 constructs. As shown in Figure 6.7, both constructs conferred comparable resistance against ethidium, acriflavine and crystal violet. These results confirm that the addition of both the 1D4 and Hisx6 tags does not affect QacC function. In addition to these WT constructs, 15 pBluescript SK II-based QacC-Cys mutants, chosen to represent mutants across the resistance spectrum, were evaluated and their growth determined (Figure 6.8).

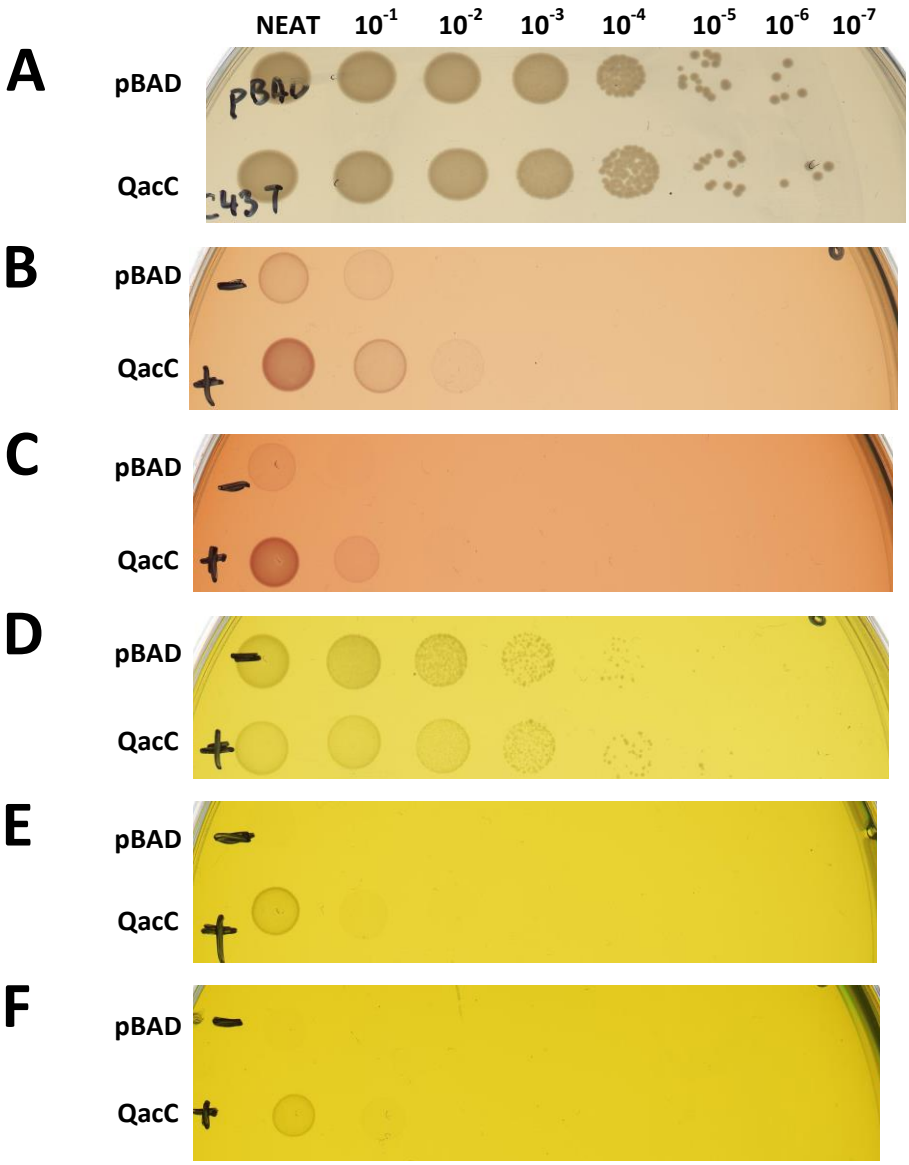

Figure 6.5: Growth phenotype of *E. coli* TOP10 cells expressing QacC and cells carrying the empty pBAD vector only

Cultures of *E. coli* TOP10 cells carrying pBAD and pBADQacC_C43T (Table 2.3) were grown to OD_{600}=0.55 in LBAmp induced with 0.02% L-arabinose and left to grow for a further hour. Cells were standardised to OD_{600}=0.6 and used in a serial 10–fold dilution (10^{-1} – 10^{-7}). (A) 5 µl of a neat sample (10^{0}) followed by each dilution was spotted on plates containing MH only (growth control), (B and C) ethidium (100 and 200 µg/ml) and (D, E and F) acriflavine (25, 50 and 75 µg/ml) and grown overnight at 37°C (Section 2.10.3). pBAD was used as a negative control for each trial. Experiments were performed in at least duplicate and a representative of the results shown above.

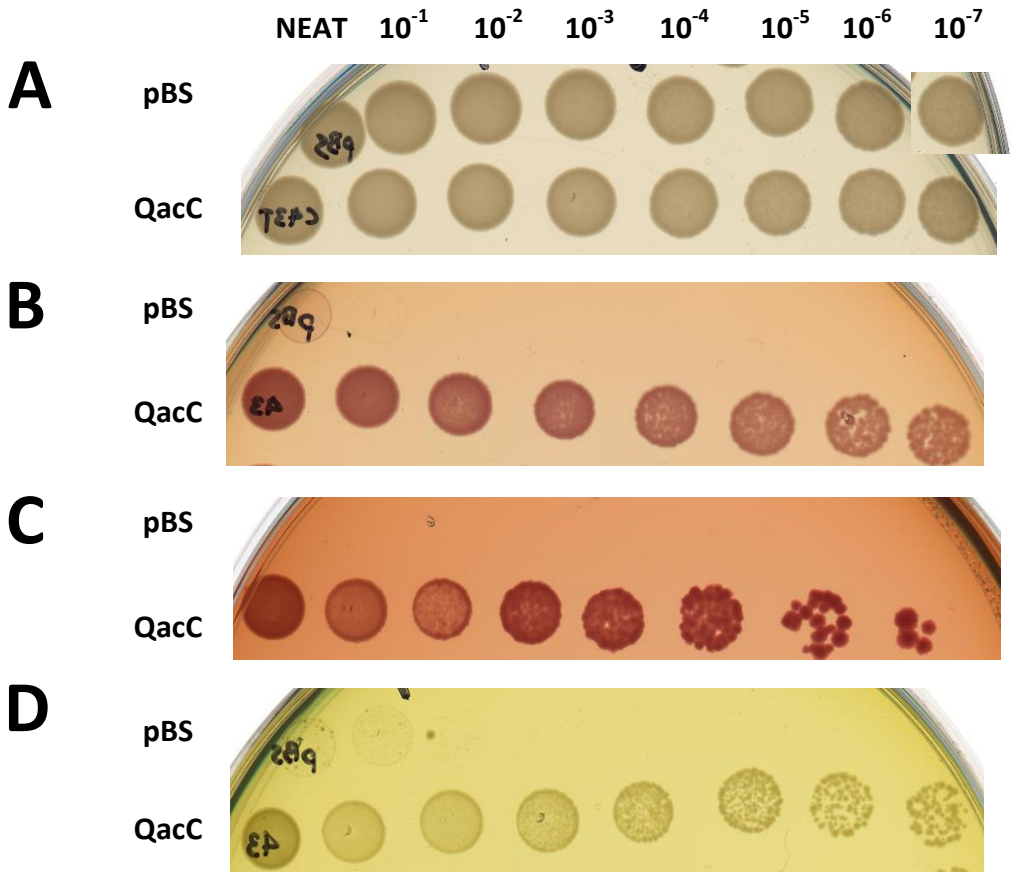

Figure 6.6: Growth of *E. coli* BL21 (DE3) cells carrying an empty pBluescript SK II vector or expressing WT QacC

Cultures of *E. coli* BL21 (DE3) carrying pBluescript SK II and pBSQacC_C43T (Table 2.3) were grown to OD_{600}=0.55 in LBAmp induced with 0.1 mM IPTG and left to grow for a further hour. Cells were standardised to OD_{600}=0.6 and used in a serial 10–fold dilution (10^{-1} – 10^{-7}). (A) 5 µl of a neat (10^{0}) sample and each dilution was spotted on MH plates containing no additive (growth control), (B and C) ethidium (100 and 200 µg/ml) and (D) acriflavine (12.5 µg/ml) and left to grow overnight at 37°C (Section 2.10.3). pBluescript SK II was used as a negative control for each trial. Experiments were performed in at least duplicate with a representative of the results shown above.

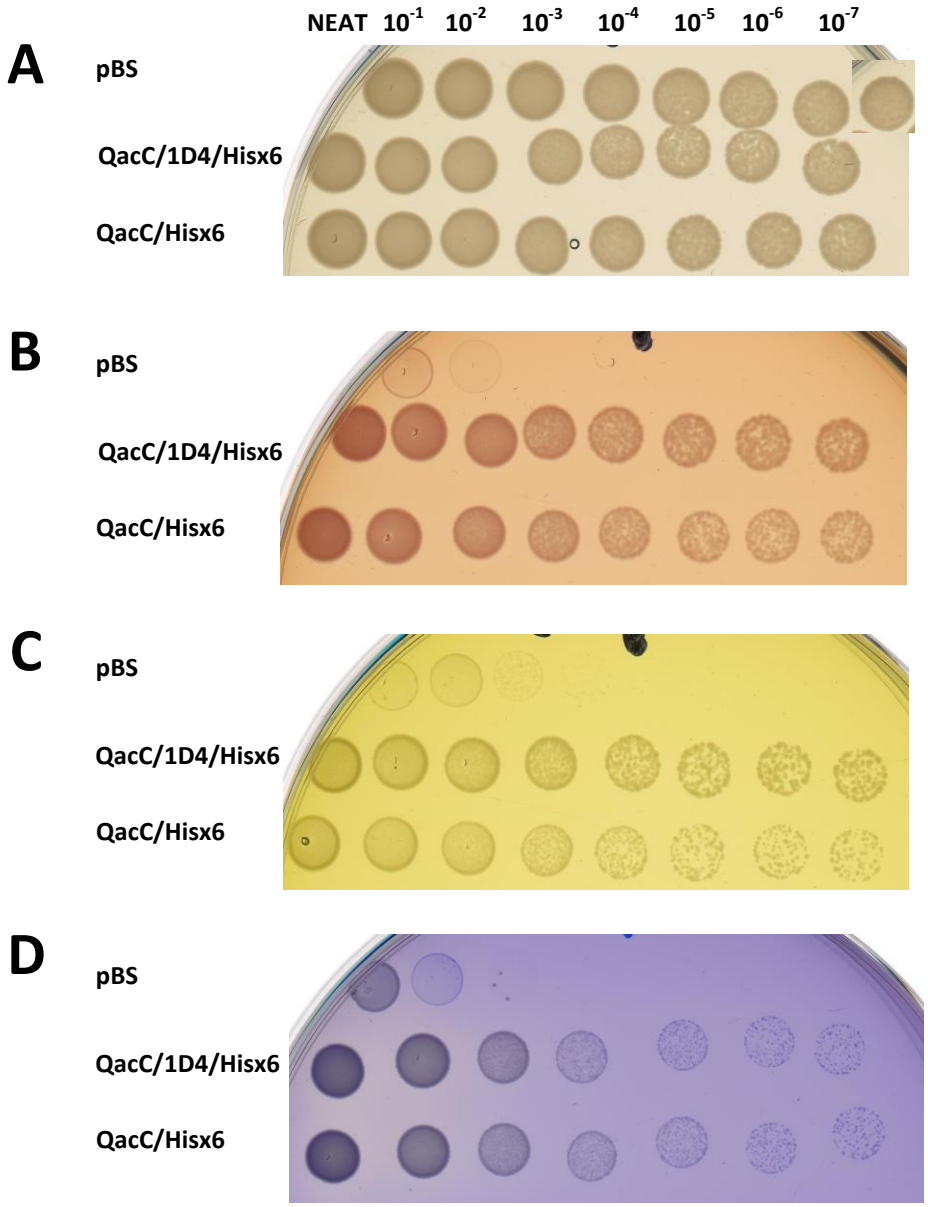

Figure 6.7: Growth phenotype of *E. coli* BL21 (DE3) cells expressing QacC/Hisx6, QacC/1D4/Hisx6 and cells carrying the empty pBluescript SK II vector only

Cultures of *E. coli* BL21 (DE3) cells carrying pBluescript SK II, pBSQacC_C43T or pBSQacC_C43THis6 (Table 2.3) were grown to OD_{600}=0.55 in LBAmp induced with 0.1 mM IPTG and left to grow for a further hour. Cells were standardised to OD_{600}=0.6 and used in a serial 10–fold dilution (10^{-1} – 10^{-7}). (A) 5 µl of a neat sample (10^{0}) followed by each dilution was spotted on plates containing MH only (growth control), (B) ethidium (100 µg/ml), (C) acriflavine (12.5 µg/ml) and (D) crystal violet (15 µg/ml) and left to grow overnight at 37°C (Section 2.10.3). pBluescript SK II was used as a negative control for each trial. Experiments were performed in at least duplicate with a representative of the results shown above.

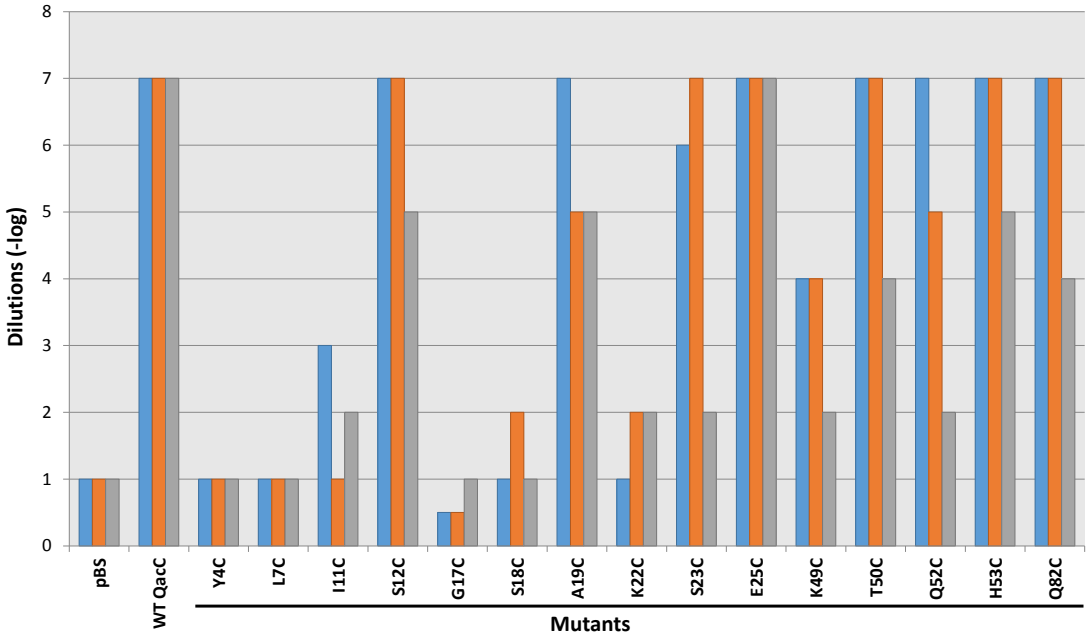

Figure 6.8: Growth of *E. coli* BL21 (DE3) cells expressing pBluescript SK II-based QacC mutants

Cells were grown in LBAmp, standardised to OD_{600}=0.55, induced with 0.1 mM IPTG and spotted on plates containing 100 µg/ml ethidium (blue bars), 12.5 µg/ml acriflavine (orange bars) and 15 µg/ml crystal violet (grey bars) as set out in Section 2.10.3. Growth was analysed following O/N incubation at 37°C. Cells carrying the empty pBluescript SK II vector were used as a negative control. Growth was observed on control plates containing no compounds for all cell dilutions. Each bar represents the maximum dilution at which growth was observed.

Results were then compared to those obtained using the microtiter method for both ethidium and acriflavine. This revealed that QacC-Cys mutants conferring high levels of resistance, such as Ser12Cys, Glu25Cys and Thr50Cys grew to the final 10^{-7} dilution, and those which displayed low levels of resistance such as Tyr4Cys, Leu7Cys, Ile11Cys, Gly17Cys, Ser18Cys and Lys22Cys grew to low dilutions only, up to 10^{-2}. Although most mutants showed similar resistance profiles using both methods there were some discrepancies. QacC mutants Ala19Cys, Ser23Cys, Lys49Cys, Gln52Cys, His53Cys and Gln82Cys, which according to the microtiter plate method were all shown to have decreased resistance to both ethidium and acriflavine, still grew to the highest dilutions (10^{-7}) on both ethidium and acriflavine plates.

6.4 Binding studies

6.4.1 Michaelis-Menten kinetics of ethidium and pyronin Y export by QacC

Utilising fluorescent transport assays as described in Section 6.2, the bioenergetics of QacC efflux, namely the kinetic parameters K_m and $V_{max,}$ were determined for the substrates ethidium and pyronin Y. Analyses were performed on whole *E. coli* TOP10 cells expressing WT QacC (Section 2.11.2), which were loaded with increasing concentrations of the substrate, 0.5 to 16 µM for ethidium and 4 to 57 nM for pyronin Y, assessed. It was determined that QacC-mediated ethidium efflux conformed to Michaelis–Menten kinetics (Figure 6.9), and that such efflux was saturable, as exemplified by the Lineweaver–Burke plot (Figure 6.10a). Thus, the kinetic parameters, K_m and V_{max} were determined and calculated to be 11.21±4.2 µM and 16.94±4.47 FU s^{-1} (fluorescence units per second), for ethidium. Although attempts to calculate the K_m and V_{max} were also undertaken for pyronin Y, due to the nature of this compound and its high fluorescence, as discussed in Section 6.2.3, the values of these could not be determined as data points obtained crossed the Y-axis zero (Figure 6.10b). Future studies examining this compound will need to be undertaken using different concentrations of pyronin Y than those examined in this experiment.

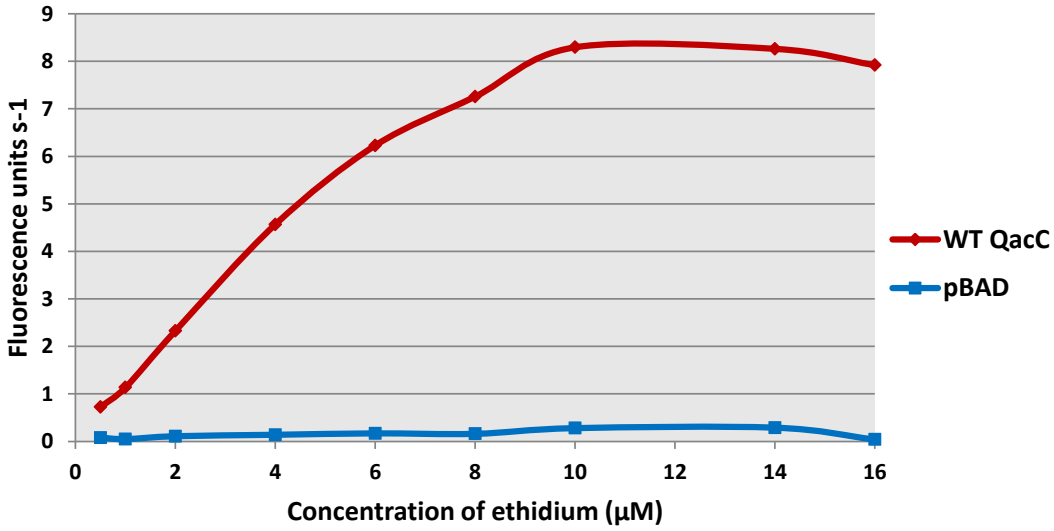

Figure 6.9: Michaelis-Menten kinetics of QacC-mediated ethidium export in *E. coli* TOP10 cells

Efflux assays were carried out utilising a range of ethidium concentrations from 0.5 to 16 µM, as described in Section 2.11.2 and discussed in Section 6.4.1. The graph depicts the initial velocity, which is represented by fluorescence units per second (s^{-1}), plotted as a function of the concentration of ethidium (µM). Cells carrying the empty pBAD vector were used as a negative control. Efflux experiments were carried out over three separate trials, with a representative shown above.

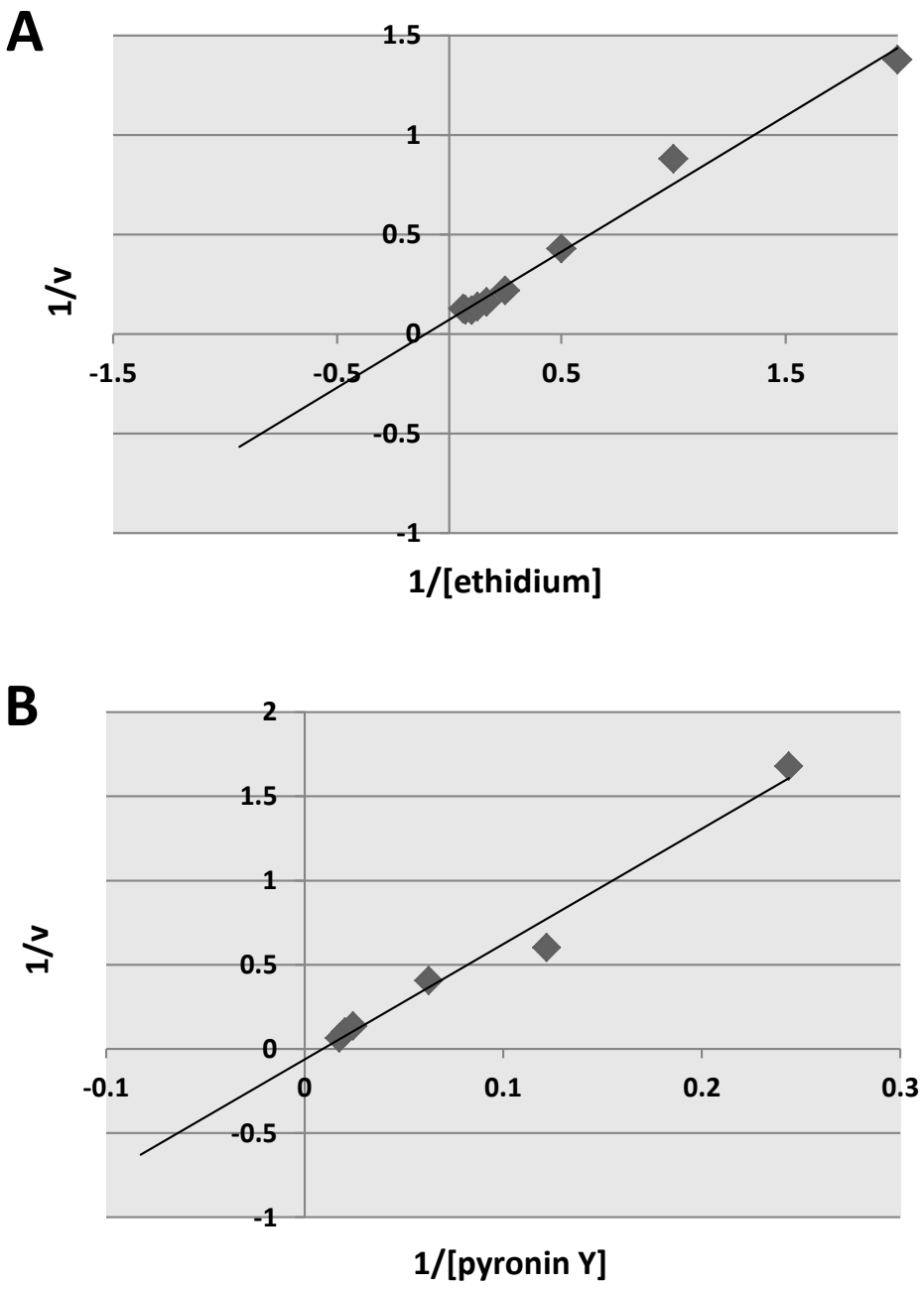

Figure 6.10: Rate of ethidium and pyronin Y export as a function of external substrate concentration plotted using Lineweaver-Burk.

The external (A) ethidium and (B) pyronin Y concentrations tested were 0.5 to 16 µM for ethidium and 4.1 to 57 nM of pyronin Y. Using these concentrations a series of efflux curves were generated and the initial velocity calculated by averaging the liner part of each curve. Plotting the inverse of the initial velocity (1/v) and inverse of substrate concentration (1/[S]) a double-reciprocal plot was generated and the kinetic parameters K_m and V_{max} determined for ethidium, however they could not be determined for pyronin Y.

6.4.2 QacC rhodamine 6G fluorescence polarisation-based binding assay

To further assess QacC substrate binding and to illustrate that QacC is functional in its tertiary conformation, a fluorescence polarisation binding assay was carried out on purified QacC, using rhodamine 6G as the substrate. This assay has successfully been applied to analyse QacA binding of rhodamine and was used in this study (Hassan et al., 2009). Although not included in resistance profiling, rhodamine 6G has been shown to be a substrate of QacC (Paulsen et al., 1995). The fluorescence polarisation assay is based on the fact that polarisation of a fluorescent molecule depends on its rate of rotation (Weber, 1953). Thus, interactions of fluorescent molecules like rhodamine 6G, with proteins such as QacC, can be readily observed and quantified by measuring changes in the rotation rate. In this study, the level of polarisation was assessed and estimated that upon binding of rhodamine 6G with a non-fluorescent partner (QacC) the rate of rotational diffusion will slow resulting in an increase in the polarisation of fluorescent emissions, which become maximal upon saturation of the fluorescent substrate (Hassan et al., 2009; Jameson and Ross, 2010). To assess QacC-mediated binding of rhodamine 6G, saturation binding curves were determined, and subsequently used to establish the dissociation constant (K_D).

Using steps outlined in Section 2.15, the interaction of rhodamine 6G with QacC was determined by titrating the QacC protein and monitoring the polarity of rhodamine 6G fluorescent emissions during the titration. Fluorescence polarisation binding assays were undertaken in dialysis buffer (Section 2.15, Table 2.1) which was used in the final step of QacC dialysis. This buffer was supplemented with 0.05% DDM to maintain QacC solubilisation. The DDM concentration of 0.05% and the final concentration of rhodamine 6G used in this study, 4 nM, were determined to be optimal in QacA rhodamine 6G binding studies (Hassan et al., 2009), and found here to be effective for studies carried out with QacC. A non-specific protein, BSA was used throughout this binding assay as a negative control. Solubilised in the same buffer as purified QacC, it was titrated into a solution of rhodamine 6G. Results showed that BSA titration had no effect on the polarity of fluorescence emissions from rhodamine 6G (Figure 6.11), indicating that it does not interact with

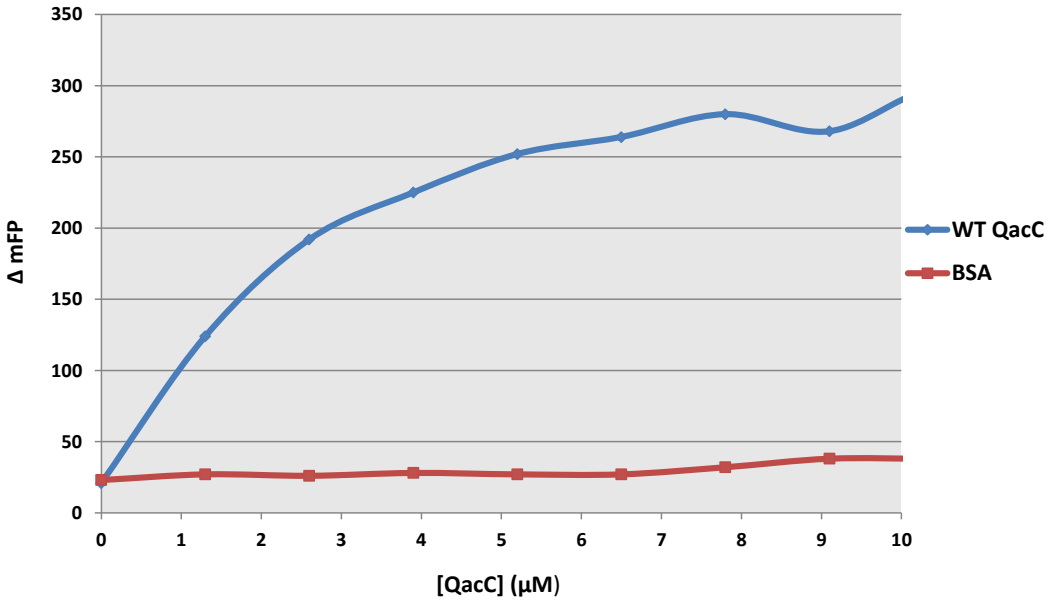

Figure 6.11: Rhodamine 6G binding to QacC analysed by fluorescence polarisation

The polarity of fluorescence emissions from 4 nM rhodamine 6G was monitored during titrations of the QacC protein (blue) and BSA (red) (Section 2.15). A representative saturation binding curve, plotting the change in rhodamine 6G millifluorescence polarisation (ΔmFP) against QacC protein concentration (µM), is shown. Using the equation stated in Section 2.15, a K_D value of 2.1±0.37 µM was determined for the QacC-rhodamine 6G interaction. Experiments were carried out over three separate trials, with a representative shown above.

the substrate. Titration of purified QacC into a solution of rhodamine 6G, however, resulted in a change in the rhodamine 6G milli-fluorescence polarisation (ΔmFP), as indicated in Figure 6.11. From these results, the K_D value determined for QacC-mediated interaction with rhodamine 6G in0.05% DDM was found to be 2.1±0.37 µM. Due to limitations in obtaining large quantities of QacC (see Sections 4.3 and 4.6) further analysis of the binding capacity of QacC with rhodamine 6G other fluorescent substrates like ethidium and acriflavine were unable to be carried out.

6.5 Discussion

The functional assessment of the QacC multidrug transporter was undertaken using a large panel of QacC-Cys mutants. Mutagenesis has been widely used to evaluate the function of residues within such transport proteins as the *E. coli* EmrE (Mordoch et al., 1999), MdfA (Adler and Bibi, 2004, 2005) and the *S. aureus* QacA (Hassan et al., 2007b; Hassan et al., 2008) and TetA(K) (Hassan et al., 2006b). In this chapter functional assessment of all QacC-Cys mutants was carried to evaluate the capacity of each mutant to actively extrude ethidium out of the cell. Optimisation of the efflux assay revealed that ideal conditions for ethidium transport occurred at 37°C using 15 µM ethidium. Evaluation of incubation temperatures indicated that temperatures above 42°C may have a destabilising effect of MPs, possibly increasing membrane fluidity to destabilising levels. Lower temperatures, on the other hand, may contribute to limiting the movement within the membrane, resulting in lower levels of transport. Concentrations of ethidium to be loaded into cells were also evaluated and found that although QacC could mediate extrusion of both 15 and 30 µM ethidium, the lower concentration of 15 µM was shown to minimise passive diffusion. Lastly, to demonstrate active efflux, both glucose and sodium formate were assessed as potential energisers for the efflux of ethidium. Results revealed that following the addition of 160 mM sodium formate, rapid active efflux of ethidium out of the cells occurred, whilst the addition of glucose, at any concentration tested did not lead to high levels of efflux. These conditions are in accordance with earlier studies carried out on QacC in proteoliposomes (Grinius

and Goldberg, 1994) and its assessment of ethidium transport in *E. coli* DH5α cells (Paulsen et al., 1995), as well as those carried out on QacA (Hassan et al., 2007b).

The results revealed that out of the 103 QacC-Cys mutants 15 were shown to have a significantly reduced capacity for ethidium efflux (0-60% of WT QacC) (Figure 6.1) and a further seven showed reduced efflux (61-70% of WT QacC). Analysis of QacC-Cys mutant expression (Section 5.2.3) showed that out of these, five had low levels of protein expression, possibly accounting for the lower levels of ethidium efflux. Examination of results also revealed that a number of the functionally impaired QacC-Cys mutants, contained mutations in regions identified to be highly conserved between SMR proteins (discussed in Section 1.4.3.1, Figure 1.9). As shown in Figure 6.1, QacC-Cys mutants of Tyr4, Ile11 and Glu14, which are putatively located in TMS1 and part of the first conserved motif, show significantly impaired efflux. Although the Tyr4Cys QacC mutant displayed lower than WT levels of expression both Ile11Cys and Glu14Cys were shown to be expressed as per WT QacC (Chapter 5). In addition to these, mutations to residues identified in the third signature sequence, Tyr60, Trp63, Gly65 and Gly67 also resulted in generating QacC-Cys mutants with an impaired capacity for ethidium efflux.

Residues within the TMS1 signature sequence (Figure 6.12) were further evaluated using fluorimetric transport assays coupled with NEM treatment (Section 6.2.2). QacC-Cys mutants shown to be reactive with fluorescein-5-maleimide (Figure 5.11) that were proposed to be located in TMS1 based on topology modelling programs, were pre-treated with NEM and their ethidium efflux capability evaluated. As shown in Figure 6.2, only Ala10Cys exhibited reduced efflux of ethidium following pre-incubation with NEM; such a reduction may be illustrative of the bound NEM interfering with binding and/or translocation. These results may also indicate that a residue at this position, in this case Ala10, may be part of the substrate binding pocket and/or translocation pathway. In contrast, pre-incubation of Ala19Cys with NEM resulted in an increased rate of transport, possibly indicating that the maleimide side chain may be a better substitute for the Cys side chain at this position. Rates of ethidium efflux for QacC-Cys mutants Ile11Cys and Phe20Cys were not affected by pre-incubation with NEM, suggesting

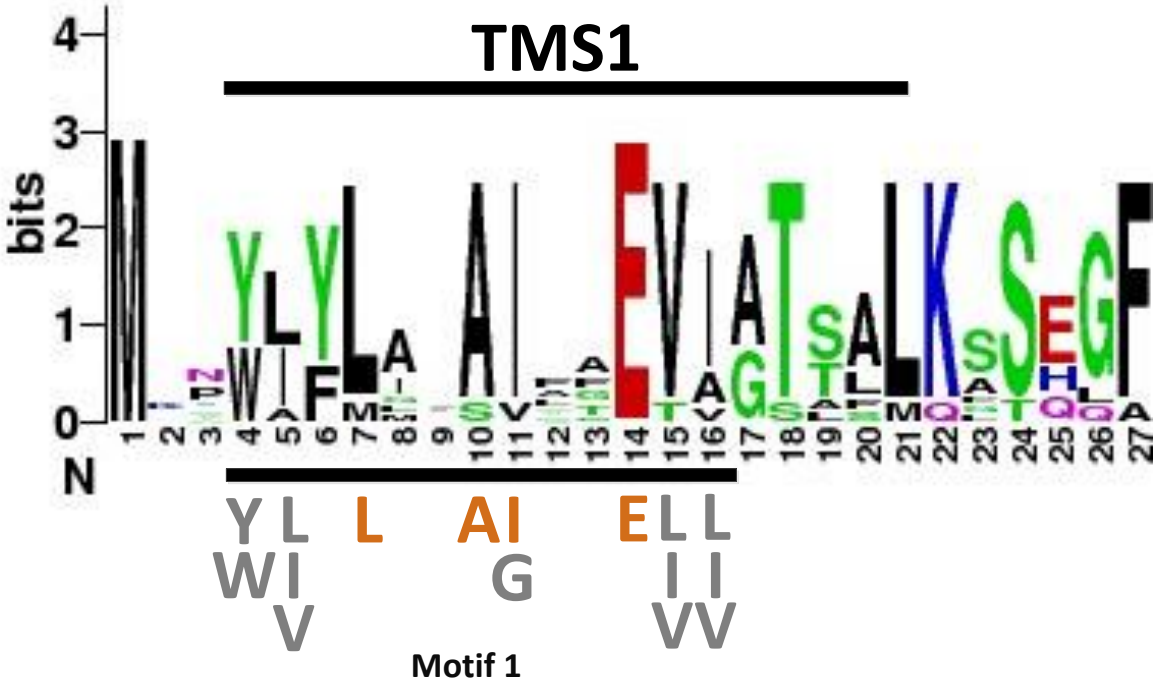

Figure 6.12: Sequence conservation in TMS1 in SMR members

Conserved residues in TMS1 as obtained from a multiple sequence alignment. Levels of residue conservation are shown on the scale which gives the certainty of finding that particular reside at that particular position and is determined by the frequency of that residue in that position. Red; acidic residues, black; nonpolar - hydrophobic residues, blue; basic residues, green; polar-uncharged, purple; polar, amide residues. A black bar above the sequence logo denotes the putative TMS1, whilst the black bar beneath the logo represents motif 1, with the residues making up the motif included in grey and orange. Orange residues in motif 1 that have been identified to be part of the substrate and proton binding site (Gutman et al., 2003). The sequence logo was generated using WebLogo http://weblogo.berkeley.edu/logo.cgi, with the amino acid position shown below the plot, and the bit score on the side.

that these positions within TMS1 do not play a role in substrate binding and/or translocation.

To assess the capacity of QacC-mediated efflux, fluorimetric transport assays were also carried out using pyronin Y (Section 6.2.3), a common SMR substrate. In contrast to ethidium, pyronin Y efflux is indicated by a rise in fluorescence, as binding of pyronin Y to RNA results in fluorescence quenching (Figure 6.3). Initial trials carried out with pyronin Y revealed that to use this substrate required significant optimisation, thus the experiments undertaken in this study served as proof of concept only. However, time permitting, pyronin Y transport assays could help identify those residues which may be specifically involved in transport of this compound. Although proteins like EmrE have been shown to have a common binding site for substrates and protons (Yerushalmi and Schuldiner, 2000a), as discussed in Section 1.4.3.3, the specific residues which contribute to this site have not been fully characterised. Mapping residues which are involved in the efflux of different substrates such as ethidium and pyronin Y, and comparing the results obtained may allow us to determine if different residues are being utilised in substrate binding and/or translocation for different compounds. In addition to pyronin Y, fluorimetric efflux assays with acriflavine were also attempted, however, only small changes could be detected in the fluorescence intensity of this substrate when the energy inhibitor CCCP was added (data not shown), and as such it was difficult to discriminate between WT QacC and mutants. Efflux assays using acriflavine as the fluorescent substrate have been used to assess such MPs as the VmrA from *Vibrio parahaemolyticus*, a Na^+ coupled MATE pump, as an alternative to ethidium (Chen et al., 2002). Although ethidium has been shown to be a common substrate for many MPs, in the case of VmrA, although a good substrate of this MP, efflux assays using ethidium did not generate good results, with only small changes in fluorescence intensity being detected. Trials using other compounds such as DAPI, another strong substrate of VmrA, also failed to generate significant results, leaving acriflavine as the only fluorescent substrate with which active drug efflux could be investigated. As such, assessment of drug efflux with certain compounds, like acriflavine in the case of QacC, may be more challenging to

carry out or simply not possible. Nonetheless, in addition to ethidium and pyronin Y, analysis of other fluorescent SMR substrates such as rhodamine 6G should, in future work, also be assessed in order to determine a more complete picture of the mechanism, or residues involved in the export of specific substrates.

In addition to efflux assays, all QacC-Cys mutants were evaluated for their ability to confer resistance against ethidium, acriflavine and pyronin Y. Results obtained by the microtiter plate assay (Section 6.3.1) revealed that of the 15 QacC-mutants found to be sensitive to all three compounds and a further 10 which displayed low QacC–mediated resistance against two of the substrates tested, four mutants, Leu7Cys, Lys22Cys, Pro32Cys and Phe79Cys, were shown to have impaired protein expression (Figure 5.4), and as such could not be fully assessed. Of these, Leu7Cys and Pro32Cys have been shown to be highly conserved in members of the SMR family (Bay and Turner, 2009), possibly implying structural and/or functional importance. For example, Leu7 has been identified in EmrE to be part of the binding domain (Sharoni et al., 2005). Hydrophobic residues such as Leu are commonly found in membrane segments and are known to have a high propensity for interhelical contacts (Adamian and Liang, 2001). In addition to Leu7Cys and Leu47Cys, Leu21Cys, Leu54Cys and Leu56Cys mutants all displayed an impaired capacity for ethidium efflux coupled with low MIC results. Given their hydrophobic nature, position within TMS and their propensity for interhelical interaction, they may play a role in protein stability, with Cys-substitutions disrupting these interactions.

Stability may also play a role in the low level of expression and subsequently impaired ethidium efflux capacity and low MIC values obtained for the QacC Pro32Cys mutant (Figures 5.4, 6.1 and 6.4). Previous MIC studies on Pro32 in QacC indicated that Pro32Ala and Pro32Gly mutants conferred reduced resistance against ethidium and crystal violet but still retained high-level resistance against QACs such as cetyltrimethylammonium bromide, benzalkonium, cetylpyridinium and dequalinium (Paulsen et al., 1995). These results indicate that Pro32 may play a role in substrate specificity, which is supported by findings that illustrate that membrane-buried Pro residues function in channel regulation via *cis-trans*

isomerisation of X-Pro peptide bonds (X denoting an unspecified residue) (Brandl and Deber, 1986). Given their compact side chain, Pro residues have also been identified to provide rigidity to protein structure and create a "kink" in the helix backbone. This kink may be of structural importance as it can provide a way for helices to be packed into their conformationally stable states. This may help to explain why a Cys-substitution in the Pro32 position, possibly influencing protein stability, leads to low levels of expression. As mentioned, a substitution of Pro32 to Ala retained partial function of QacC (Paulsen et al., 1995), however, a natural substitution of Pro to Ala in the medium subunit of the photosynthetic reaction centre from *Rhodopseudomonas viridis* revealed that this central residue maintains the overall shape of the helices and resembles the homologous protein containing the Pro residue in this location (Michel et al., 1986; von Heijne, 1991), which may explain the partial functionality found in QacC with this substitution. This may also explain why Pro55Cys, putatively located in loop 3, is still expressed at WT QacC levels but is unable to confer resistance against any of the three substrates tested. Its loop location excludes its possible role in structure and/or stability, but the Cys–substitution appears to disrupt its ability to mediate resistance.

As can be seen in Table 6.2, of the 103 QacC–Cys mutants, the function of 14 has been shown to be significantly impaired. These include the QacC mutant Glu14Cys, the highly conserved negatively charged residue identified to be at the centre of the substrate binding/translocation site (Grinius and Goldberg, 1994), Leu7Cys and Pro32Cys discussed above. Of the 14 mutants, seven are shown to be highly conserved and part of SMR signature sequences (Figure 1.9), indicating functional importance. Although not highly conserved a Cys-substitution of Gly17 rendered *E. coli* cells completely susceptible to all three substrates tested, even at levels below those determined for cells carrying the empty pBAD vector only. It is unclear as to why this particular mutation would have such a negative impact on the expressing cells, however, given such low levels of resistance it may be hypothesised that this mutation results in cells which are more permeable, allowing low levels of substrate to quickly enter the cells resulting in cell death. Given its position in TMS1, proximity to Glu14 and its inherent flexibility, the removal of this

Table 6.2: Table of impaired QacC-Cys mutants.

			MIC (Fold Resistance)		
Mutants[ab]	**Expression[c]**	**Efflux (% WT)**	**Ethidium**	**Acriflavine**	**Pyronin Y[d]**
Controls					
pBAD		34	1	1	1
QacC	3	100	4	2	16
Significantly impaired ethidium efflux and resistance					
L7C	1	35	1	1	2
E14C	3	27	1	1	1
G17C	3	32	1	1	1
S18C	3	42	1	1	4
P32C	1	54	1	1	1
F40C	2	25	1	1	1
M51C	3	35	1	1	2
L54C	3	58	1	1	1
L56C	3	58	1	1	1
Y60C	3	32	1	1	1
W63C	3	42	1	1	1
G65C	3	55	1	1	1
G67C	3	50	1	1	1
I94C	3	41	1	1	1
Reduced ethidium efflux and low resistance					
Y4C	2	39	2	1	4
L21C	3	38	2	1	2
L47C	1	43	2	2	1
I58C	1	33	2	1	2
Impaired ethidium efflux					
I42C	3	58	4	1	16
F44C	3	41	2	2	4

a: Red - identified to be part of SMR signature sequences (Section1.4.3.1, Figure 1.9)
b: Blue - SMR family conserved residues identified by Bay and Turner, 2009
c: Protein expression levels are indicated as 1: low, 2: medium and 3: as per WT QacC
d: A four-fold change in resistance to pyronin Y is not considered significantly impaired

			MIC (Fold Resistance)		
Mutants[ab]	Expression[ce]	Efflux (% WT)	Ethidium	Acriflavine	Pyronin Y[d]
Controls					
pBAD		34	1	1	1
QacC	3	100	4	2	16
Significantly impaired resistance					
F27C	3	113	1	1	1
K49C	3	64	1	1	2
H53C	3	87	1	1	4
P55C	3	69	1	1	1
A61C	3	67	1	1	2
L68C	3	61	1	1	2
F79C	1	71	1	1	1
K80C	3	81	1	1	2
E81C	3	86	1	1	1
Q82C	3	87	1	1	2
I86C	3	88	1	1	4
L93C	3	106	1	1	4
G97C	3	62	1	1	2
Low resistance					
K22C	1	69	2	1	1
G26C	3	105	2	1	2
S28C	3	99	2	1	2
Y45C	1	74	2	2	1
T62C	3	85	2	1	2
A64C	3	96	2	1	2
T71C	3	64	2	1	1
S75C	1	64	2	1	2
N84C	3	74	2	1	1

a: Red - identified to be part of SMR signature sequences (Section1.4.3.1, Figure 1.9)
b: Blue - SMR family conserved residues identified by Bay and Turner, 2009
c: Protein expression levels are indicated as 1: low, 2: medium and 3: as per WT QacC
d: A four-fold change in resistance to pyronin Y is not considered significantly impaired

residue may have led to an inability to switch between the protein's open and closed conformation (discussed in Section 1.4.3.3), thus allowing for compounds to pass through more easily leading to complete sensitivity of the cells.

In addition to the mutations which significantly impaired both transport and resistance, a number of Cys-substitutions resulted in mutants that did not transport ethidium at significant levels, but nevertheless maintained low levels of resistance to two of the three compounds tested (Table 6.2). Two mutants, Ile42Cys and Phe44Cys showed an impaired capacity for efflux, but nonetheless conferred moderate levels of resistance to substrates analysed. Of these, Phe44 is shown to be highly conserved in members of the SMR family, and appears to be located in close proximity to other aromatic residues, such as Phe40, Tyr45 and Phe46. It has been suggested that the role of aromatic residues is to anchor the protein into the membrane, by forming interactions with the lipid head groups with their aromatic rings. To carry this out, their preferential location within the protein is thought to be in the interfacial region (Ulmschneider and Sansom, 2001). In QacC, both Phe44 and Tyr45 are putatively located at the C-terminal end of TMS2 and thus may play a role in protein stability. Cys-substitutions at these positions may destabilise QacC, possibly explaining low expression levels as well as low levels of conferred resistance observed for the Tyr45Cys mutant, and the impaired capacity for transport of ethidium observed for Phe44Cys. The role of Phe44 in stabilisation has also been suggested in studies exploring the binding domain on EmrE, where it was speculated that the carboxyls of Glu14 are stabilised by interactions with the aromatic residues contributed from TMS2 and TMS3, as well as from Phe44 from TMS2 (Sharoni et al., 2005). Recent studies on EmrE have also proposed that this residue, amongst others located in TMS2, may play a role in substrate specificity, as a substitution of this residue with Trp eliminated ethidium resistance activity (Wang et al., 2014).

A number of mutations within QacC also resulted in mediating high levels of ethidium efflux despite conferring no or low levels of resistance against the compounds evaluated (Table 6.2). Six of these were found to be conserved residues within SMR signature motifs (Figure 1.9) whilst others like Lys49 and Lys80,

(discussed in Section 5.5) may play a role in the orientation of QacC monomers within its dimeric form.

From the results obtained by the microtiter serial dilution plate method (Section 2.10.1 and 6.3.1) both ethidium and pyronin Y were shown to be good substrates with which resistance could be analysed, whilst acriflavine was determined to be a poorer choice. Results also indicated that high expression of QacC from a pBAD and/or TOP10 background may have been a contributing factor to the overall results obtained in these studies. Previous studies assessing QacC conferred resistance were carried out in *E. coli* DH5α cells (Paulsen et al., 1995) and *E. coli* BHB2600 cells (Paulsen et al., 1993), on solid media. Solid media, using the plate dilution method (Section 6.3.2), has also been used extensively to assess conferred resistance to a range of compounds by such MPs as EmrE (Seppala et al., 2010; Yerushalmi and Schuldiner, 2000) and MdfA (Fluman et al., 2012; Tirosh et al., 2012). In this study, the plate dilution method and expression from the pBluescript SK II vector in *E. coli* BL21 (DE3) cells may aid in the generation of more representative results which are able to highlight subtle differences between each mutant, and its ability to confer resistance against various compounds.

In order to examine the resistance profile on solid media a serial dilution method was used, all strains grew in media supplemented with ethidium at 100 µg/ml and with acriflavine at 15.65 µg/ml, with the exception of Gly17Cys. As such, growth at the 10^{-1} dilution was expected for all *E. coli* BL21 (DE3) cells used in the assessment, including those carrying the empty pBluescript SK II vector. This allowed for a more sensitive evaluation of resistance between mutants. For example, His53Cys and Gln82Cys conferred no resistance to either ethidium or acriflavine when assessed by the microtiter plate method (Section 6.3.1, Table 6.2), however, grew at the highest dilution of 10^{-7} on plates supplemented with either ethidium or acriflavine (Section 6.3.2, Figure 6.8). Given that both mutants mediated high levels of ethidium efflux (both 87% of WT QacC) these results appear to corroborate observations obtained by the fluorimetric transport assays. Using this method also allowed for screening against acriflavine, which as noted, proved challenging utilising the microtiter plate method.

Although more sensitive, a significant contribution to the success of the plate dilution method was also due to the change of the expression vector and host *E. coli* cells. Essentially, *E. coli* TOP10 cells with QacC expressed from pBAD failed to outgrow control cells carrying the empty pBAD vector, on the selective media assessed (Figure 6.5). This is in contrast to *E. coli* BL21 (DE3) cells expressing QacC which grew at the highest dilution tested on all selective plates, unlike the control cells (Figure 6.6). As such, the use of the pBluescript SK II-based expression system was shown to be more reliable for measuring resistance (Figure 6.6). This may be due to the observation that *E. coli* TOP10 cells were not as viable post induction, as shown by a drop in growth following the addition of L-arabinose (Figure 4.3).

Lastly, kinetic experiments using fluorescence transport assays were performed to determine the K_m and V_{max} values for QacC-mediated export of ethidium and pyronin Y (Section 6.4.1). Although unable to determine the values for pyronin Y as a results of its high fluorescence and limitations with the equipment used, QacC-mediated efflux of ethidium produced K_m values in the low micromolar range indicating that QacC interacts with this substrate with high affinity. Studies into SMR family proteins like EmrE and SugE have also demonstrated such high affinities for substrates such as ethidium (K_m of 6 μM) (Sikora and Turner, 2005). Analysis of binding was also carried out using the fluorescence polarisation binding assay, this time with rhodamine 6G as the substrate tested (Section 6.4.2). Using this method, the binding of QacC to rhodamine 6G in 0.05% DDM was observed with a dissociation constant (K_D) in the micromolar range (2.1 ± 0.37 μM). This K_D value falls within the range observed for binding of EmrE to ethidium (6.81 ± 0.53 μM), methyl viologen (43.6 ± 3.80 μM), TPP (23.6 ± 7.10 μM) and cetylpyridinium (6.61 ± 2.20 μM) (Winstone et al., 2005).

6.6 Conclusion

The focus of this chapter was to assess the capacity of QacC and mutants to efflux ethidium out of the cell. Additionally, each mutant was evaluated for its ability to confer resistance against ethidium, acriflavine and pyronin Y. Overall results of fluorescent transport assays (Section 6.2.1) and resistance analysis

(Section 6.3.1) revealed that out of the 103 assessed QacC residues, only 14 were shown to be essential for function (Table 6.2), indicating that although small in size, many of residues within QacC can be mutated without deleterious effects. The importance of a number of these residues is highlighted by their strong conservation amongst members of the SMR family, and identification as essential for function, as seen for Glu14 and other residues forming the substrate binding site. Further investigation into these residues and their possible roles in other membrane transporters, from all drug membrane transport families, may help to determine if these play a specific role within a protein, within a family or whether their function is specific in all drug efflux pumps. Nonetheless, the results obtained here further support and broaden our collective understanding of the amino acids within membrane drug transporters, and the possible role or roles they may play in substrate binding and transport, protein stability and possible dimerisation.

In addition to evaluating the function of QacC, fluorimetric analysis and fluorescence polarisation assays were also used to determine the kinetic parameters for QacC. Found to be in the low micromolar ranges, the binding of ethidium, pyronin Y and rhodamine 6G were all found to be in accordance with values found for other SMR proteins like EmrE and SugE, indicating that QacC has a high affinity for its substrates. As different membrane mimetics can be used to investigate ligand binding, further research into various mimetic environments and how they can influence binding should also be carried out in order to ascertain if the membrane itself contributes to binding affinity. In addition to the methods used in this study, employing other methods which can be utilised to characterize the thermodynamics of ligand binding, such as isothermal titration calorimetry, could be beneficial in broadening our current knowledge of QacC-ligand binding.

CHAPTER 7 – QACC OLIGOMERISATION

7.1 Introduction

Much has been discussed regarding the quaternary structure of SMR proteins such as EmrE, particularly regarding their possible topologies otherwise referred to as “topoforms” (Schuldiner, 2012). It has now been established that EmrE functions as an oligomer, with studies indicating that the minimal functional unit is dimeric in form (Butler et al., 2004; Chen et al., 2007; Rotem et al., 2001; Ubarretxena-Belandia et al., 2003) (Section 1.5). Higher oligomeric states have also been proposed for EmrE, including the possibility that the functional unit is an oligomer which is formed by two or more dimers (Rotem et al., 2001). Like EmrE, the transport protein Hsmr of *H. salinarum* has been shown to function as a dimer, however, unlike EmrE the Hsmr dimer is resistant to denaturant treatment (Ninio and Schuldiner, 2003), making analysis of the physiologically relevant oligomeric form of the protein less challenging.

Much of the work carried out on membrane transporters like QacC is on a detergent-solubilised form of the protein, which when studying oligomerisation, has some shortcomings. For example, the structure of a detergent-solubilised protein may not be representative of its native lipid-bound form, as the molecular organisation of the protein in the bacterial membrane may be different (Friesen et al., 2000), resulting in inaccurate determination of the oligomeric form. Also, when working with detergents, the ratio of detergent to protein needs to be determined since coating of the protein with detergent will result in alterations in the molecular mass, leading to inaccurate molecular weight estimations of possible protein complexes (Heuberger et al., 2002).

Although challenging, there are number of techniques with which the quaternary structure of proteins in detergent solution can be studied. These include analytical ultracentrifugation, size-exclusion chromatography, blue native gel electrophoresis (BN-PAGE), co-purification and co-immunoprecipitation, all of which present their own challenges and drawbacks. In addition to these, chemical cross-linking in conjunction with various reactive cross-linking reagents, has also

been shown to be a general approach used to determine the quaternary structure of a protein (Lloris-Garcera et al., 2012; Morrison et al., 2012; Pakula and Simon, 1992). Amongst the reagents which can be used in this approach is formaldehyde, with its short cross-linking span of 2.3–2.7 Å and subsequent ability to easily permeate cell membranes (Nadeau and Carlson, 2007; Sutherland et al., 2008). Containing a single aldehyde group, formaldehyde is able to connect to amino side chains via a two-step reaction, resulting in a stable cross-link. Its primary amino acid targets are Lys and Trp residues, although the side-chains of residues such as Cys and Tyr, amongst others, have also been reported (Sutherland et al., 2008). In order to utilise formaldehyde in protein–protein interaction studies, optimisation of the procedure is required to identify short incubation periods and low concentrations of formaldehyde to limit the extent of non-specific cross-linking (Sutherland et al., 2008).

Formaldehyde cross-linking has been used extensively to analyse the *Shigella flexneri* polysaccharide co-polymerase membrane-bound protein Wzz_{SF} (Daniels and Morona, 1999; Papadopoulos and Morona, 2010) and TonB, and their interactions with other cell wall envelope proteins (Skare et al., 1993). In this study, formaldehyde cross-linking, using QacA as a monomeric control, was used to evaluate QacC oligomerisation. In addition to oligomerisation analysis, it was applied to assess the importance of the dimerisation motif ^{90}GLXLIXXGV^{98} identified within TMS4 of SMR proteins, which as discussed in Section 1.4.3.1, is thought to mediate TMS4-based SMR dimerisation (Poulsen et al., 2009). Additionally, BN-PAGE studies were carried out on DDM solubilised QacC, a method which has previously been used to determine the oligomeric state of MPs such as LacS (Heuberger et al., 2002), the *E. coli* carnitine transporter CaiT (Vinothkumar et al., 2006) and EmrE (Lloris-Garcera et al., 2012).

7.2 QacC–QacC interactions assessed by formaldehyde cross-linking

The use of formaldehyde as a cross-linking reagent was evaluated and optimised to determine the conditions which would lead to a high yield of oligomer formation and reduce non-specific cross linking. As mentioned previously, three

parameters have been shown to play a crucial role in formaldehyde cross-linking, these include; the temperature, duration of the reaction and formaldehyde concentration. Previous studies have determined that temperatures between 37°C and 25°C did not result in any observable differences pertaining to cross-linking, with RT serving as an ideal temperature (Klockenbusch and Kast, 2010). In this study, incubation time was trialled at one and two hours, and formaldehyde concentrations of 0.5% and 1% assessed. To evaluate cross-linking, whole *E. coli* TOP10 cells expressing QacC were divided into treated and untreated samples. Both treated and untreated samples were solubilised by the addition of 2X SDS-PAGE sample buffer (Table 2.1) followed by heating at 37°C for 30 minutes. As heating at 100°C has been shown to reverse formaldehyde cross-linking (Papadopoulos and Morona, 2010), a sample of treated cells were also heated at 100°C for 5 minutes, allowing for verification that protein-protein interaction is occurring only as a result of formaldehyde cross-linking. All samples were resolved by 15% SDS-PAGE and transferred to a PVDF membrane for Western Blot analysis (Sections 2.8.1, 2.9.2, 2.9.3 and 2.16.1).

Assessment of QacC-QacC interaction, and thus dimerisation, by Western immunoblotting, revealed that following incubation at RT for one hour with 1% formaldehyde, two bands were detected for the treated and solubilised at 37°C samples (Figure 7.1A). Only one band with a MW of approx. 15 kDa, was detected for the treated and heated at 100°C sample (Figure 7.1A). Although QacC monomer interaction was observed with treatment using 0.5% formaldehyde following one hour incubation at RT, this interaction was not as obvious as when incubated with 1% formaldehyde. However, at this higher concentration, it may be possible that non-specific cross-linking occurred, which may have led to the smearing observed in Figure 7.1B.

7.3 Assessment of the SMR oligomerisation motif in QacC

Gly is considered an unique amino acid as it lacks a side-chain and as a result can confer conformational flexibility to proteins (Curran and Engelman, 2003; Yan and Sun, 1997) . Frequently found in transmembrane helices of MPs, Gly residues

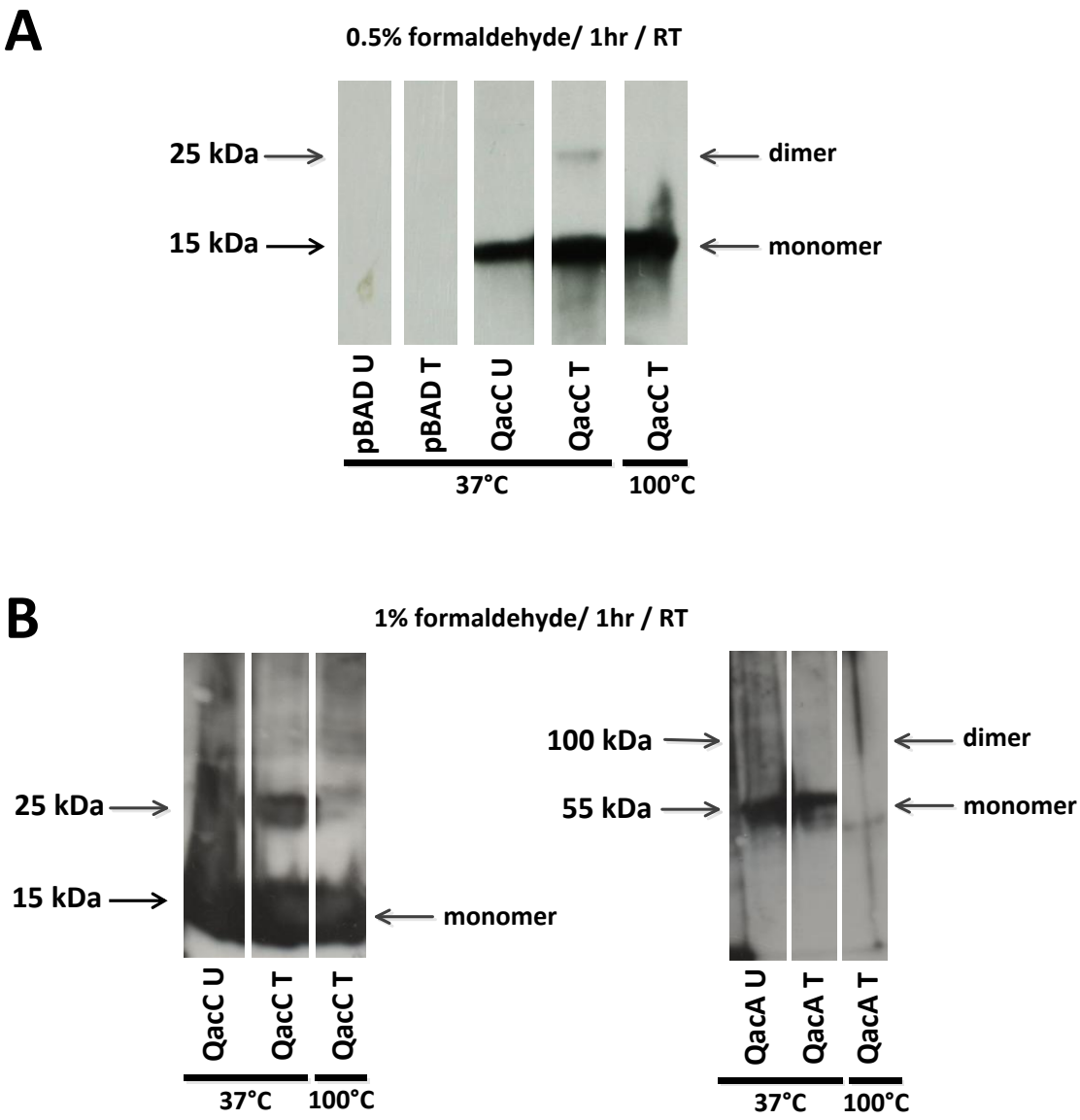

Figure 7.1: Analysis of QacC and QacA cross-linking with formaldehyde

Cultures of *E. coli* TOP10 and DHα cells carrying the pBADQacC_C43T clone and pSK7201 (Table 2.3) were expressed (Section 2.7.1) and harvested as set out in Section 2.16.1. Cells were washed in 10 mM KPO_4 and treated (T) by exposure to (A) 0.5% or (B) 1% formaldehyde at RT for one hour. Samples were heated at 37°C or 100°C and subjected to SDS-PAGE (Section 2.8.1) transferred to a PVDF membrane (Section 2.9.2) and probed with an anti-HISx6 antibody (Section 2.9.3). Untreated cells, heated at 37°C were also included to serve as controls (U). Bands indicative of the size of monomeric and dimeric QacC/QacA are indicated, whilst the 15, 25, 55 and 100 kDa marker size is shown with an arrow on the left of the Western blot.

permit short interhelical separation, possibly participating in Cα-H···O bonds (Senes et al., 2004; Senes et al., 2001). They have also been identified to be part of universal scaffolds which allow for the assembly of two transmembrane helices, and thus participate in helix-helix interactions (Melnyk et al., 2004). This scaffold is referred to as the GXXXG motif, where two Gly residues are separated by three amino acids (Treutlein et al., 1992). This motif has been identified in a number of studies to serve as an oligomerisation motif (Arselin et al., 2003; Lee et al., 2004; Lu et al., 2014; Mendrola et al., 2002). A longer version of this motif, GXXXXXXG, has also been identified to act as a scaffold for dimerisation, and was shown in studies on EmrE to play a role in the ability of the protein to form dimers (Elbaz et al., 2008).

The QacC protein has a total of eight Gly residues at positions 17, 26, 35, 41, 65, 67, 97 and 105, which with the exception of Gly26 and Gly105 are all putatively located within TMSs. Of these, four are highly conserved (Figure 1.9) and have been shown in this study (Section 6, Table 6.2) to lead to reduced function when substituted with a Cys residue. To further analyse QacC Gly residues, in particular Gly97 and its role in dimerisation, and the SMR proposed dimerisation motif GXXXXXXG, site-directed mutagenesis in combination with formaldehyde cross-linking was performed.

7.3.1. Mutagenesis of dimerisation motif residues

Identified within EmrE (Elbaz et al., 2008) and defined in the *H. salinarum* SMR protein Hsmr to be composed of ^{90}GLXLIXXGV^{98} (Poulsen et al., 2009) the assembly motif within QacC is ^{90}SIVLIIVGV^{98} (Figure 7.2). Compared to the proposed Hsmr motif, only four residues are conserved between the Hsmr and QacC sequences; Leu93, Ile94 Gly97 and Val98, with a Ser replacing Gly at position 90. However, as illustrated by Bay and Turner, (2009), and discussed further in Section 7.9, although shown to highly conserved, variations to Gly90 have been identified in SMR family members (Figure 7.3), with other small residues such as Ser and Ala, found at this position. As QacC possesses only one Gly residue in this proposed motif, to assess its possible role in dimerisation Ser90 was substituted to either Gly or Ile.

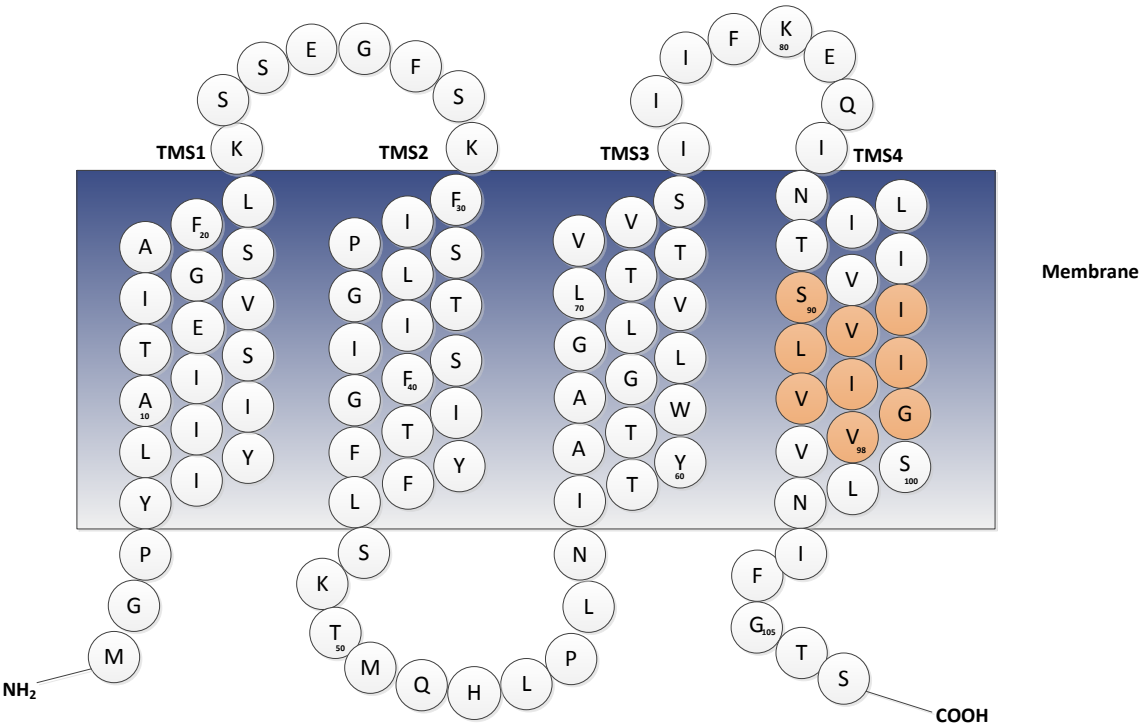

Figure 7.2: The proposed QacC dimerisation motif

A secondary structure model of QacC with residues belonging to the proposed SMR dimerisation motif, ^{90}SIVLIIVGV^{98}, shaded in orange.

A

```
AbeS|YP_001713109.1| 66 IILISAIGWIFYKQHLDLAACIGLALMIAGIVIINVFSK---------- 105
QacE|YP_855865.1|    74 ILVVSTLGMAFYGQHPDTAAMIGMAVIASGIVIMNLFSKMGSEEGTDASA 124
QacF|AAD22143.1|     67 IVLVAAIAWIFHGQKLDFWAFIGMGLIVSGVAVLNLLSKVSAH------- 110
QacH|AAZ42322.1|     67 IVLVAAIAWIFHGQKLDFWAFIGMGLIVSGVAVLNLLSKVSAH------- 110
QacH|AAX56371.1|     67 IVLVAAIAWIFHGQKLDLWAFVGMGLIVSGVAVLNLLSKVSAH------- 110
QacG|AAS79147.1|     67 IVLVTAIAWVLHGQKLDMWGFVGVGFIISGVAVLNLLSKASVH------- 110
EmrE|NP_951765.1|    67 TALVAVVAWVFMGQRLDLPAVFGILLIIAGVLVLNLCSTSATH------- 110
QacE|AAL52226.1|     67 VTLVALIGWLVFGQKLDLPAIVGMGLIIAGVIVLNLLSNTAQH------- 110
EmrE|NP_386521.1|    70 IVLISIAGYVVFGQKLDLPAILGLALIIAGVLVLNLFSKSTFH------- 113
QacF|YP_358817.1|    67 IVLITLMGWIVHKQTLDLAAMLGMGLIVSGVLVINIFSKSGAH------- 110
EmrE|NP_253677.1|    67 IVLVSLVAMFVYGQRLDPAALLGIGLIIAGVLVIQLFSRASGH------- 110
EbrB|YP_235623.1|    77 ITLITVIAAFRYQQVPTLMAIGGISLIIAGIILVNLAKEQAT-------- 119
EbrA|CAB13613.1|     67 TALTAIVGFLLFGETISLKGVFGLTLVIAGVVVLNQSKAHAEDKKQTACE 117
EbrB|EDN74090.1|     67 LVLTAIVGVVVFGEKVDFWGIASIGLILAGVIMMNTLSKMGGH------- 110
QacG|O87866.1|       66 LVLTTIISVIVFKENVNLISIISIGLIVIGVVLLNVFGESH--------- 107
QacJ|CAD55144.1|     66 LVLTTIVSVVIFKESVNLISIFSIILIIIGVVLLNVFGSSH--------- 107
QacC|AAM94142.1|     66 LVLTTVVSIIIFKEQINLITIVSIVLIIVGVVSLNIFGTSH--------- 107
Hsmr|OE3652F|        66 IVATALVGVVVFEESVDVAGVVGLALIVAGVVVLNVASDAYTPAH----- 112
Mmr|P9WGF0.1|        67 TAAIVLVAVLFLGSPISVMKVVGVGLIVVGVVTLNLAGAH---------- 107
QacE|NP_070524.1|    67 -----LLAFLLQLNRLKAFLFISAGVCIAGVVVLLASIGIIRLL------ 137
```

B

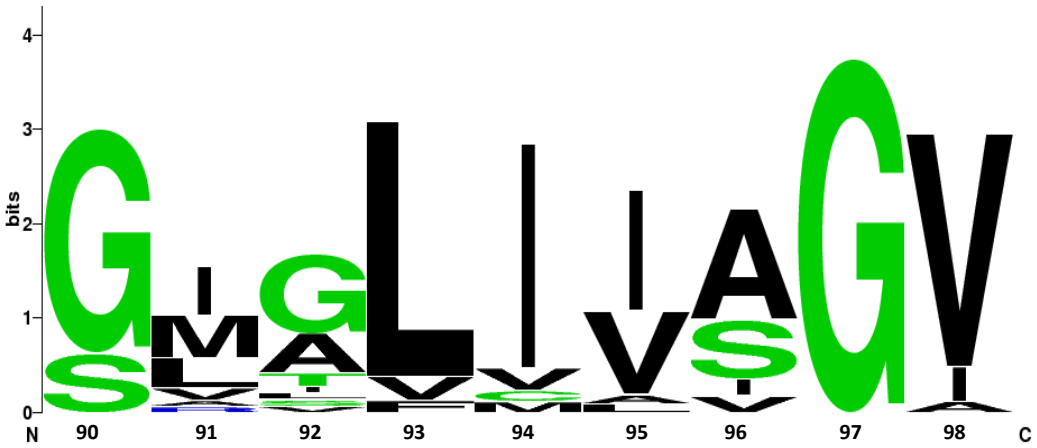

Figure 7.3: Multiple sequence alignment and sequence conservation plot of the SMR dimerisation motif

(A) Multiple sequence alignment of selected SMR proteins showing the conserved dimerisation motif in red. The name of each protein and Genbank accession number is given at the beginning of each sequence. (B) Sequence conservation plot of the dimerisation motif obtained from the above alignment using WebLogo (http://weblogo.berkeley.edu/logo.cgi). The plot shows the Weblogo bit score on the side and the amino acid position below the plot.

Gly was chosen to determine if this change would impact on or improve dimerisation and/or function, whilst Ile was chosen to assess whether the addition of an amino acid with a larger side-chain would alter dimerisation. Similarly, Gly97 was also substituted with Ile. In addition to these, a double mutant Ser90Cys_Gly97Cys was generated to determine if both Ser and Gly are needed for dimerisation and/or function. All mutagenesis was performed as described in Section 2.6.2 and discussed in Section 5.2. Once verified by sequencing (Section 2.5.4) protein expression levels of the QacC mutants were assessed by a Western blot (Section 2.9.3), where it found that although somewhat lower than WT QacC, all mutants were expressed at detectable levels (Figure 7.4).

7.3.2 Functional analyses of QacC Ser and Gly mutants located in the putative dimerisation motif

The capacity of QacC Ser90 and Gly97 mutants to efflux ethidium out of the cell was determined by a fluorimetric transport assay (Section 2.11.1). As seen in Figure 7.5, although cells expressing the Ser90Cys derivative were shown to mediate efflux of ethidium at essentially WT levels, a substitution to Gly at this position resulted in lower levels of ethidium transport (80% of WT QacC). Furthermore, the substitution to Ile resulted in mutants exhibiting ethidium efflux at only 60% of WT QacC. A substitution of Gly97 to Ile produced similar results to Ser90Ile, with reduced (60% of WT QacC) levels of ethidium transport. Although a substitution of Ser90 with Cys did not affect levels of ethidium exported out of the cell, as compared to WT QacC, a Cys substitution of Gly97 had a negative impact on the level of ethidium transport (60% of WT QacC), displaying levels of efflux on par with Gly97Ile. Lastly, low levels of ethidium efflux were also observed for the double mutant Ser90Cys_Gly97Cys (60% of WT QacC). Thus, although Ser90Cys mediated high levels of efflux, the additional mutation of Gly97Cys reduced the overall capacity of this mutant to export ethidium.

To extend the functional assessment of the mutants, the capacity to confer resistance against ethidium, acriflavine and pyronin Y was analysed using the

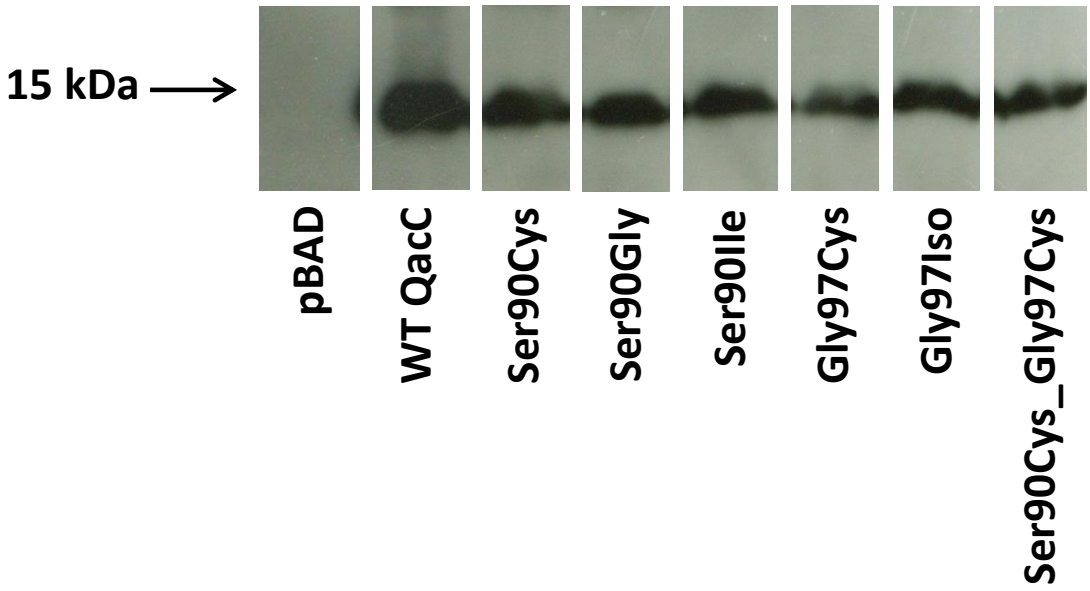

Figure 7.4: Expression of QacC–Ser90 and Gly97 mutants

Cultures of *E. coli* TOP10 cells carrying pBADQacC_C43T and QacC Ser90 and Gly97 mutant expression plasmids (Table 2.3) were grown to OD_{600}=0.55 in LBAmp and induced for two hours with 0.02% L-arabinose (Section 2.7.1). Cells were collected and membrane fractions isolated (Section 2.7.3). Membrane fractions were solubilised with 1% (w/v) DDM, and 20 µg of each sample was subjected to SDS-PAGE (Section 2.8.1), transferred to a PVDF membrane (Section 2.9.2) and probed with an anti-HISx6 antibody (Section 2.9.3). Locations of the QacC and QacC mutant protein bands (at 15 kDa) are indicated on the left-hand side of the Western blots.

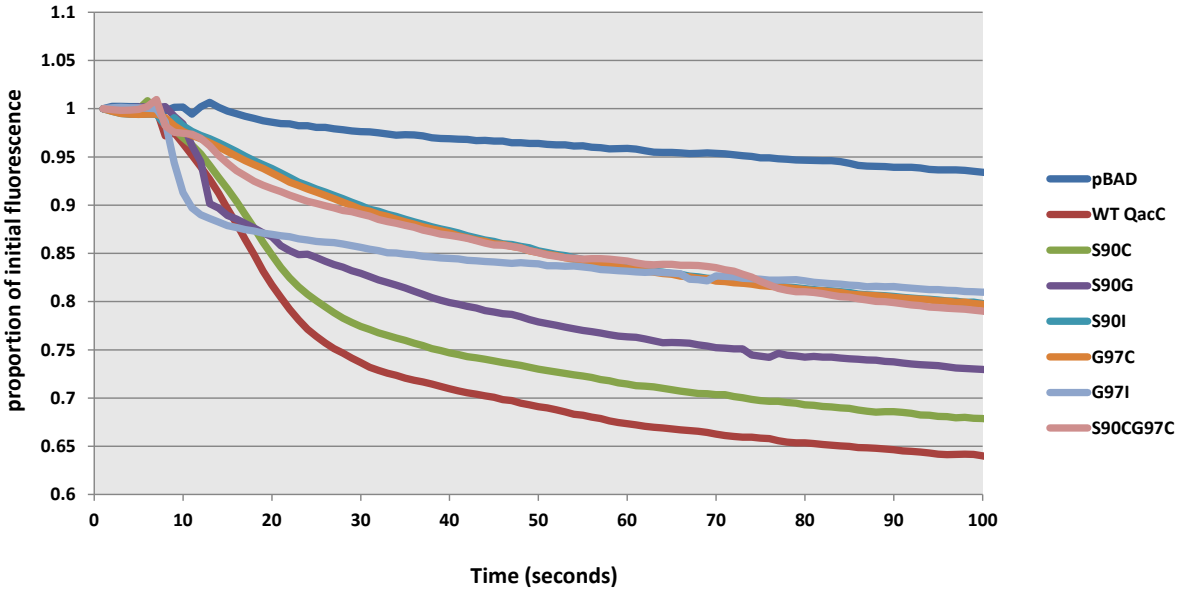

Figure 7.5: Ethidium efflux mediated by QacC and QacC Ser90 and Gly97 mutants

Fluorimetric assays were carried out on *E. coli* TOP10 cells carrying pBADQacC_C43T and pBAD-based vectors encoding Ser90 and Gly97 mutants (Table 2.3). Cultures were grown to OD_{600}=0.55 in LBAmp and induced for two hours with 0.02% L-arabinose (Section 2.7.1). Cells were collected and loaded with 15 µM ethidium bromide (Section 2.11.1) and ethidium efflux energised by the addition of 160 mM sodium formate at time zero (Section 2.11.1). *E. coli* TOP10 cells carrying the empty pBAD vector were used as a negative control (shown in blue). Efflux experiments were performed in at least triplicate with a representative shown.

microtiter plate dilution method (Section 2.10.1). As seen in Figure 7.6, cells expressing Ser90 mutants showed reduced levels of resistance against all three compounds, with Ser90Gly being the only mutant to show residual levels of resistance to ethidium and pyronin Y and comparable to WT QacC levels for acriflavine. These results are similar to those obtained with Ser90Cys. All QacC Gly97 mutants displayed poor overall drug resistance capacities. Lastly, a two-fold resistance against ethidium was observed for the double mutant Ser90Cys_Gly97Cys, reflecting the low levels of ethidium efflux observed in the transport assay (Figure 7.3).

7.3.3 Assessment of the dimerisation motif by formaldehyde crosslinking

To assess the possibility that the Ser90 and Gly97 mutations affected QacC dimerisation, formaldehyde cross-linking was carried out on each mutant. Following the steps outlined in Section 2.16.1 and using the optimised conditions discussed in Section 7.2, *E. coli* TOP10 cells expressing Ser90Cys, Ser90Gly, Ser90Ile, Gly97Ile and Ser90Cys_Gly97Cys were exposed to 1% formaldehyde for one hour at RT. Protein interactions were detected by Western blot analysis (Section 2.9.1 and 2.9.2) using anti-Hisx6 antibodies (Section 2.9.3) (Figure 7.7). As can be seen, in all treated reactions following heating at 37°C two bands can be observed, denoting a monomeric and a dimeric form of QacC, whilst in reactions following heating at 100°C, the upper dimeric band is missing. As seen in previous experiments (Figure 7.1), using formaldehyde at 1% appears to have an effect on the clarity of the results obtained by Western blot. Nevertheless, results obtained indicate that Ser90 and Gly97 mutations do not disrupt QacC monomer interactions.

7.3.4 Blue Native-PAGE analysis

To further analyse the oligomeric state of QacC and confirm the results obtained from formaldehyde cross-linking, BN-PAGE was used to separate QacC complexes under non-denaturing conditions. Samples of QacC were purified and dialysed (Section 2.7.4) and DDM added to a final concentration of 0.05%. Samples were prepared and loaded on a 4-16% Bis-Tris BN-gradient gel following steps

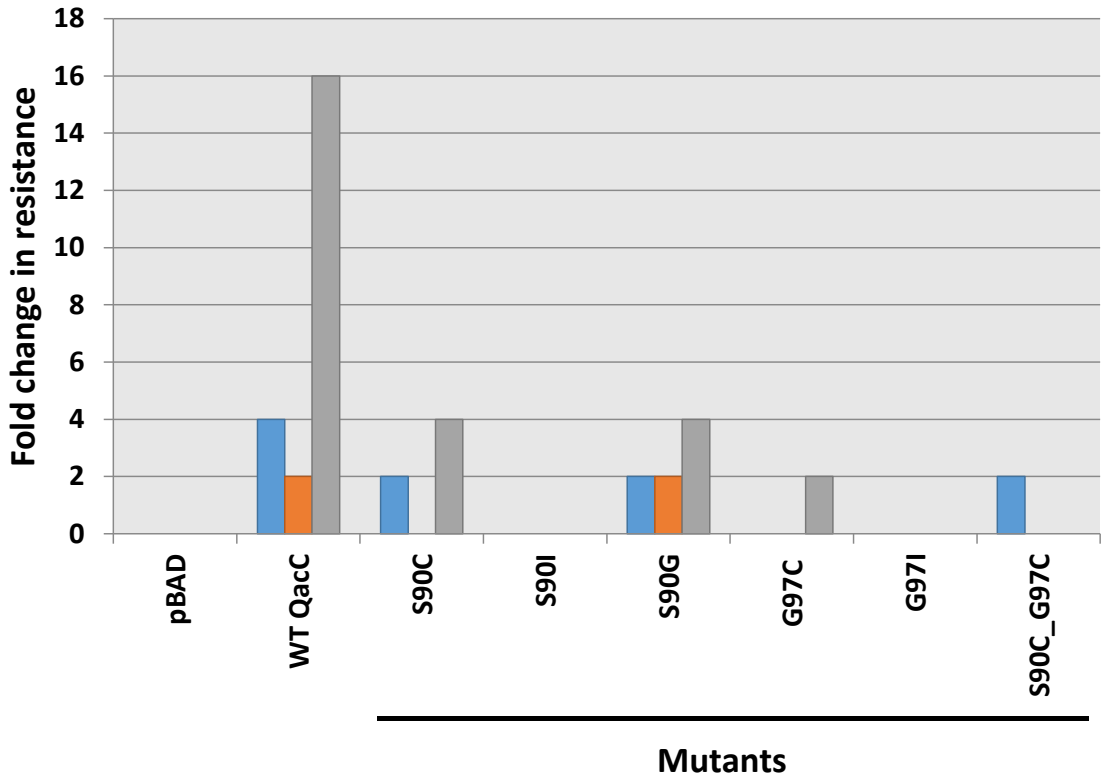

Figure 7.6: MIC_{90} results shown as fold change in resistance for ethidium, acriflavine and pyronin Y as obtained for *E. coli* TOP10 cells expressing QacC and the Ser90 and Gly97 dimerisation motif mutants

E. coli TOP 10 cells carrying pBAD-based QacC plasmids (Table 2.3) were grown in LBAmp and induced for two hours with 0.02% L-arabinose. Cells were adjusted to OD_{600}=0.6, diluted 1:100 and added to serially diluted compounds (Section 2.10.1). Absorbances were determined after overnight incubation at 37°C. MIC concentrations were compared with results obtained for cells carrying the empty pBAD vector only and fold change in resistance determined. Data shown are for ethidium (blue bars), acriflavine (orange bars) and pyronin Y (grey bars). Experiments were carried out in triplicate with the average shown.

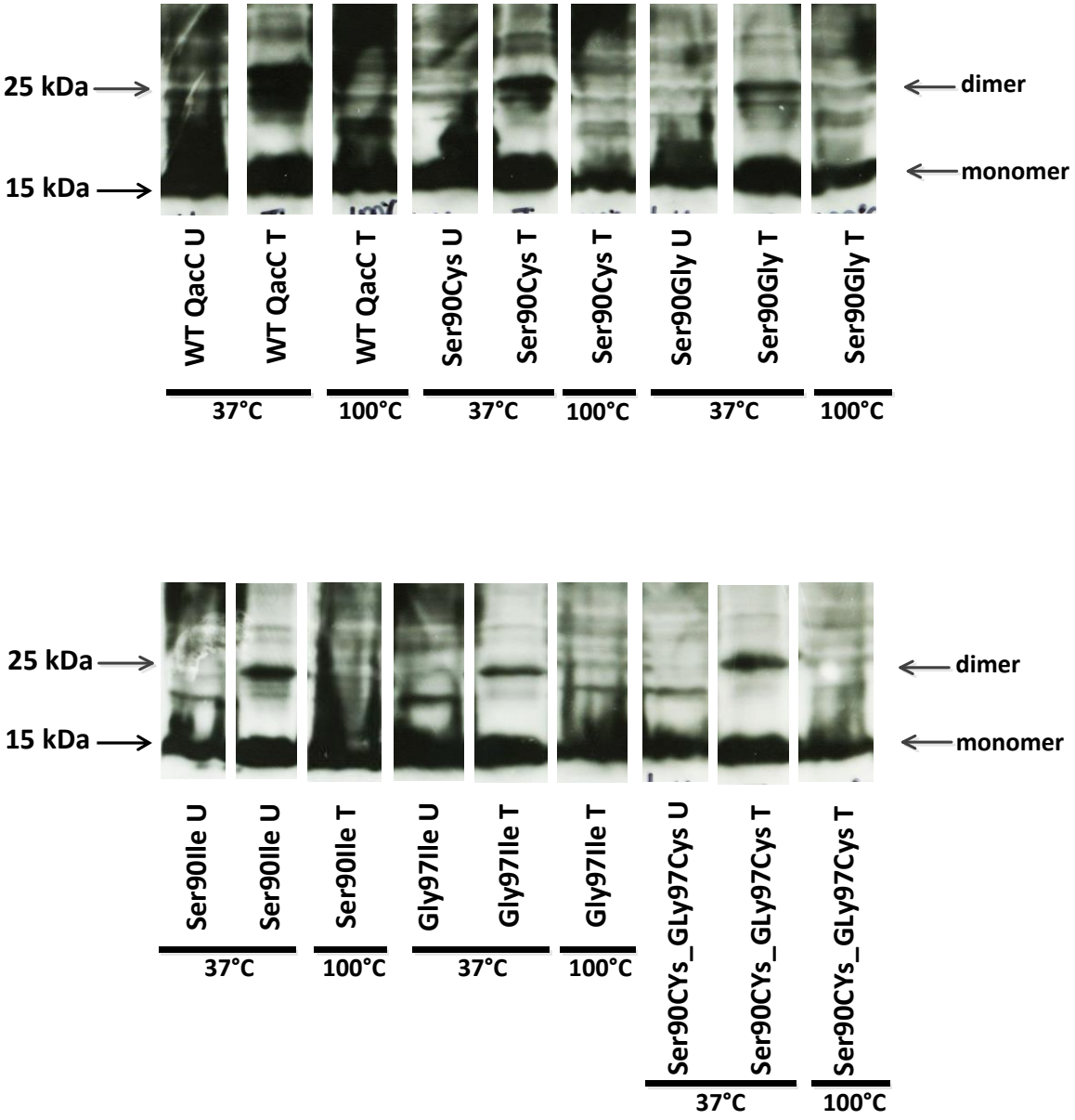

Figure 7.7: Analysis of the QacC and QacC–Ser90 and Gly97 mutants by formaldehyde cross-linking

Cultures of *E. coli* TOP10 cells carrying pBADQacC_C43T and QacC Ser90 and Gly97 mutant expression plasmids (Table 2.3) were grown to OD_{600}=0.55 in LBAmp and induced for two hours with 0.02 % L-arabinose (Section 2.7.1). Cells were washed in 10 mM KPO_4 and exposed to 1% formaldehyde (T) at RT for one hour. Samples were heated at 37°C and 100°C and subjected to SDS-PAGE (Section 2.8.1), transferred to a PVDF membrane (Section 2.9.2) and probed with an anti-HISx6 antibody (Section 2.9.3). Untreated cells, heated at 37°C were also included to serve as a control (U). Bands indicative of monomeric and dimeric QacC and QacC mutant proteins are indicated on the right, whilst the 15 kDa marker size is given with an arrow on the left of the Western blot.

outlined in Section 2.8.2. Following separation by BN-PAGE, samples were analysed by staining with Coomassie Brilliant Blue R-250 (Figure 7.8). To assess the monomeric form of QacC, solubilised samples (1% DDM) were denatured by the addition of sample buffer containing SDS (Section 2.8.1).

As can be seen in Figure 7.8, two bands are visible at an apparent MW of 66 and 146 kDa (population 1 and 2, respectively), for pBAD expressed QacC when solubilised in 1% DDM. However, no band is visible for the monomeric, denatured QacC.

7.4 Discussion

Studies into SMR proteins like EmrE and Hsmr have demonstrated that the minimal functional unit of these proteins is a dimeric form (Chen et al., 2007; Ninio and Schuldiner, 2003; Poulsen et al., 2009; Rotem et al., 2001). To evaluate the oligomerisation of QacC two separate approaches were undertaken; formaldehyde cross-linking was used to investigate the interactions between QacC monomers and BN-PAGE was used to assess the oligomeric state of QacC under non-denaturing conditions. Formaldehyde cross-linking along with mutagenesis was also used to evaluate a putative dimerisation motif which is thought to mediate dimerisation of SMR proteins.

Detection by Western blot revealed that following the addition of 1% formaldehyde to *E. coli* TOP10 cells expressing QacC (Section 2.16.1), an additional protein band running above the 25 kDa marker was present in treated samples which is missing in both the untreated and treated, and heated at 100°C samples. As only a single band (population 2) with a MW approximately twice that of the proposed monomer (population 1) is detected, it indicates that like EmrE and Hsmr, the native form of QacC is possibly dimeric. To assess the likelihood that this is a non-specific interaction, formaldehyde cross-linking was also carried out on QacA, which functions as a monomer (Hassan, 2007). Unlike QacC, only one band is visible for QacA treated with 1% formaldehyde, reconfirming its monomeric form.

Formaldehyde cross-linking was also used to assess the proposed SMR dimerisation motif, ^{90}GLXLIXXGV^{98} identified in TMS4 of SMR family members

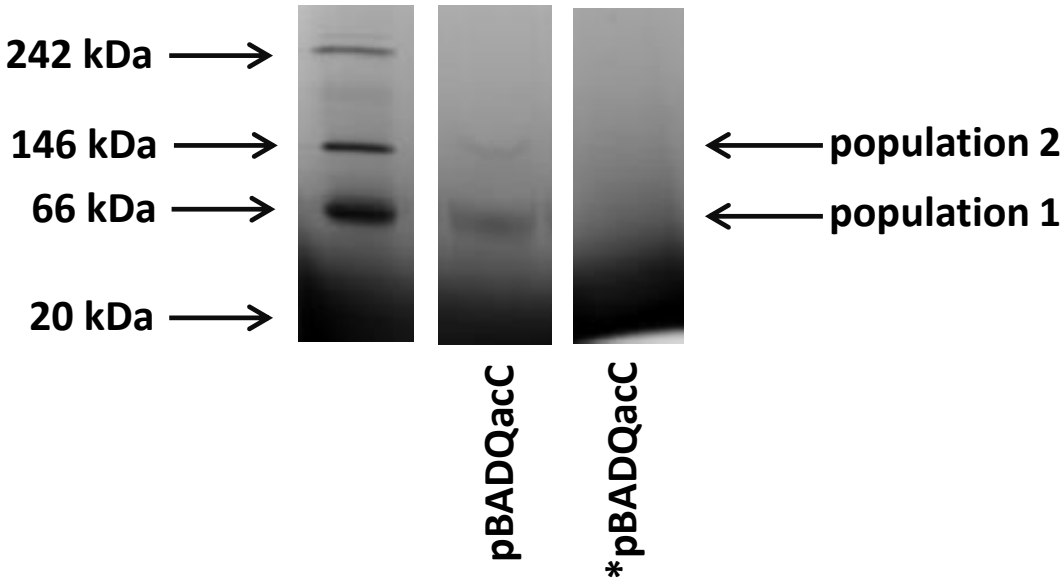

Figure 7.8: Analysis of QacC oligomerisation by BN-PAGE

Cultures of *E. coli* TOP10 cells carrying pBADQacC_C43T (Table 2.3) were grown to OD_{600}=1.0 in LBAmp and induced for 30 minutes with 0.02% L-arabinose (Section 2.7.1). Cells were collected and membrane fractions isolated and solubilised with 1% (w/v) DDM (Section 2.7.3). Proteins were purified and dialysed as set out in Section 2.7.4. To show QacC in its monomeric from, one sample, identified by an asterisk, was denatured by the addition of 2X SDS-PAGE sample buffer (Table 2.1) and heating at 37°C for 30 minutes. Both samples were subjected to BN-PAGE (Section 2.8.2). MW markers and their sizes are given on the left. Arrows on the right indicate the two populations of QacC observed in the BN-PAGE gel.

(Poulsen et al., 2009). In EmrE, it was shown that two essential Gly residues, Gly90 and Gly97, form a dimerisation motif in which they are separated by six other residues, and as such is referred to as the GG7 motif (Elbaz et al., 2008). According to the aforementioned study, a substitution of either of these Gly residues results in an impaired ability of EmrE to form dimers. Although shown to be essential in the EmrE and Hsmr and conserved amongst SMR members, variations to Gly at position 90 have also been identified. These include Ser, Ala, Cys and Phe, with a higher propensity for Ser occurring in non SugE members, implying certain flexibility in this position (Figure 7.3). A higher level of conservation can be seen for position 97, with only a limited number of SMR members displaying a Ser or an Ala residue for this position (Bay and Turner, 2009). Given this, it may be understandable why a Cys substitution for Ser90 does not impair QacC function and possible dimerisation; however, a replacement with an Ile residue, which has a much bulkier side-chain, leads to a reduced capacity for QacC to efflux ethidium out of the cell (Figure 7.5) and abolished resistance against ethidium, acriflavine and pyronin Y (Figure 7.6). As levels of conservation appear to be higher for Gly in position 90 (Figure 7.3), a Gly substitution was constructed for Ser90 in QacC. The QacC Ser90Gly mutant did not outperform WT QacC, instead it showed lower levels of efflux and reduced MIC levels for ethidium and pyronin Y (Figures 7.5 and 7.6). These results may indicate that although Ser is not essential in this position to retain function, it can only be replaced with another small residue. In order to investigate this further, additional substitutions of small and large amino acids at this position could be undertaken and dimerisation assessed.

Evaluation of Gly97, unlike Ser90, indicates that the Gly residue cannot be replaced by any other residue without significant reduction to protein expression and function (Figures 7.4, 7.5 and 7.6). In comparison to Ser90 QacC mutants, all substitutions to Gly97 resulted in reduced expression levels. In addition to expression, functional importance was also observed for this residue, as Gly97 QacC mutants were able to mediate only low levels of ethidium efflux, approx. 60% of WT QacC. Resistance, nonetheless, was shown to be severely compromised for these mutants, with the exception of Gly97Cys, which displayed very low levels of

resistance against pyronin Y (two-fold change); all other Gly97 mutants were shown to be susceptible to all compounds examined. Together these results indicate an importance for a Gly residue at this position. Although functionally significant, studies evaluating its potential role in dimerisation, as determined by formaldehyde cross-linking, failed to establish its necessity in dimerisation. As shown by Western blotting (Figure 7.7), bands corresponding to dimeric QacC were seen for all treated QacC Ser90 and Gly97 mutants and absent in the untreated samples. The use of alternative crosslinking methodologies may further help to fully evaluate this motif and whether it is essential for QacC function and dimerisation. Alternatively, additional analysis utilising techniques such as negative-dominance, analytical ultracentrifugation and size-exclusion chromatography would need to be carried out in order to fully evaluate SMR oligomerisation and the proposed dimerisation motif.

In addition to formaldehyde cross-linking, QacC oligomerisation was evaluated using BN-PAGE. Results revealed that under non-denaturing conditions, two forms of QacC exist, these possibly being monomeric and dimeric. Analysis of oligomerisation, particularly the oligomeric state of MPs which need to be solubilised in detergents by BN-PAGE is in itself challenging. Bound lipids, detergent and Coomassie Brilliant Blue R-250 dye all add mass to the protein being analysed, leading to considerable discrepancies in their migration (Heuberger et al., 2002). As a result of this discrepancy, determination of the mass and oligomeric state of MPs becomes problematic. As determined in Figure 7.8, two bands migrating at 66 kDa and 146 kDa were observed, indicating that QacC takes on a higher oligomeric form. However, difficulties arose when trying to discern how many monomeric subunits are present in these multimers. Studies have reported a conversion factor of 1.8, or an 80% mass increase in the MW of a protein resolved on a BN-PAGE when compared to its true MW (based on the amino acid sequence) (Heuberger et al., 2002). Although this conversion factor appears to account for the MW size discrepancy observed for large MPs such as LacS, which at 71 kDa runs at 142 kDa on a BN-PAGE, it may not be ideal for the conversion of smaller proteins such as QacC. Running at 66 kDa and adjusted by the 1.8 conversion factor, would indicate

that QacC, in its monomeric form is 36 kDa, which does not match the proposed MW of 66 kDa. Two possibilities may be considered, the first being that the 66 kDa band actually represents the dimeric form of QacC, with the 146 kDa representing a tetramer. However, it may be conceded that QacC, solubilised in 1% DDM, leads to a greater mass increase than the 80% reported for larger proteins, and that that 66 kDa band is in fact representational of the monomeric form of QacC. Limitations with this study make analysis of the results obtained difficult to interpret.

One of these limitations pertains to the small nature of QacC and the lowest detectable MW marker included in the gel. At 20 kDa, the soybean trypsin inhibitor runs at the very edge of the Coomassie Blue G-250 dye front, which all but masks it making detection difficult. The front obscures all results which may be present at the lower margins of the gel. Attempts to shorten the duration of the electrophoresis failed to minimise the width of this dye front. There is thus the possibility that the monomeric denatured QacC added as a control is simply too small to be detected by BN-PAGE, explaining its absence in Figure 7.8.

The mass protein standards currently used in BN-PAGE also represent a limitation, especially when analysing the oligomerisation of MPs. Soluble standard proteins like the soybean trypsin inhibitor mentioned above, have been shown to have a different migration behaviour to that of MPs, as they do not contain bound lipid, detergent and Coomassie Blue G-250 dye, which as mentioned can equal or even exceed the mass of MPs (Heuberger et al., 2002; Wittig et al., 2010; Wittig and Schagger, 2008).

It is obvious that further analysis into QacC oligomerisation is needed, especially when using BN-PAGE. Alternative methods improving the sensitivity of QacC detection on BN-PAGE may need to be applied. These could include radiolabelling, as was carried out on studies investigating EmrE oligomerisation (Lloris-Garcera et al., 2012). Nevertheless, the results in this study have shown that QacC can form homo-oligomers, as shown by formaldehyde cross-linking and BN-PAGE; however, its exact oligomeric form is still unknown. Future attempts at electrophoresis may include the use of High Resolution Clear Native Electrophoresis (hrCNE) (Wittig et al., 2007), which omits the use of the Blue Coomassie G-250 dye

from the sample and cathode buffer. This approach may allow for the resolution and detection of smaller proteins such as QacC, which are obscured by the dye front in the traditional BN-PAGE. Also, the inclusion of MPs as migration markers instead of the currently used soluble standard proteins, could also lead to a better estimation of the MW of membrane transport proteins in BN-PAGE.

7.5 Conclusion

The aim of the work described in this chapter was to evaluate QacC oligomerisation and to identify and assess the proposed SMR dimerisation motif. *In vivo* formaldehyde cross-linking of QacC indicated that QacC forms dimers. Using this methodology together with site-directed mutagenesis, allowed for the SMR dimerisation motif GG7 in QacC to be analysed, where it was found that position 90, which in QacC is Ser90, is essential but not specific to Ser or Gly, as a Cys substitution does not impair functionality, nor does it disrupt dimerisation. The position of Gly97, however, does appear to be essential, with all substitutions function impairing. Dimerisation, as indicated by formaldehyde cross-linking, does not appear to be affected by mutations to Gly97, however, it would appear that although still possibly forming a dimer, QacC forms an inactive structure, as indicated by its loss of function.

In addition to formaldehyde cross-linking, QacC oligomerisation was also assessed by BN-PAGE. Results obtained further supported QacC oligomerisation, with two forms of QacC detected, possibly monomeric and dimeric. Due to the limitations of this approach, especially when analysing small proteins such as QacC, it was difficult to completely interpret and establish whether the results obtained were indicative of dimerisation, or of higher oligomeric forms. As such, refinement which includes using different markers and hrCNE instead of BN-PAGE may lead to clearer results.

CHAPTER 8 – ANALYSIS OF THE PUTATIVE MULTIDRUG RESISTANCE PROTEIN SEPA

8.1 Introduction

The *S. aureus* SepA protein (GenBank protein id: BAB83936.1) was identified as a member of the SMR family of transport proteins. Composed of 157 amino acids, it was first discovered in 2002 in antiseptic-resistant isolates of *S. aureus* found in Japan (Narui et al., 2002). These strains displayed a low-level of resistance to antiseptics, which upon assessment could not be attributed to QacA, QacC nor mutations in NorA. Analysis of these strains, involved digesting chromosomal DNA and cloning the obtained fragments into *E. coli* DH5α cells, which identified a clone conferring acriflavine resistance. Nucleotide sequencing of this clone identified an insert of 4064 bp consisting of five ORFs. However, only three ORFs were complete, with the first (ORF1) missing the 5'-terminal region and the last (ORF5) missing the 3'-terminal region. Following further assessment using fluorimetric transport assays, it was revealed that one of the complete ORFs (ORF2, 3 or 4) was responsible for the extrusion of ethidium out of the cells. Deletion of ORF2 and ORF3 and further screening with acriflavine identified ORF4 as the gene giving rise to antiseptic resistance. This ORF4 was named *sepA* (staphylococcal efflux pump gene) and was shown to consist of 474 bps which encoded a protein of 157 amino acids with a MW of approximately 18900 Da (Narui et al., 2002).

Further analyses of SepA, showed that cells expressing this protein were able to extrude ethidium out of the cell. Nonetheless, the MIC results compiled for ethidium, amongst other compounds, revealed identical MIC values of 64 μg/ml for both *E. coli* DH5α cells harbouring the plasmid with the cloned chromosomal fragment and *E. coli* DH5α cells carrying an empty plasmid (pUC18), used as a negative control. Although these results appear to contradict those obtained for the ethidium efflux assay, it has since been established that ORF3 encodes the *S. aureus* SdrM (GenBank protein id: BAB43261.1) multidrug transport protein. Belonging to the MFS family (Section 1.3.2.2), this chromosomally-encoded membrane transporter has been shown to confer resistance to acriflavine and to a

much lesser degree, ethidium (Yamada et al., 2006a). As such, although still not explaining MIC results obtained for ethidium, the efflux observed may be attributed to the SdrM MP, and not to SepA.

Attempts to reproduce the findings obtained by Narui et al., (2002) proved to be unsuccessful (Sandercock, 2010). Investigation into the acriflavine resistance levels mediated by SepA showed levels of resistance *on par* with the negative control, with further analyses unable to support findings that SepA conferred resistance to acriflavine (Section 8.6). Despite this, given its conservation throughout the staphylococcus genome, the function of SepA raises some interesting questions, in particular with a focus on its structure. Although considered to be a member of the SMR family, SepA does not fit the "typical" structural conformation of other members of the SMR family (Section 1.4). As illustrated in an alignment of SepA with other SMR transporters (Figure 8.1), residues which have been identified as essential, and thus conserved amongst members of the SMR family (Figure 1.9), although being present in SepA, are further along the sequence, located within the last three putative TMS. This indicates that although sharing some SMR protein characteristics, such as a small size and conserved residues, SepA appears unique with a function that has yet to be fully determined.

Following its initial documentation as a multidrug efflux transporter (Narui, Noguchi et al. 2002), SepA has been included in a number of publications; however none have confirmed its function as a drug efflux pump. Studies evaluating overexpression of *S. aureus* multidrug efflux pumps, including SepA, following single and multiple *in vitro* exposures to biocides such as benzalkonium, chlorhexidine, acriflavine and ethidium failed to observe increased expression of SepA (Huet et al., 2008). Screening studies looking at the expression of multidrug transporters in *S. aureus* in the presence of ethidium, also failed to identify SepA amongst those capable of effluxing this compound (Patel et al., 2010). Thus, although screened for and further investigated, no supporting evidence identifying SepA as an active multidrug transporter exists.

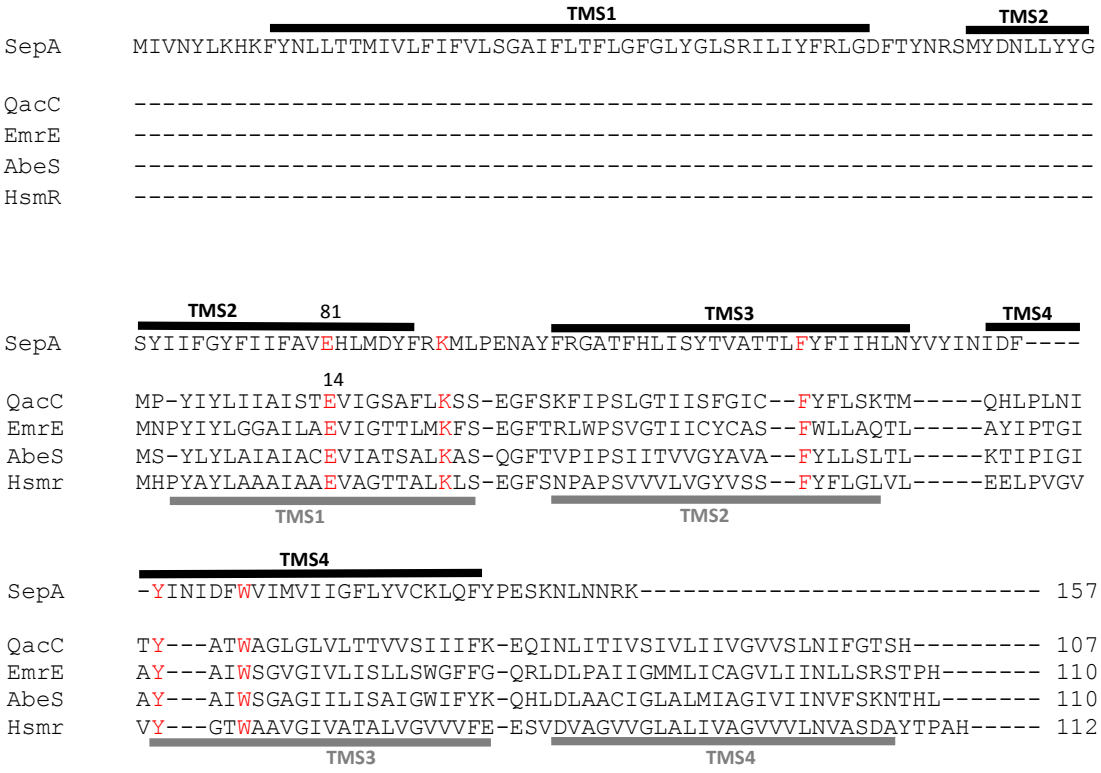

Figure 8.1: Multiple sequence alignment of SepA with related members of the SMR family

Multiple sequence alignment of SepA and related members of the SMR family generated using ClustalW (Section 2.17.1). Putative TMS are indicated with horizontal bars above and below the alignments (black-SepA, grey-QacC). Well conserved residues identified in the above listed SMR proteins (QacC, EmrE, AbeS and Hsmr) are used as markers (indicated in red) to align and illustrate the relative positions of the putative TMS. Sequences were obtained from the GenBank database using the following proteins accession numbers: SepA (BAB83937.1); QacC (NP_647561.1); EmrE (P23895.1); AbeS (FJ843079) and Hsmr (YP_001689722.1).

This chapter describes the expression and purification of the SepA membrane transport protein. Given the position of the essential residues identified in related SMR export proteins and where they lie in SepA, various constructs were generated to test the importance of the first TMS. To evaluate the role of SepA and its variants in antimicrobial resistance, the MICs of selected antimicrobials, detergents and dyes were examined. Finally, fluorimetric efflux assays were performed to confirm its ability to efflux substrates out of the cell.

8.2 Sequence analysis of SepA

The *sepA* nucleotide sequence (GenBank accession number: AB078343), was used in a standard BLAST search to assess its conservation in the *S. aureus* genomes currently available on the NCBI database. This identified the *sepA* sequence to be well conserved (90-100%) throughout all sequenced *S. aureus* genomes. To further investigate the classification of the SepA protein as a multidrug transporter, the SepA amino acid sequence (GenBank protein identification: BAB83937.1) was used in a BLASTp search (Section 2.19.1), generating a number of pairwise alignments, with a large majority annotated as being the multidrug drug efflux pump SepA. Although predominantly identified in *S. aureus*, hits were also found for other staphylococcal species. One identified a hypothetical protein known as SmrB, found in *Staphylococcus lugdunensis* HKU09-01 which shows high sequence homology with SepA (71% identity). However, no experimental analysis exists on this protein, and as such its function remains unknown. A search excluding staphylococcus sequences (taxid:1279) returned three BLAST hits, including a multidrug resistance protein SepA in *Macrococcus caseolyticus*, with an identity score of 46%, as well as two hypothetical proteins from Hymenobacter species (identity score of 27%) and *Criblamydia sequanensis* (identity score of 33%). None of these, however, have been experimentally analysed and their functions remain unknown.

Given the above results, an amino acid alignment with SepA and a number of confirmed SMR members was carried out using ClustalW (Thompson et al., 2002) (Figure 8.1). To determine the homology between these proteins, pair-wise alignments between SepA and each protein were also generated using the NCBI

BLASTp suite-2-sequences (Table 8.1). Results revealed that SepA shares very little homology with any of the proteins assessed. However, the conserved functionally-important residues found in SMR proteins are present in SepA, although deviating from the relative location in which they are found. To examine the possibility that SepA may have undergone changes which have contributed to the shift of essential SMR residues, two variants of SepA were constructed. The first, SepAΔ1, consists of TMS2 to TMS4 and was constructed to examine the possible role and/or necessity of the first TMS of SepA, which as seen in Figure 8.1 does not share any sequence homology with other members of the SMR family. The second construct, SepAΔ1QacCTMS4 was generated in which TMS4 from QacC was fused to the last TMS of SepAΔ1 (TMS4 of SepA). This was carried out in order to determine if the "correct" placement of the residues identified to be essential in other SMR members would impact and/or possibly restore SepA function. A multiple sequence alignment of all SepA constructs, along with QacC, illustrates the positions of the putative TMSs as well as the location of essential residues of SMR family members (Figure 8.2).

Although displaying similarity to membrane transporters belonging to the SMR family, such as size, number of TMS, as well as certain essential residues found in SMR members, it is possible that SepA belongs to a yet unknown novel family. As discussed in Section 1.3.5, recent analysis of the *A. baumannii* genome revealed a novel family of chlorhexidine pumps, the PCE family (Hassan et al., 2013). As members of this family were predicted to be comprised of four TMS, and contained a conserved Glu15 residue shown to be essential for transport, an alignment between AceI, a member of the PCE family (Section 1.3.5) and SepA was also performed to evaluate the possibility that SepA may belong to this family (data not shown). The alignment revealed very low levels of homology between these two proteins, with only 6% sequence coverage. To evaluate an alignment with SepA and other PCE family members a multiple sequence alignment was carried out aligning Glu15, an essential PCE protein residue (Section 1.3.5) to Glu81 in SepA. Unlike the alignment between SepA and members of the SMR family, shown in Figure 8.1,

Table 8.1: Summary of results obtained from a pairwise alignment of SepA[a] with QacC, EmrE, AbeS and Hsmr generated by the NCBI protein-protein BLAST suite using the BLOSUM62 matrix

Protein	Query cover	E value	Identity	Accession number
QacC	18%	2.9	24%	194775
EmrE	6%	0.42	40%	252145
AbeS	22%	0.26	31%	12003
Hsmr	26%	0.89	31%	93381

a: SepA accession number used in the alignment: BAB83937.1

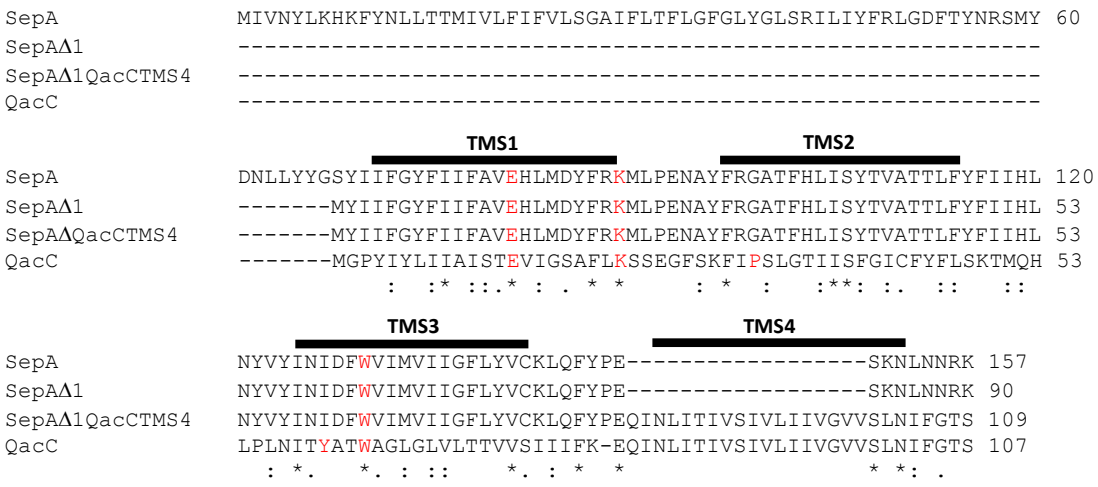

Figure 8.2: Multiple sequence alignment of SepA with SepAΔ1, SepAΔ1QacCTMS4 and QacC

Multiple sequence alignment of SepA, SepAΔ1, SepaAΔ1QacCTMS4 and QacC generated by ClustalW (Section 2.17.1). The putative TMS as determined for QacC are displayed by horizontal black bars. Asterisks indicate fully conserved residues, colons indicate strongly similar residues and dots indicate weakly similar residues. Dashes represent gaps. Essential residues found in members of the SMR family are indicated in red. Sequences were obtained from the GenBank database by the use of the following protein accession numbers: SepA (BAB83937.1) and QacC (NP_647561.1). Sequences of SepAΔ1 and SepAΔ1QacC are based on the SepA and QacC sequences, with the deletion (SepAΔ1) and addition of sequences (TMS4 of QacC) based on SMR sequence analysis.

aligning Glu15 to Glu81 in SepA did not result in a better fit between these proteins (Figure 8.3). Whether this is relevant or not is unclear, as at this point is unknown if Glu81, or any other Glu residue is essential for SepA function.

8.3 Construction of SepA, SepAΔ1 and SepAΔ1QacCTMS4 variants

8.3.1 Cloning of SepA and variants into pBluescript II SK

To assess the structure and function of SepA, genomic DNA from *S. aureus* NCTC 8325 was extracted and used as a template from which to amplify *sepA* by PCR using oligonucleotides SepAFor and SepARev (Table 2.4) (Section 2.6.1). This amplified PCR product also encodes a C-terminal Hisx6 tag for subsequent purification of SepA. The resulting PCR product was digested with *Nco*I and *Xba*I (Section 2.6.4) and cloned into a similarly digested pBluescript II SK vector.

To construct the SepAΔ1 derivative that has the first TMS of SepA removed, the pBluescript II SK-based SepA plasmid was used in a PCR with a forward primer, SepA Truncated and the pBluescript II SK sequencing reverse primer, M13_rev (Table 2.4). The PCR amplicon was digested with *Kpn*I and *Pst*I and cloned into a similarly digested pBluescript II SK plasmid (Section 2.6.4). Using this variant, the SepAΔ1QacCTMS4 derivative was then constructed, see Figure 8.4 for details. All constructs were transformed into *E. coli* DH5α cells (Section 2.4.2), verified by sequencing (Section 2.5.4) and subsequently used for expression and functional assessment.

8.4 Expression and detection of the pBluescript II SK-based SepA, SepAΔ1 and SepAΔ1QacCTMS4 proteins in *E. coli*

Initial protein expression trials on SepA in pBluescript II SK were carried out in *E. coli* DH5α cells and tested for “leaky” expression (Section 4.5). However, analysis of protein expression from whole cell lysates (Section 2.7.2) on a Western blot (Section 2.9.2) using the anti-Hisx6 monoclonal antibody failed to detect expression of SepA (data not shown). In order to induce expression of the pBluescript II SK-

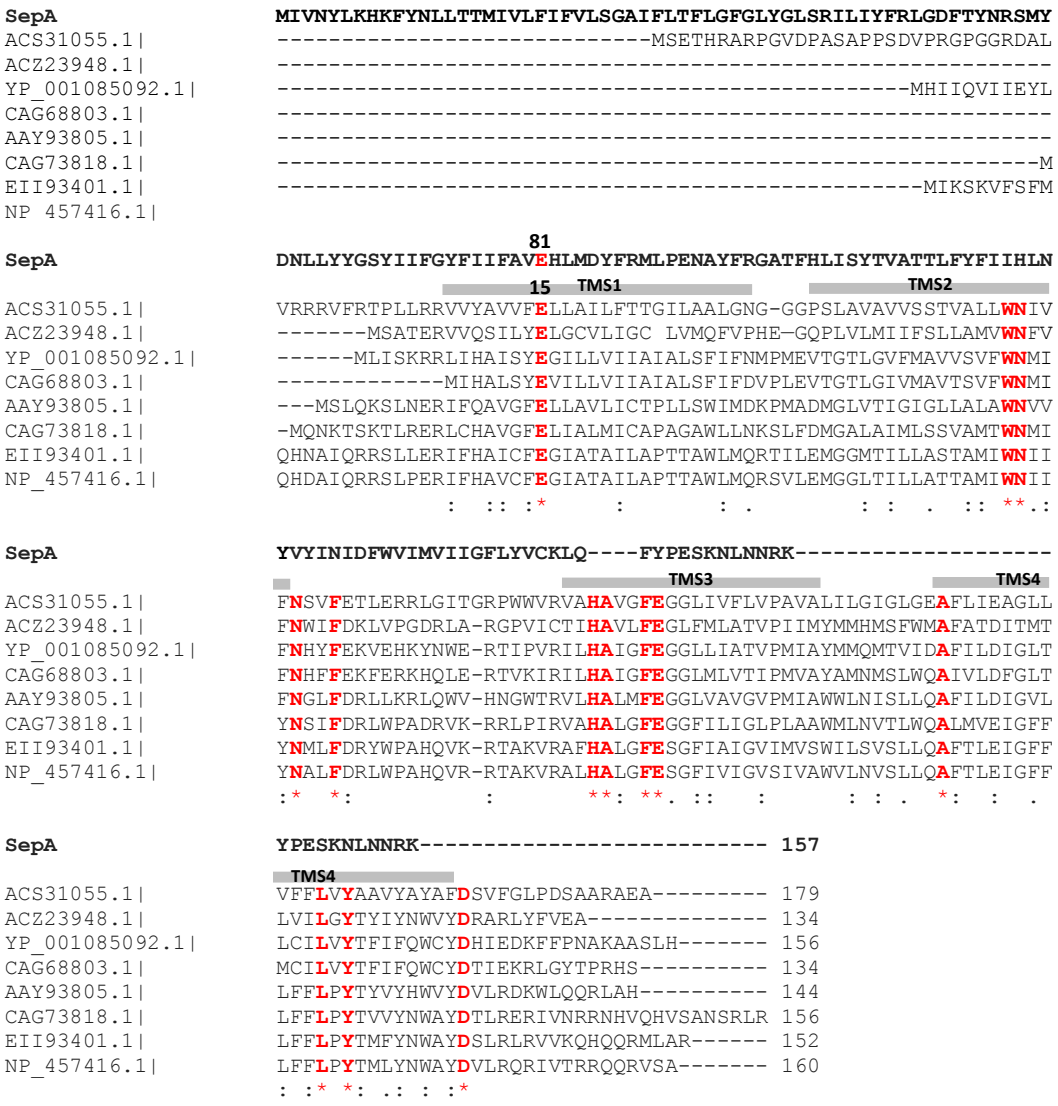

Figure 8.3: Multiple sequence alignment of SepA with members of the PCE family

Multiple sequence alignment of SepA and members of the PCE family generated with ClustalW2 (Section 2.17.1). The putative TMS as determined for QacC are displayed by horizontal grey bars. Asterisks indicate fully conserved residues, colons indicate strongly similar residues and dots indicate weakly similar residues. Dashes represent gaps. Well conserved residues in the PCE proteins are used as markers (indicated in red) to align and illustrate the relative positions of the amino acids within PCE members and SepA. Sequences were obtained from the GenBank database using the accession numbers listed on the right hand side of each sequence. GenBank protein sequence accession numbers; SepA-BAB83937.1, AceI-YP_00108592.1, all other sequence accession numbers belong to uncharacterised proteins.

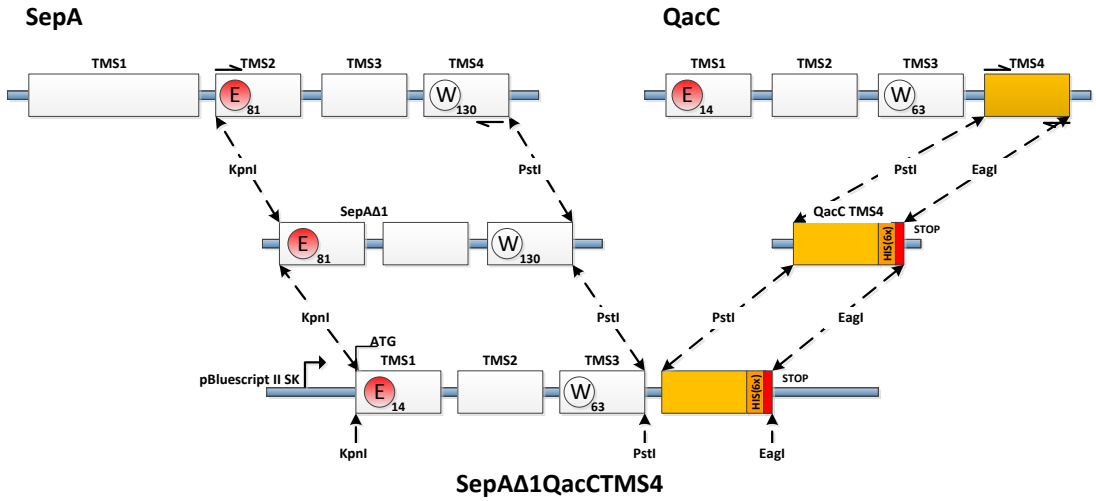

Figure 8.4: Construction of the pBluescript II SK-based SepAΔ1QacCTMS4 variant

Linear map representing the relevant positions of the TMS of SepA and QacC proteins and in the final SepAΔ1QacCTMS construct. Half arrows above and below SepA TMS2, TMS4 and QacC TMS4 represent primers used to generate the sub-sections of SepA and QacC fused to generate SepAΔ1QacCTMS4. The location of the restriction sites used in the ligation of the two segments, *Kpn*I, *Pst*I and *Eag*I are indicated with broken arrows. Two conserved SMR residues are shown in circles, with the essential Glu14 shown in red. The QacC TMS4 segment used in the construction of the SepA variant is shown in orange, with the C-terminal Hisx6 tag and stop codon boxed in dark orange and red, respectively.

based SepA and SepA constructs, *E. coli* BL21 (DE3) cells were used and expression induced with IPTG. Using the optimal induction conditions that were established for QacC (Section 4.5), cells with pBluescript II SK-based SepA, SepAΔ1 and SepAΔ1QacCTMS4 were induced at OD_{600}=0.55 with 0.1 mM IPTG followed by two hours of growth at 37°C. Membrane fractions obtained for all three constructs (Section 2.7.3) were resolved on an SDS-PAGE gel (Section 2.8.1), transferred (Section 2.9.2), and probed with an anti-Hisx6 antibody (Section 2.9.3). Results revealed that the SepAΔ1 protein was expressed at levels comparable to SepA (Figure 8.5). Migrating at a faster rate than SepA in SDS-PAGE; the molecular mass of the SepAΔ1 construct appeared to be approximately 12 kDa when compared to soluble protein markers (data not shown). These results were supported by computational analysis of the MW of SepAΔ1-based on its amino acid sequence (http://pir.georgetown.edu/pirwww/search/ comp mw.shtml), which predicted the MW to be 11.82 kDa. In comparison to both SepA and SepAΔ1, the pBluescript II SK-based SepAΔ1QacCTMS4 construct was not expressed at high levels. Low expression levels may be indicative of unstable or misfolded proteins which could result in aggregation of the protein as inclusion bodies. To explore this further, a Western blot (Section 2.9.2 and 2.9.3) was carried out on whole cell lysates (Section 2.7.2) of induced pBluescript II SK-based SepAΔ1QacCTMS4 along with SepA and SepAΔ1 (data not shown). Results obtained were then compared to those shown in Figure 8.5, where it was seen that levels of expressions from whole cells were the same as those obtained from membrane fractions. These results indicate that the low levels of expression of SepAΔ1QacCTMS4 are as a result of the inherent property of the construct itself. Despite the lower expression levels seen for SepAΔ1QacCTMS4, as expression was still detectable, this variant was included in the functional analysis of all three constructs.

8.5 Functional assessment of SepA and SepA variants

Initial functional assessment of SepA and variants focused on the evaluation of the transport capacities of all proteins. Fluorimetric efflux assays (Section 2.11.1)

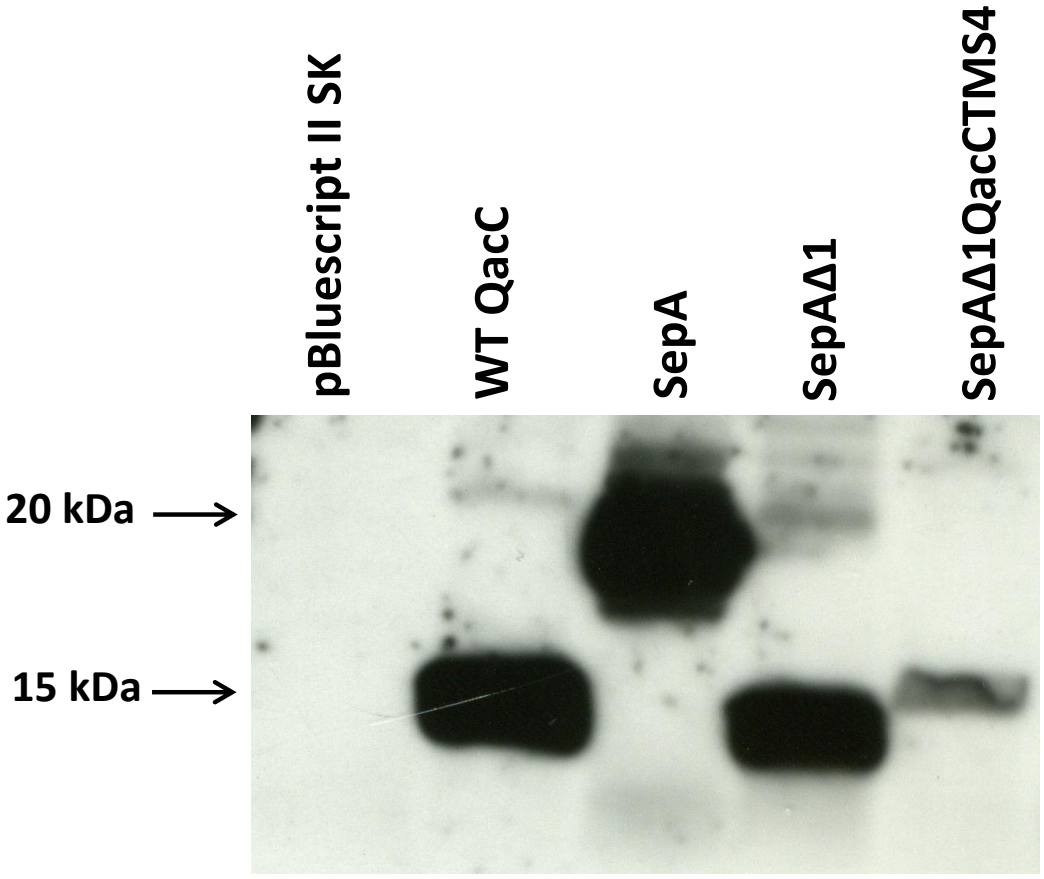

Figure 8.5: Expression of SepA and variants from pBluescript II SK in *E. coli* BL21 (DE3) cells

Cultures of *E. coli* BL21 (DE3) cells carrying the pBluescript II SK-based SepA and SepA variant expression plasmids (Table 2.3) were grown to OD_{600}=0.55 in LBAmp and induced for two hours with 0.1 mM IPTG (Section 2.7.1). Cells were collected and membrane fractions isolated (Section 2.7.3). Membrane fractions were solubilised with 1% (w/v) DDM, and 20 µg of each sample was subjected to SDS-PAGE (Section 2.8.1), transferred to a PVDF membrane (Section 2.9.2) and probed with an anti-HISx6 antibody (Section 2.9.3). Locations of the SepA and variant protein bands (at 15 and 20 kDa) are indicated on the left-hand side of the Western blot.

conducted in *E. coli* BL21 (DE3) cells carrying the pBluescript II SK plasmids expressing SepA and variants, were initially carried out to determine whether these proteins could efflux the common SMR substrate ethidium. However, despite the significant levels of expression of both SepA and SepAΔ1, results showed that none of the SepA proteins could facilitate the transport of ethidium out of the cell (data not shown). To further assess the transport capacity of these proteins, an acriflavine efflux assay was carried out (Section 8.5.1). The choice of acriflavine as substrate was based on the findings observed in the initial analysis of SepA, which showed that *E. coli* DH5α cells expressing this protein displayed elevated resistance levels against acriflavine (Narui et al., 2002) (Section 8.1).

8.5.1 Acriflavine efflux assay

To assess whether SepA and variants were capable of actively transporting substrates out of the cell, an acriflavine efflux assay was carried out on *E. coli* BL21 (DE3) cells expressing each SepA protein. In contrast to ethidium, the addition of acriflavine and subsequent binding to DNA results in a decrease in fluorescence intensity. Efflux of acriflavine subsequently decreases the concentration of intracellular acriflavine and leads to its dissociation from DNA, resulting in an increase in fluorescence (Chen et al., 2002).

Initial trials examining the final concentration of acriflavine to be used for assessment of efflux revealed that, like pyronin Y (Section 6.2.3) when added in high concentrations, fluorescence intensity reaches levels beyond detection for the Perkin Elmer LS 55 fluorescence spectrophotometer. Thus, refinement of the final concentration of acriflavine which could be used and efficiently monitored was needed. Further analysis revealed that the optimal concentration of acriflavine to load the cells was 96 nM (Section 2.13). Assessment into the amount of CCCP to be added was also evaluated, although the standard concentration of 10 µM was used, trials using higher concentrations were also explored. Based on a study carried out on the *Vibrio parahaemolyticus* VmrA efflux pump, where 30 µM was used to deenergise the cell (Chen et al., 2002), concentrations of 10 to 48 µM were trialled.

As assessment of resistance against acriflavine (Section 8.6) by the microtiter plate (Section 8.6.1) and agar plate dilution methods demonstrated that neither SepA nor its variants conferred acriflavine resistance, it was difficult to ascertain whether the acriflavine transport assays were effectively measuring the true efflux capacity of the proteins. This uncertainty was further compounded by the lack of a proper positive control which could verify the effectiveness of the transport assay itself. Although attempts to utilise the QacC multidrug transporter were undertaken, its low level of resistance against acriflavine was insufficient to establish consistently clear results which could be used as a comparison for active acriflavine efflux. Results of the transport assay revealed that both SepA and SepAΔ1, and SepAΔ1QacCTMS4 mediated acriflavine efflux at levels that were 18-20% and 10%, respectively, higher than cells carrying the empty pBluescript II SK vector (Figure 8.6). These results indicate that SepA and the SepAΔ1 protein are capable of effluxing acriflavine, albeit at a very low rate.

8.6 Resistance profiling of SepA and SepA variants

8.6.1 Microtiter plate analysis

To evaluate the role of the SepA protein in antimicrobial resistance, the MICs of various detergents and dyes were determined in MH broth in triplicate, using the microtiter plate analysis method (Section 2.10.2). Initial assessment of MICs focused on substrates shared in common with other members of the SMR family, including ethidium, acriflavine and pyronin Y, however, as these failed to show any significant resistance levels, the panel of compounds tested was expanded. Additional compounds that were evaluated included detergents such as SDS, Triton X and Tween 20, and the cationic dye acridine orange. Nonetheless, MIC results revealed that *E. coli* BL21 (DE3) cells expressing SepA or its derivatives, did not exhibit resistance to any of the compounds assessed (data not included).

Studies of multidrug transporters such as MdfA from *E. coli* have shown that changes in the external pH have an effect on resistance towards certain drugs (Lewinson et al., 2003). External pH of 6.5, 7.0 and 7.5 was thus also evaluated in

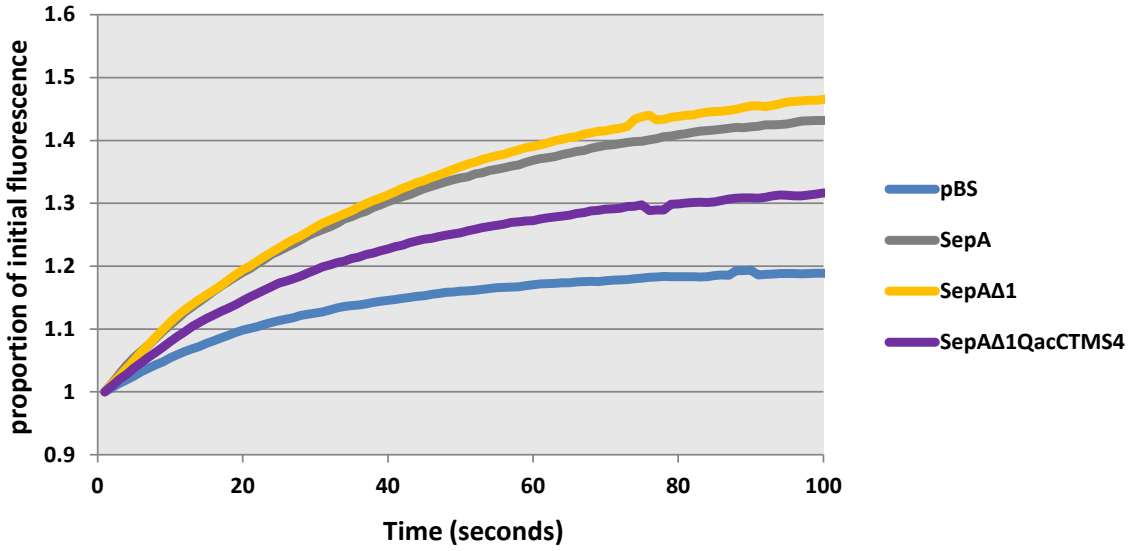

Figure 8.6: Acriflavine efflux assay of *E. coli* BL21 (DE3) cells expressing SepA, SepAΔ1 and SepAΔ1QacCTMS4

Fluorimetric assays were carried out on *E. coli* BL21 (DE3) cells harbouring pBluescript II SK-based plasmids encoding SepA, SepAΔ1 and SepAΔ1QacCTMS4 (Table 2.3). Cultures were grown to OD_{600}=0.55 in LBAmp and induced for two hours with 0.1 mM IPTG (Section 2.7.1). Cells were collected and loaded with 96 nM of acriflavine and 10 mM of CCCP. Cells were energised by the addition of 160 mM sodium formate at time zero (Section 2.13). *E. coli* BL21 (DE3) cells carrying the empty pBluescript SK II (pBS) vector were used as a negative control (shown in blue). Efflux experiments were performed in at least triplicate with a representative shown.

this study to assess whether it played a part in antimicrobial resistance mediated by SepA (Section 2.10.2). As can be seen in Figure 8.7 and 8.8, *E. coli* cells expressing SepA and variants thereof did not grow at concentrations of acriflavine and pyronin Y comparable to those expressing WT QacC. A comparison of the growth of cells expressing SepA and it variants, showed that only those expressing the SepAΔ1QacCTMS4 fusion protein were able to grow at higher concentrations of pyronin Y, at all pH levels examined, than those carrying the empty pBluescript SK II vector. Similar results were observed for acriflavine; however, higher growth rates of cells expressing SepAΔ1QacCTMS4 were observed at pH 7.0 only, with growth at pH 6.0 and 7.5 being comparable to the negative control. Analysis of growth at three different pH levels examined indicates that unlike the MdfA multidrug transporter, external pH does not appear to have an effect on SepA-mediated resistance against acriflavine or pyronin Y.

8.6.2 Plate dilution analysis

To further evaluate SepA resistance, the more sensitive agar plate-dilution method (Section 2.10.3), as described in Section 6.3.2, was used in this study. Growth of diluted cultures of SepA and derivatives, on solid media containing acriflavine (10 and 12.5 μg/ml) and pyronin Y (5 and 10 μg/ml) at pH 6.5 and 7.0 were compared to *E. coli* BL21 (DE3) cells carrying the empty pBluescript II SK vector (Figure 8.9). Cells were diluted 10 fold from 10^{-1} to 10^{-6} and spotted onto the selective MH plates; a control MH plate without compounds revealed all cells grew at all dilutions at pH 6.5. However, at pH 7.0 cells expressing SepAΔ1 displayed reduced growth at 10^{-5} and 10^{-6} dilutions, whilst all others grew to the final 10^{-6} dilution, indicating that at this elevated pH, even without selection, cells expressing this variant were compromised.

As shown in Figure 8.10, growth at higher dilutions was observed for all *E. coli* BL21 (DE3) cells at pH 6.5, when compared to those growing at pH 7.0, with the exception of cells expressing QacC which grew to the final dilution of 10^{-6} on all plates assessed. Although growing to the final dilution of 10^{-6} on 12.5 μg/ml acriflavine plates, growth of cells expressing both SepA and SepAΔ1QacCTMS4

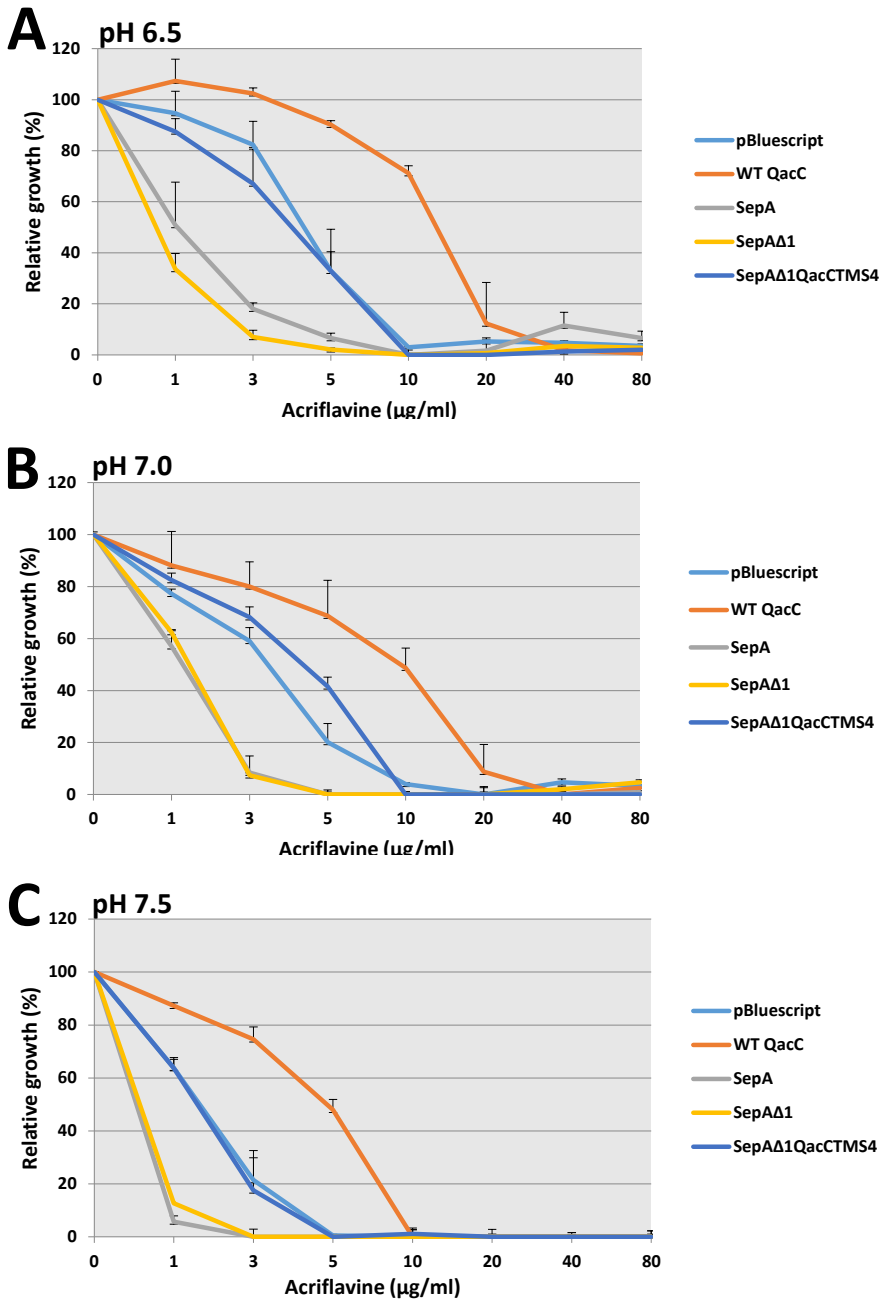

Figure 8.7: Acriflavine resistance of the SepA and variants expressed in *E. coli* BL21 (DE3) cells grown at pH 6.5, 7.0 and 7.5

Relative growth of *E. coli* BL21 (DE3) cells carrying pBluescript II SK-based SepA and variants in MH broth at (A) pH 6.5, (B) 7.0 and (C) 7.5, in increasing concentrations of acriflavine. Relative growth was calculated from the cell density, measured at OD_{600}. QacC was also assayed and included for comparison. Representative data shown with error bars indicate standard deviation from triplicate assays.

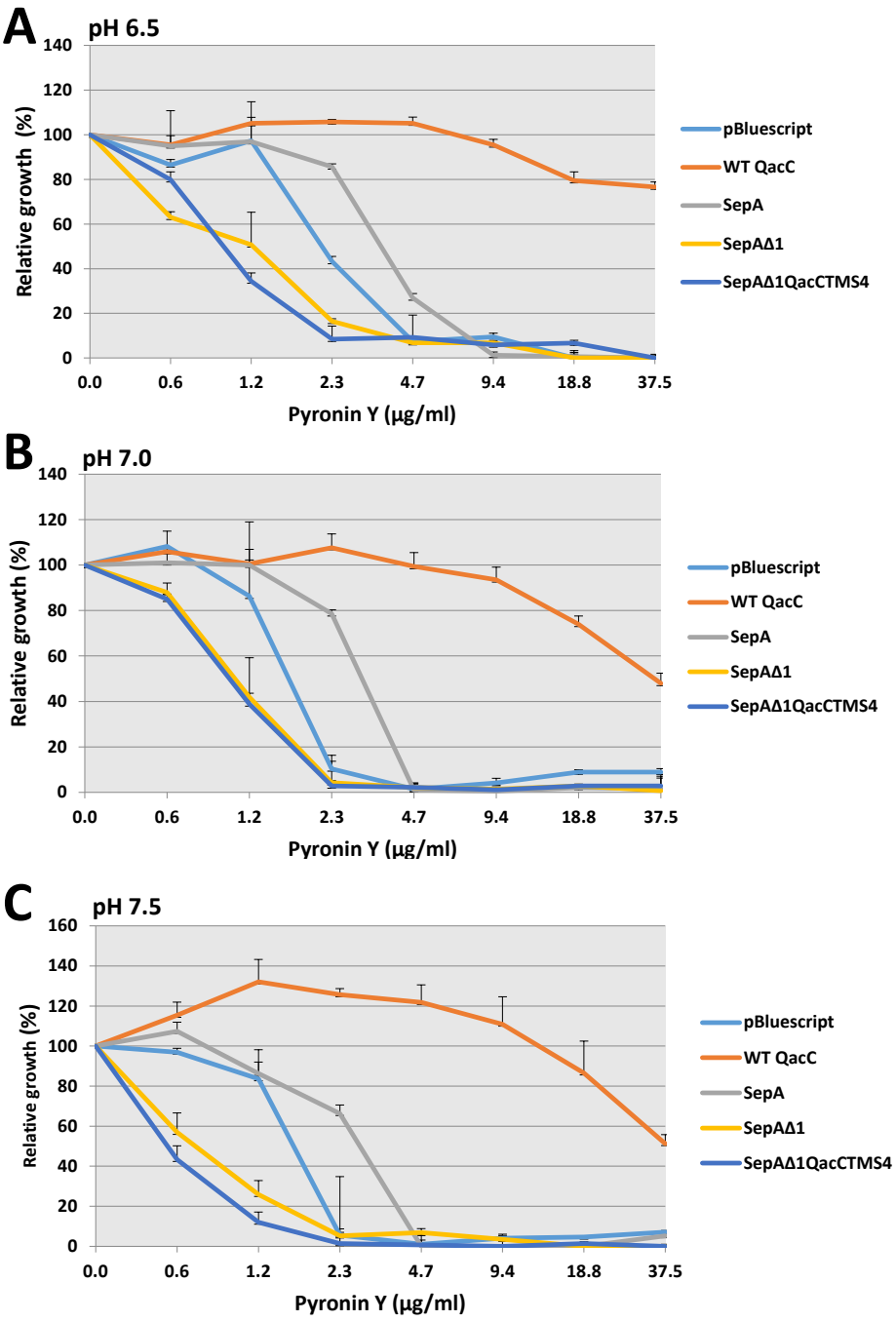

Figure 8.8: Pyronin Y resistance of the SepA and variants expressed in *E. coli* BL21 (DE3) cells grown at pH 6.5, 7.0 and 7.5

Relative growth of *E. coli* BL21 (DE3) cells carrying pBluescript II SK-based SepA and variants in MH broth at (A) pH 6.5, (B) pH 7.0 and (C) pH 7.5, in increasing concentrations of pyronin Y. Relative growth was calculated from the cell density, measured at OD_{600}. QacC was also assayed and included for comparison. Representative data shown with error bars indicating standard deviation from triplicate assays.

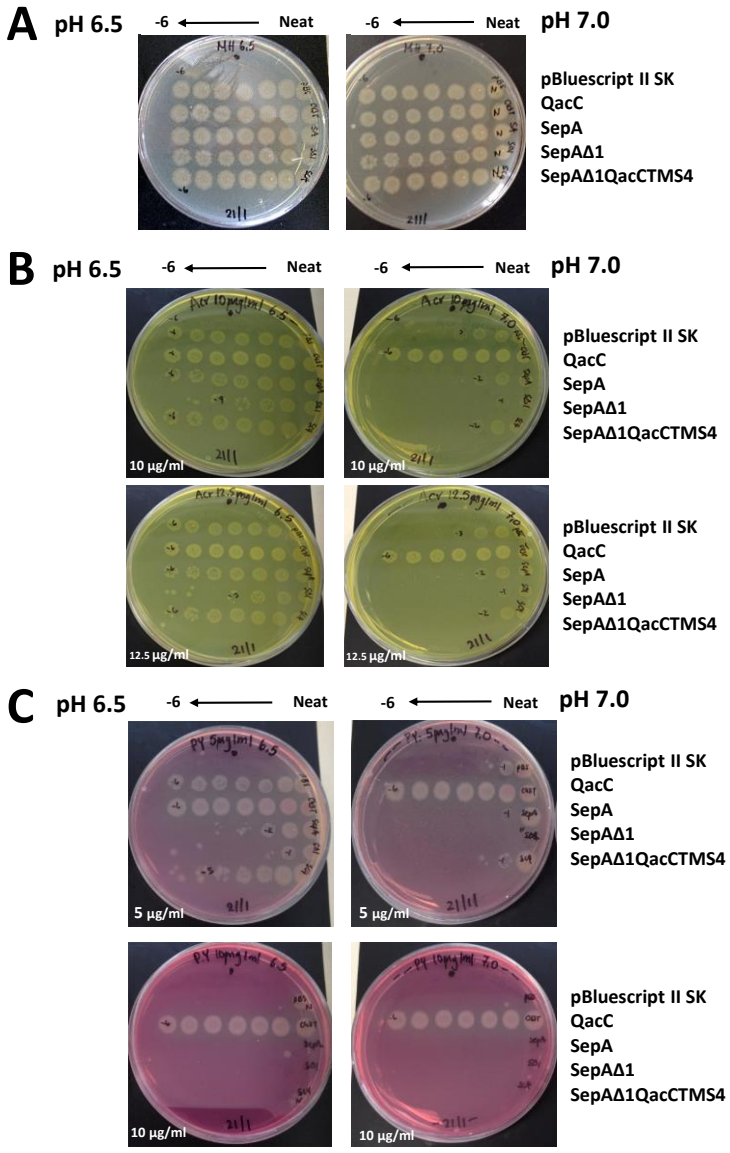

Figure 8.9: Growth of serial log_{10} dilutions *E. coli* BL21 (DE3) cells expressing SepA and variants on pH 6.5 and 7.0 MH agar plates supplemented with acriflavine or pyronin Y

Growth phenotype of *E. coli* BL21 (DE3) cells carrying pBluescript II SK-based SepA and variants. Cells were grown at 37°C in LBAmp, induced at OD_{600}=0.55 for one hour at 37°C with 0.1 mM IPTG (Section 2.7.2). After adjustment of cells to OD_{600}=0.6, cells were serially diluted and a 5 µl volume of neat to 10^{-6} dilutions of the culture was spotted on (A) control MH plates, (B) MH plates containing 10 or 12.5 µg/ml of acriflavine and (C) 5 or 10 µg/ml of pyronin Y, at pH 6.5 and 7.0. Growth was analysed after overnight incubation at 37°C. *E. coli* BL21 (DE3) cells carrying the pBluescript II SK-based QacC protein were included for comparison and served as a positive control.

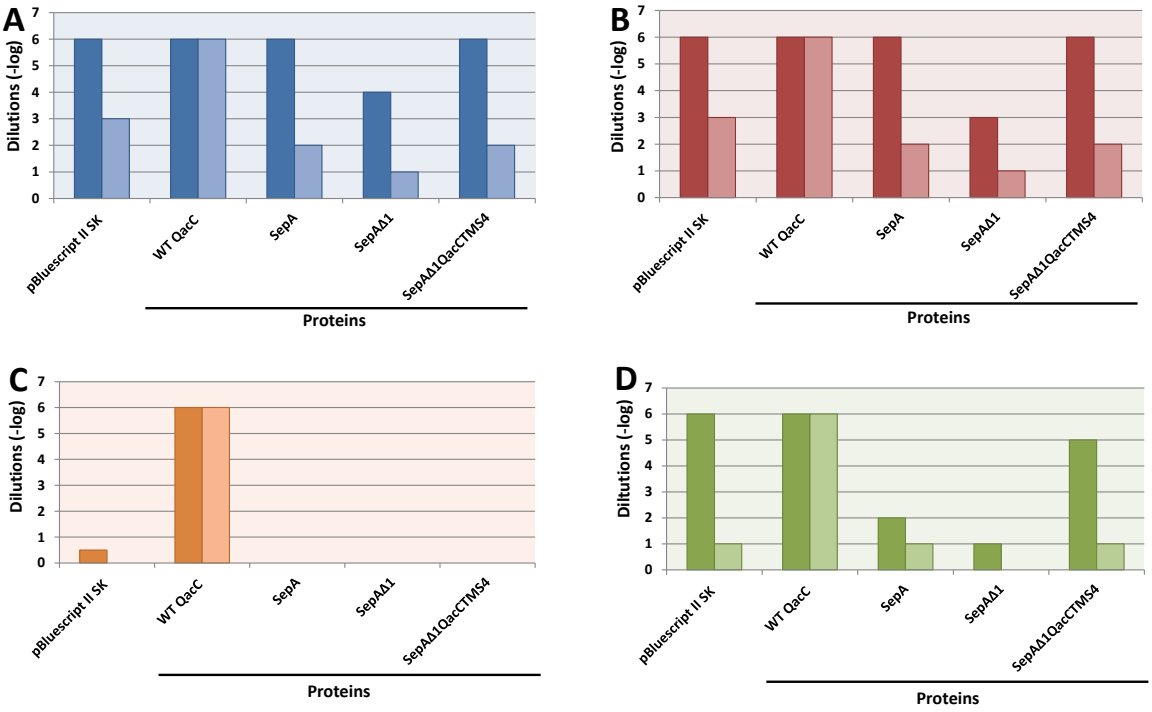

Figure 8.10: Resistance activity of SepA and variants against acriflavine and pyronin Y at pH 6.5 and 7.0

Cells were grown at 37°C in LBAmp, induced at OD_{600}=0.55 for one hourr at 37°C with 0.1 mM IPTG (Section 2.7.2). After adjustment of cells to OD_{600}=0.6, cells were serially diluted and a 5 µl volume of neat to 10^{-6} dilutions of the culture was spotted on (A and B) MH plates containing 10 or 12.5 µg/ml of acriflavine and (C and D) 5 or 10 µg/ml of pyronin Y, at pH 6.5 (darker shade) and 7.0 (lighter shade). Growth was analysed after overnight incubation at 37°C. *E. coli* BL21 (DE3) cells carrying the pBluescript II SK-based QacC protein were included for comparison and served as a positive control. Growth at all dilutions was observed on control plates for all cells (data not shown). Each bar represents the maximum dilution that growth was observed.

began to diminish at the 10^{-5} dilution. Instead of confluent growth, single colonies were visible for both SepA and SepAΔ1QacCTMS4 expressing cells (Figure 8.9), indicating lower levels of conferred resistance to cells expressing these two proteins. *E. coli* BL21 (DE3) cells expressing SepAΔ1 grew at lower dilutions of 10^{-4} and 10^{-1} on the 10 μg/ml acriflavine plates at both pH 6.5 and 7.0, respectively. On the 12.5 μg/ml acriflavine plates growth at pH 6.5 was observed to be at a lower dilution of 10^{-3}, indicating that this variant conferred the lowest levels of resistance.

Results observed from the pyronin Y plates displayed similar patterns, with growth at higher dilutions observed on the 5 μg/ml pyronin Y plate at pH 6.5 when compared to growth at pH 7.0. As seen in Figure 8.10, of the SepA variants assessed, cells expressing the SepAΔ1QacCTMS4 protein grew to a dilution of 10^{-5}, higher than both SepA and SepAΔ1. However, *E. coli* BL21 (DE3) cells carrying the empty pBluescript II SK vector grew to the higher dilution of 10^{-6}, indicating that this variant did not confer any significant resistance against pyronin Y at any pH assessed. Additionally, the pyronin Y concentration of 10 μg/ml appeared too high, as no growth was observed at any dilution for SepA and variants thereof.

In conclusion, although shown to be a more sensitive method of evaluating resistance, the plate-dilution method was still unable to identify any significant increase in the MIC for strains carrying either SepA or any of the SepA variants. However, as indicated in Figure 8.7 and discussed in Section 8.6.1, the lower pH concentration of 6.5 appears to be more favourable for growth of all *E. coli* BL21 (DE3) cells, including those expressing SepA and variants thereof. Nonetheless, lower pH levels do not affect the levels of resistance that are conferred to the expressing cells by those proteins.

8.7 Discussion

To date, very little is known regarding the *S. aureus* SepA membrane transport protein. A BLAST analysis of the *sepA* gene clearly identifies the genome sequence to be highly conserved in all currently sequenced *S. aureus* genomes. Such a high conservation implies functional importance of the *sepA* product. To further investigate its funtion, the SepA protein sequence was used in a BLASTp search;

however, results identified it simply as SepA, a *S. aureus* multidrug efflux protein, with no close homologues. A BLASTp search excluding staphylococcus revealed only three other possible hits, though these were shown to have very low identity scores and were identified as hypothetical proteins only. The current description of SepA as multidrug transporter comes from only one publication (Narui et al., 2002).

Sequence alignments (Figure 8.1 and 8.2), clearly show that SepA shares little homology with other members of the SMR family. Although SepA possesses certain residues which have been identified as essential in SMR members (Figure 1.9), the relative position of these residues does not compare to those found in other SMR efflux pumps. The first TMS does not contain the essential Glu14 residue found to be highly conserved in most members of the SMR family (Bay and Turner, 2009). Analysis of the SepA amino acid sequence revealed that most of the identified and experimentally verified essential residues found in members of the SMR family, can be identified in the last three TMS of SepA. However, SepA still appears to be missing a number of essential residues found for SMR proteins. These have been shown to be involved in function, substrate binding and translocation, as well as being part of a dimerisation motif.

Given its size and formation of four TMS, it was also assessed whether SepA could belong to the newest family of efflux proteins, the PCE family (Section 8.2). However, an amino acid sequence alignment of SepA and AceI, the best characterised member of the PCE family, revealed very low sequence homology. A multiple sequence alignment between SepA and other members of the PCE family revealed that Glu15, identified as an essential residue for transport in AceI, did not align to a Glu residue, or any other negatively charged residue within SepA. Alignment of Glu15 with Glu81 in SepA (Figure 8.3) did not result in an improved alignment between members of the PCE family and SepA. Assessment of this alignment and the alignment of SepA and members of the SMR family (Figure 8.1), revealed that SepA contained more essential resides that were identified in members of the SMR family than with those belonging to the PCE family.

Based on the alignment between SepA and other members of the SMR family (Figure 8.1), and the relative positions of the essential residues, a truncated SepA

construct was created, containing the last three TMS only (SepAΔ1). To assess whether the arrangement of the residues within the protein was necessary for function of a transporter, SepAΔ1 was then fused with the last TMS4 of QacC, giving rise to the SepAΔ1QacCTMS4 fusion protein. As can be seen in Figure 8.2, a sequence alignment of all SepA constructs and QacC, and the relative positions of the essential residues, indicated that this construct may be functional, as the addition of QacC TMS4 moves the essential SMR residues into the correct position within the polypeptide chain.

Following the construction of all SepA variants, levels of expression were analysed for each construct. High levels of expression of both SepA and SepAΔ1 were detected by Western blot, however, SepAΔ1QacCTMS4 expression levels were low in comparison (Figure 8.5). A Western blot on whole cell lysates and membrane fractions revealed similar low levels of SepAΔ1QacCTMS4 expression, indicating that this construct was incorporated into the membrane and not lost in inclusion bodies. It is uncertain as to why levels of expression for this construct were so low, possible reasons could include protein instability due to the fusion of two different proteins, leading to degradation by host proteases or protein aggregation due to an alternative non-native conformation of the fusion protein itself. Time permitting it would have been of interest to do a pulse-chase experiment, to follow the fate of SepAΔ1QacCTMS4 within the cell.

Following on from previous ethidium efflux studies (Narui et al., 2002), fluorimetric efflux assays were performed with both ethidium and acriflavine in order to verify its role as a drug transport protein. Although ethidium has shown to be a common substrate of many drug membrane transporters, including those within the SMR family, here, ethidium efflux was not observed by any SepA protein (Section 8.5). Given these results and those of the initial SepA study which showed that *E. coli* cells expressing SepA displayed elevated resistance levels against acriflavine (Narui et al., 2002), acriflavine was chosen to further examine the transport capacity of the SepA proteins.

As shown in Figure 8.6, upon the addition of sodium formate, an increase in fluorescent intensity was observed for SepA and both SepA variant proteins,

however, the level of efflux for all three proteins appeared to be low. Of the three constructs assessed, SepAΔ1 displayed the lowest levels of efflux, indicating that this variant is unable to effectively extrude acriflavine out of the cell. The results of the remaining SepA proteins are indicative of low acriflavine efflux only. Given these results and those obtained from fluorimetric transport assays using ethidium as a substrate, both acriflavine and ethidium do not appear to be suitable substrates to fully ascertain whether SepA is a drug efflux pump, and as such other compounds would need to be assessed to verify its transport capacity.

In addition to fluorimetric transport assays, resistance profiles of SepA and variants thereof were assessed (Section 8.6). Using two different techniques, analysis of SepA drug specificity, determined by assaying the MIC levels of various antimicrobials (Section 8.6.1), revealed that neither SepA nor variants thereof were able to confer resistance to any of the compounds tested. The possibility that the external pH may have an effect on resistance mediated by the SepA proteins was also evaluated (Section 8.6.1). Results, however, indicated that external pH levels did not affect the resistance against either acriflavine or ethidium (Figure 8.7 and 8.8).

The above results highlight the difficulties in assigning a possible function to unknown proteins. Although conserved throughout the staphylococcal genome, highlighting a possible importance of the protein, its lack of homology to any other known protein makes functional determination difficult. An alternative approach to predict a proteins possible function involves the use of computational analysis. However, computational prediction of a proteins possible function depends on comparative modelling, which allow for the construction of a 3D model of a protein of an unknown structure. Once this is achieved binding regions can be proposed and ligands that have a high probability of binding can be identified by using protein ligand docking servers, such as SwissDock (Grosdidier et al., 2011). Nonetheless, as mentioned, such predictions are dependent on comparative modelling, and ideally a crystal structure of a homologous protein must be available for such modelling to be carried out. In the case of SepA, given its lack of homology

with any other known protein, let alone one which has had its structure solved, makes even computational prediction of function extremely difficult.

Its vague resemblance to proteins belonging to the SMR family alludes to the possibility that it too may be a membrane drug transporter. However, results in this study indicate that although this may still be possible, as SepA could extrude a substrate not as yet tested for, it does not share the same drug profile as MPs belonging to the SMR family, as shown by both fluorimetric transport assays and MIC results.

Further analysis into SepA is needed to confirm its possible role within the Staphylococcus genus. Given its high conservation rate within the Staphylococcus genome (Section 8.2), future studies focussing on the construction of a *sepA* knockout may help to elucidate whether SepA is essential for *S. aureus* growth and/or conferred resistance. In addition, as the identification of a possible substrate profile for a given protein of unknown function can be time consuming, high-throughput experiments in which large amounts of data can be generated and then analysed would need to be carried out on SepA in order to be able to fully assess its possible function. One possible approach to assess SepA may be through the use of protein chips (MacBeath and Schreiber, 2000; Zhu et al., 2000; Zhu and Snyder, 2001), which were used to analyse yeast protein kinases from *Saccharomyces cerevisiae* (Zhu et al., 2000). This method allows for the high-throughput analysis of biochemical activities of proteins, using only small amounts of proteins, and as such has potential for SepA screening. Alternatively, in order to examine the potential substrate range of SepA the use of Biolog on a *sepA* knockout could also be carried out.

8.8 Conclusion

Despite initial reports purporting SepA to be a multidrug efflux pump belonging to the SMR family, analysis carried out in this study failed to reproduce the findings stated in the Narui et al., (2002) study. It is possible that SepA is a multidrug protein that has a resistance profile that has yet to be elucidated or that its activity is restricted to Gram-positive organisms such as *S. aureus*. Given the recent

determination of a new family of membrane transporters, the PCE family, it is also likely that SepA belongs to a novel, perhaps a yet to be discovered drug membrane transporter family. Nonetheless, it is obvious that more research needs to be carried out to determine the possible function of SepA and to ascertain if this protein is in fact a multidrug transporter.

CHAPTER 9 - DISCUSSION

Identified to be a major cause of both hospital and community acquired infections, *S. aureus* is a highly adaptive opportunistic bacterium that displays a wide spectrum of pathogenicity, which it owes largely to its ability to acquire mobile genetic elements that encode virulence and resistance determinants (Alibayov et al., 2014; Lindsay and Holden, 2004). This adaptive power has given rise to the emergence of MRSA, and continues to contribute to the development of new strains which have been shown to be resistant against vancomycin, a last resort antibiotic (Hiramatsu et al., 1997; Limbago et al., 2014). Studies have identified a number of determinants which contribute to the pathogenicity of *S. aureus*, however, they have not identified any potential genes or gene combinations which could be targeted by new therapies (Lindsay and Holden, 2004; Saravolatz et al., 2012). Whole genome sequencing of *S. aureus* and comparative genomics aimed at trying to develop new therapies against this bacterium, have shown that approximately 22% of the *S. aureus* genome is composed of variable regions (Fitzgerald et al., 2001), which were shown to correspond to horizontally transferred elements (Lindsay and Holden, 2004). Amongst these elements are plasmids such as pTZ20, which carries the *qacC* gene encoding the QacC multidrug transporter (Sasatsu et al., 1989), allowing *S. aureus* to become more resistant against biocides and other antimicrobials.

In addition to QacC, a number of multidrug transport proteins have been identified in *S. aureus.* These include the chromosomally encoded transport proteins, LmrS (Floyd et al., 2010), NorA (Neyfakh et al., 1993; Yoshida et al., 1990), NorB (Truong-Bolduc et al., 2005), NorC (Truong-Bolduc et al., 2006), SdrM (Yamada et al., 2006a), MepA (Kaatz et al., 2005) (Section 1.3.4.2) and Sav1866 (Dawson and Locher, 2006) (Section 1.3.1.1) and the plasmid encoded QacA (Tennent et al., 1985), QacB, (Paulsen et al., 1996a) (Section 1.3.2.3) QacJ (Bjorland et al., 2003) and QacG transporters (Wong et al., 2013) (Section 1.4.4). As shown above, the number of drug efflux pumps found within *S. aureus* is abundant, raising the question as why so many drug transporters are needed within this bacterium. In contrast to Gram-negative bacteria, Gram-positive bacteria like *S. aureus* lack an

outer membrane, making them more susceptible to an attack. As such, drug efflux pumps may serve as a first line of defence before other, more established mechanisms can be developed, allowing for survival in a hostile environment (Costa et al., 2013). Given its classification as an opportunistic nosocomial pathogen (Archer, 1998; Tong et al., 2015), its long term exposure to hospital biocides may also have significantly contributed to its acquisition of numerous membrane drug transport proteins.

Working synergistically with other resistance mechanisms, efflux pumps have been shown to play a major role in antimicrobial resistance. With their wide substrate specificity and ability to produce elevated levels of resistance when overexpressed, they contribute to the high level of resistance exhibited by *S. aureus.* As such, they have become targets for the development of inhibitors which, when successful would restore the activity of the antimicrobial (Li and Nikaido, 2004). Although inhibition studies have led to the development of compounds capable of inhibiting such pumps, for example inhibition of the Tet efflux pumps (Gibbons and Udo, 2000; Hirata et al., 1997; Nelson and Levy, 1999), the development of inhibitors is challenging and complicated by the need to combine them with antimicrobial agents that show similar pharmacokinetic profiles (Li and Nikaido, 2004).

The key to the development of inhibitors lies in the understanding of the molecular mechanism of multidrug transporters. Nonetheless, such an understanding, especially regarding transporters within the SMR family, is still limited as much of the knowledge gained thus far on SMR proteins is based predominantly on the *E. coli* EmrE membrane transporter (Section 1.5). Due to its size, stability and retention of function when solubilised in detergent, EmrE is considered to be a unique archetype for structural studies (Schuldiner, 2007; Schuldiner et al., 2001a), and thus has been extensively analysed (Lloris-Garcera et al., 2012; Schuldiner, 2009; Yerushalmi et al., 1995). In addition to biochemical analysis, four EmrE structures have thus far been solved and submitted to the PDB archive (http://www.rcsb.org) (Chen et al., 2007; Fleishman et al., 2006). This makes findings observed for EmrE the only results by which functional and

structural results obtained for other SMR transporters can be compared to. Nowhere is this limitation more apparent than when trying to predict a 3D structure of a protein such as QacC when using homology modelling (Section 3.4). As EmrE is the only QacC homologue with a solved structure, all predictive modelling of QacC was based on this one protein, generating results which may not necessarily be illustrative of the final 3D QacC structure. As seen in this study, it is apparent that although sharing homology, differences also exist, as exemplified by the GG7 dimerisation motif which appears to be more defined for EmrE when compared to QacC (Section 7.3.3). Such differences are demonstrative of the fact that although EmrE is a representative member of the SMR family, it does not fully exemplify all SMR multidrug transporters.

What is obvious from the lack of solved SMR MP structures is that obtaining high resolution crystal structures is challenging. This was demonstrated by attempts to conduct crystallographic and structural studies on the QacC protein (Dunham et al., 2001). Using the QacC construct utilised in this study, small crystals were grown, however, these were shown not to diffract well and as such could not be used for further assessment. Problems associated with structural studies such as crystallisation, are due to the inherent nature of MPs, which include their high hydrophobicity and difficulties in obtaining sufficient quantities of protein, as was discussed in Chapter 4. Overexpression of proteins such as QacC can lead to cell toxicity, as was seen following the induction of QacC from the pBAD vector in *E. coli* TOP10 cells (Section 4.3). As such, the need to determine a good overexpression system was evaluated in this study, where it was found that expression from the T7 promoter in the pBluescript II SK expression system in *E. coli* BL21 (DE3) cells was more effective and allowed cell growth even after induction (Section 4.4) compared to when cloned and expressed from the pBAD vector (Section 4.3). Although obtaining higher yields of QacC in the pBluescript II SK system and utilising Pro-Bond™ Ni-resin for purification resulted in a level of protein, which while sufficient for small scale analysis of QacC like Cys–scanning mutagenesis (Chapter 5), it was not adequate for analysis requiring large amounts of purified QacC, such as for fluorescence polarisation-based binding assays (Section 6.4.2). Nonetheless, levels

of protein expression and purification obtained in this study were sufficient for Cys–scanning solvent–accessibility studies, fluorimetric transport assays (Section 6.2) and MIC analyses (Section 6.3), which were carried out on QacC and the large number of mutants constructed in this study.

9.1 Structural modelling of the QacC transporter

Due to the above mentioned difficulties in overexpression and purification of MPs like QacC, which make crystallisation studies not practical, computer modelling was used as a viable alternative approach for structural analysis. Such analysis can facilitate the characterisation of a protein's molecular mechanism, and further our understanding of ligand binding, all of which may aid in the development of inhibitors or other pharmacological agents targeting QacC (Hopf et al., 2012). As such it was used in this study to evaluate both the secondary and tertiary structure of QacC.

Two approaches, the *de novo/ab initio* and comparative modelling, were undertaken for predictive structural analysis of QacC (Chapter 3). Given the lack of structural information available for MPs, computational structure prediction is still predominantly based on knowledge obtained for water-soluble proteins, which may not accurately depict the structural constraints applicable to MPs. As a result of this, work has gone into the refinement of these approaches, with new algorithms being written which are specific for MPs (Hopf et al., 2012; Punta et al., 2007).

Computational analysis of the secondary structure of QacC revealed that QacC consists of four α-helices (Section 3.2). Some ambiguity regarding the orientation of the N–terminal region was observed, however, given that SMR proteins like EmrE have been found to function as dimers, with monomers taking an antiparallel orientation in the membrane (Rapp et al., 2006), this was not unexpected. It has been postulated that topology of MPs is influenced to a great extent by the "positive–inside" rule which stipulates that positively charged amino acids tend to be located in the cytoplasmic part of the protein (Heijne, 1986; Rapp et al., 2007). However, like EmrE, QacC lacks such topological determinants as it displays a weak

Lys and Arg bias (Section 5.5, Figure 5.13). This weak bias, which has been identified in “dual–topology” proteins (Rapp et al., 2006), may not only explain the ambiguity observed regarding the possible location of the N–terminus, but could also explain fluorescein-5-maleimide and eosin-5-maleimide labelling of QacC mutants, which indicated that QacC monomers insert in both orientations within the membrane (Section 5.3). Such an arrangement would allow Glu14, identified to be at the centre of the substrate binding and translocation pathway in both QacC (Grinius and Goldberg, 1994) and EmrE (Yerushalmi and Schuldiner, 2000), to be in close proximity to itself (in a dimer) allowing for the formation of the binding site. This structural arrangement and possible orientation of Glu14 was indicated in the MODELLER predicted QacC model, which for comparison is shown aligned with EmrE (Section 3.6, Figure 3.13). This model illustrates and supports both the orientation of the monomers within the membrane as well as the placement of each TMS within the dimer, allowing for the formation of a substrate binding pocket and translocation pathway. Fluorescein-5-maleimide accessibility studies (Section 5.4) further evaluated this proposed translocation pathway, with results indicating that the Cys introduced into the QacC Glu14Cys derivative is accessible to fluorescein-5-maleimide, further supporting its placement within a possible hydrophilic substrate translocation site. Moreover, functional assays of cells expressing this mutant (Chapter 5) revealed that a substitution of Glu14 with Cys abolished efflux of ethidium and resistance against ethidium, acriflavine and pyronin Y. All together these results confirm the role of Glu14 as essential for function, and support its predicted location and position orientation within the QacC dimer.

9.2 QacC Cys–scanning solvent–accessibility studies

One of the key goals in this study was to evaluate the topology of QacC and to define its possible secondary structure. As discussed above, computational predictive methods proposed that QacC has four TMS, which was further supported by Cys-scanning mutagenesis and fluorescein-5-maleimide labelling (Section 5.3.1). Using the Cys-scanning solvent-accessibility method allowed for the sequential

mapping of all QacC residues within the protein, making it possible to form a topological map which allocated each residue to the membrane or loop region within QacC (Figure 5.18). As indicated in Chapter 5, an overlap of the allocation of residues within QacC by fluorescein-5-maleimide and by predictive methods such as TMHMM is visible, indicating that the proposed model and biochemically assessed secondary structure of QacC is representational of the native structure of QacC. Lastly, conformational consensus is also seen between TMS assigned residues determined in this study and those by the NMR study (Poget et al., 2010), with slight variations seen to the allocation of residues to TMS1 and 2. However, as stated in the NMR analysis, an underrepresentation was seen for TMS3 and TMS4, as only 11 amino acids were assigned to TMS3 and 9 to TMS4. Studies have shown that the number of residues needed to make up an α-helix that is able to span the 30 Å hydrophobic core of the membrane is ~20 (Fleishman and Ben-Tal, 2006; White and Wimley, 1999). However, the length of the transmembrane α-helix can vary between 12 and 40 residues (De Marothy and Elofsson, 2015). Nonetheless, as the number of residues allocated to TMS3 and TMS4 are so low, further assessment would need to be carried out on these segments.

Cys-scanning solvent-accessibility has been used extensively to evaluate both structure and function of a number of membrane transporters (Hassan et al., 2007b; Hruz and Mueckler, 1999; Loo and Clarke, 1999; Wilson et al., 2014), and is ideal for studies of small proteins such as QacC, where the systemic replacement of each residue with Cys can be carried out on the entire protein. Utilisation of Cys has become very attractive for site–specific mutation due its rarity in proteins as well the ease at which it can be introduced into a specific site of the protein without perturbing the function of the protein (Kim et al., 2008). In this study, with the aid of the thiol-reactive fluorescein-5-maleimide reagent, all QacC residues were evaluated with respect to their accessibility to this membrane–impermeant compound. Results of this study allowed for the formation of a detailed representative map of all QacC residues and their relative position within the QacC protein (Figure 9.1). Residues shown to be reactive with fluorescein-5-maleimide

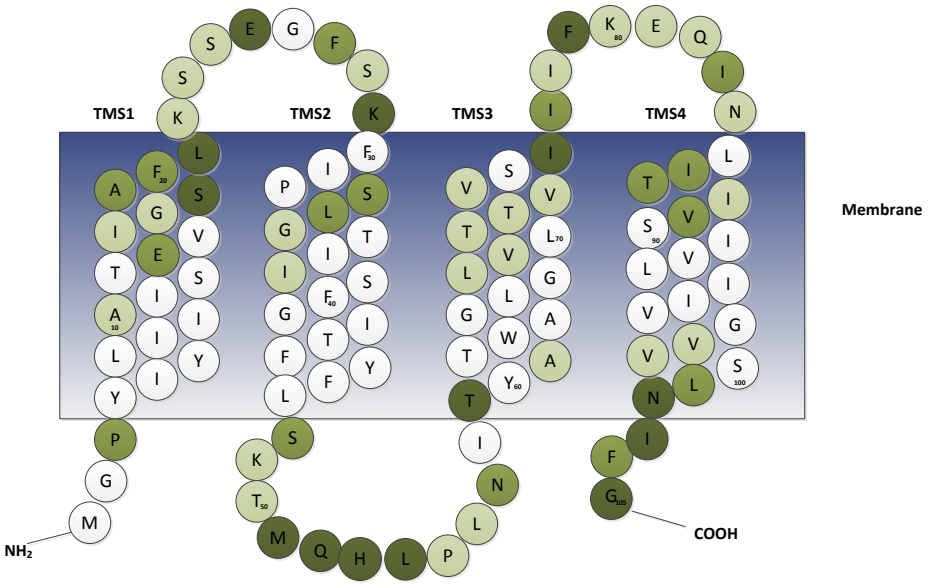

Figure 9.1: Accessibility of fluorescein-5-maleimide to QacC–Cys mutants

A final composite model of the secondary structure of QacC showing the putative positions of all Cys–replaced QacC residues showing low reactivity with fluorescein-5-maleimide (+ shaded in light green), medium (++ shaded in green) and high reactivity (+++ shaded in dark green).

were thus determined to lie outside of the membrane. Those found reactive, but putatively assigned to reside at the boundary of the membrane, were thought to represent partially buried residues whose side chain was still accessible to fluorescein-5-maleimide. Lastly, accessible residues within the membrane represent those thought to be part of the substrate translocation pathway.

As previously discussed, QacC Glu14Cys was shown to be accessible to fluorescein-5-maleimide (Section 5.3.1) and blocked by pre-incubation with pyronin Y prior to labelling (Section 5.4), supporting its possible role in substrate binding and its location in the translocation pathway. Moreover, results indicate that reactive residues Ser18Cys to Leu21Cys found in TMS1 may also form part of this pathway, although not be directly involved in substrate binding. Other reactive residues, located at the boundary of the bacterial membrane include Lys21 (end of TMS2), Thr59 (beginning of TMS3), Asn84 (beginning of TMS4) and Asn102 (end of TMS4). Non-reactive residues allocated to QacC loop regions, may indicate the tight packaging of QacC restricting accessibility to these residues. Nonetheless, the results obtained give a possible glimpse of residue packaging within QacC, as expressed and inserted into the *E. coli* membrane. Future comparison studies analysing the structure of QacC in the *S. aureus* membrane would complement these studies, especially in the light of recent findings which show that the membrane environment plays a key role in protein topology and dimerisation (Dutta et al., 2014; Vitrac et al., 2013).

Given the proposition that SMR proteins function as homodimers, accessibility studies were also carried out on whole cells expressing QacC (Section 5.3.1). Results obtained in this study support findings that like EmrE, QacC monomers appear to insert in an antiparallel orientation, as QacC–Cys residues putatively located to loop regions on both sides of the membrane were all accessible to fluorescein-5-maleimide. Such reactivity could occur only if half/some of the monomers were inserted in the opposite orientation, allowing for each loop located residue to be exposed to the reagent. Due to its bulkier nature and inability to pass through the membrane, eosin-5-maleimide was chosen to represent a more stringent evaluation of monomer orientation. However, accessibility studies with eosin-5-

maleimide were hampered by its low fluorescence levels, making determination of reactivity difficult to ascertain. Therefore, future studies focussing on whole cells may need to include bacterial strains with a leaky outer membrane which will allow for the entry of bulkier reagents such as eosin-5-maleimide. Alternatively, other reagents such as methoxypolyethylene glycol maleimide (MAL-PEG), a membrane impermeable maleimide which has a mass of approximately 5,000 Da, could also be used to determine surface exposed Cys residues. Unlike using a fluorescent probe such as fluoresceine-5-maleimde, the binding of this reagent to exposed Cys residues can be assessed using a gel-shift assay, where labelled proteins can then be detected by Western blotting. This method has been successfully utilised in studies investigating the topology of the *E. coli* twin-arginine transport proteins TatA and TatB in intact cells (Koch et al., 2012).

In addition to using different reagents, accessibility studies using liposomes could also be used to assess the topology of QacC, eliminating the need for reagents to cross the *E. coli* outer membrane. The outer membrane, although more permeable than the inner membrane, may have contributed to the low reactivity levels observed with eosin-5-maleimide, as the outer membrane may have inhibited access of eosin-5-maleimide into the periplasmic space, where it could react with exposed Cys residues.

9.3 QacC oligomerisation

Although initially thought to be the perfect system in which multidrug efflux could be evaluated, studies of EmrE, in particular those assessing its oligomeric structure, uncovered that it forms a dimer composed of two monomers that take on an antiparallel orientation (Chen et al., 2007; Fleishman et al., 2006). This discovery initiated an ongoing debate regarding the possibility of such an orientation (Section 1.5), especially in light of biochemical evidence which seemed to indicate that EmrE monomers take on a symmetrical (parallel) orientation (Schuldiner et al., 2001a; Soskine et al., 2006; Soskine et al., 2002). Nonetheless, EmrE has become synonymous with such terms as “dual-topology” (Rapp et al., 2006; Seppala et al., 2010) and “topoforms” (Ong et al., 2013; Schuldiner, 2012)

indicating that although initially thought to be highly unlikely, especially when considering the mechanism behind the insertion and opposite orientation of each monomer of a membrane within a protein dimer, the notion that EmrE takes on this unique structure has become accepted and supported by numerous studies (discussed in Section 1.5.1).

Similar to EmrE, it was postulated that QacC also functions as an oligomer. As such, preliminary work was carried out in this study to evaluate the possible oligomeric structure of QacC. Two different approaches were undertaken to assess oligomerisation, formaldehyde cross-linking (Section 7.2) and BN-PAGE analysis (Section 7.3.4). Formaldehyde cross-linking has successfully been used to assess protein-protein interactions and has been found to be an ideal cross-linking reagent as it has a short cross-linking span, is able to permeate cell membranes and formed cross-links that are reversible upon heating (Nadeau and Carlson, 2007). It has been used to evaluated the interaction of Wzz_{SF} (Daniels and Morona, 1999; Papadopoulos and Morona, 2010) and TonB, a cytoplasmic MP (Skare et al., 1993) and has been shown to be an effective method to assess protein-protein interactions, as such it was used in this study. In addition to formaldehyde cross-linking, QacC oligomerisation was evaluated with BN-PAGE, a commonly employed method used to observe oligomerisation. BN-PAGE has been extensively used in the oligomeric assessment of a large number of MPs, including the lactose transporter LacS (Heuberger et al., 2002) and EmrE (Lloris-Garcera et al., 2012).

To date, no experimental data pertaining to the oligomeric state of QacC has been reported. Initial findings obtained from the studies described in this thesis indicate that like EmrE, QacC functions as a dimer. Cross-linking analysis with formaldehyde on *E. coli* TOP10 cells expressing QacC suggests that QacC can form dimeric complexes (Section 7.2, Figure 7.1). Although it has been proposed that SMR proteins may form higher oligomeric structures such as tetramers (Ubarretxena-Belandia and Tate, 2004), formaldehyde cross-linking failed to identify QacC-QacC interactions which were higher than dimeric in form.

In accordance with the above mentioned results, QacC oligomerisation was also observed by BN-PAGE analysis (Section 7.3.4). However, limitations pertaining to

the amount of lipids, detergent and Coomassie Brilliant Blue dye R-250 which bind to the protein, contributed to the difficulty in determining the mass and migration behaviour of QacC. When purified and analysed by BN-PAGE, two QacC protein bands at approximately 66 kDa and 146 kDa were observed. These molecular masses were different from the predicted masses of the monomeric and dimeric forms of QacC based on its amino acid composition, which are approximately 15 kDa and 30 kDa, respectively. Taking into account the presence of both DDM and Coomassie Brilliant Blue dye R-250, and using the conversion factor of 1.8, proposed for MPs solubilised in detergent (Heuberger et al., 2002), the MW of monomeric QacC determined by BN-PAGE would thus be 36 kDa and dimeric 80 kDa, which still does not correspond to the masses seen in BN-PAGE. It is possible that the above stated conversion rate of 1.8, which was established on studies using larger proteins such as LacS, which has a MW of 71 kDa and the xyloside transporter XylP from *Lactobacillus pentosus* with a MW of 91.5 kDa, is possibly not applicable to small membrane transporters such as QacC. In addition to the added mass conferred by bound lipids, detergent and dye which make protein mass estimation difficult, the use of soluble proteins as markers, which are not affected by the aforementioned conditions, may also contribute to errors in mass estimation of small proteins. As such, the identification of the two QacC bands observed by BN-PAGE, whether they represent monomeric and dimeric form of QacC, or higher oligomeric structures is uncertain. Given the results obtained by formaldehyde cross-linking, which indicate that the highest oligomeric form of QacC appears to be dimeric, it is highly probable that the two BN-PAGE QacC bands represent monomeric and dimeric forms of QacC, however, further analysis using analytical ultracentrifugation or size-exclusion chromatography, amongst others, is needed in order to confirm these results.

In addition to the assessment of QacC oligomerisation, the proposed SMR dimerisation motif ^{90}GLXLIXXGV^{98}, identified in TMS4 of SMR proteins such as Hsmr (Poulsen et al., 2009) and the GG7 motif identified in EmrE (Elbaz et al., 2008) was analysed in this study by site-directed mutagenesis and formaldehyde cross-linking (Section 7.3.3). Following a Cys-substitution of each of the eight residues it was

observed that within the ^{90}GLXLIXXGV^{98} motif, which in QacC is ^{90}SILIGVIG^{97} only Il94 and Gly97 are crucial for QacC function. Both Ile94Cys and Gly97Cys QacC mutants were unable to efflux ethidium or confer resistance against ethidium, acriflavine or pyronin Y (Table 6.2). Further assessment of these residues, in particular Ser90 and Gly97, revealed that although a Cys–substitution is permissible for Ser90, a replacement of this residue with a larger residue such as Ile significantly disrupts QacC function (Figure 7.5 and 7.6). Interestingly, although shown to be disruptive for QacC function when applied to Gly97, a substitution of Ser90 with another small residue such as Cys and Gly did not impair QacC function, possibly indicating that any small residue at position 90 is permissible for QacC function and dimerisation. Thus, although the GG7 motif, which in this study is identified as the SmG7 motif (Sm for any small residue), appears to be essential for QacC function. Nonetheless, results from formaldehyde cross-linking used to assess dimerisation did not indicate a disruption to dimerisation when QacC residues Ser90 and Gly97 were substituted with Cys, Gly or Ile (Figure 7.7). However, dimerization disruption may be limited to both Ser90 and Gly97 being replaced with larger residues, as a Cys replacement, given its small nature, may still allow for dimerization to occur. Clearly further analysis into the proposed dimerisation motif would need to be carried out to fully ascertain the role of these residues in dimerisation.

Oligomerisation of small membrane transporters such as QacC and EmrE continues to raise interesting questions regarding their evolution, more specifically why, given their unique homodimeric organisation, have these proteins not fused into larger membrane transporters? Investigation into larger MPs has revealed that many of them have evolved by internal gene duplications (Reddy et al., 2012; Shimizu et al., 2004), as discussed in Section 1.3.2. Within the DUF606 family, which contains proteins that also adopt dual topologies, internally duplicated proteins, whose first and second halves take on opposite orientations, have been identified (Rapp et al., 2006). However, proteins belonging to the SMR family appear to be “evolutionarily stuck” in a dimeric state, with no fused, internally duplicated proteins found to date (Lloris-Garcera et al., 2014). Possible explanations, based on

studies on genetically fused EmrE proteins, indicate that although functional, fused EmrE proteins display varying degrees of activity (Steiner-Mordoch et al., 2008). Thus, a fusion event may result in proteins which are less active (Lloris-Garcera et al., 2014), possibly explaining why such fusion events have not been observed for SMR proteins like QacC or EmrE. Nonetheless, given the existence of PSMRs, discussed in Section 1.4.3.2, it is highly possible that transporters such as QacC and EmrE function as oligomers with an antiparallel monomeric orientation. Members of both sub-families have been shown to be small, consisting of around 100 to 120 amino acids and are predicted to form four TMS. Their substrate profile overlaps, and both have a highly conserved Glu residue located in TMS1 shown to be essential for protein function. However, where SMRs like EmrE are purported to form homo-oligomers and are not thought to possess topological constraints which determine the orientation of each monomer within the membrane, PSMR protein function is reliant on hetero-oligomerisation, with the conformation of each monomer within the membrane possibly determined by the existence of longer hydrophilic loops found in one monomer and a hydrophilic C-terminus extension in the other (Kikukawa et al., 2006). In PSMR proteins, the existence of these hydrophilic regions appears to constrain these proteins and may lock them into forming antiparallel asymmetric dimers. Whether these constraints are needed for the correct orientation within the membrane for all SMR proteins is unknown, however, the existence of PSMR proteins may indicate that small proteins such as those belonging to the SMR family, must oligomerise in order to function, and that the orientation of each monomer within the membrane may need to take on the antiparallel orientation in order to for substrates to be bound and transported out of the cell. With current knowledge pertaining to PSMR proteins being limited, especially regarding their structure and topology, it is difficult to draw comparisons between members of the different subfamilies.

The limitations of our understanding of PSMR proteins and other members of the SMR family, is compounded by the fact that more studies are carried out on EmrE than any other SMR protein, limiting the potential for comparative analyses. It is clear, however, that studies on EmrE have provided us with surprising insight

into both function and structure, and with the relative rarity of "dual–topology" proteins, it is obvious that further analysis of SMR protein oligomerisation and topology is greatly needed in order to determine if EmrE is truly representational of all members of the SMR family, or if these proteins are unique and adopt dual–topology. As such, studies on QacC oligomerisation, as presented here, are necessary in order to fully understand this process and to ascertain if other SMR proteins follow the same fate as EmrE does, and take on an antiparallel configuration. To elucidate these further, studies using solid-state and liquid state NMR spectroscopy, Förster resonance energy transfer and electron paramagnetic resonance spectroscopy may need to be undertaken as these could offer more insight into oligomerisation states and the overall dynamics of these transporters.

9.4 Functional assessment of residues within QacC

In addition to solvent accessibility studies, Cys–scanning mutagenesis was used to assess and identify amino acids important for QacC function (Figure 6.1). Of these, a number belonged to conserved motifs which are specific to the SMR family (Section 6.5, Table 6.2). In total, of the 14 QacC–Cys mutants found to be unable to mediate ethidium efflux or confer any resistance against any of the compounds tested, seven are located within the SMR specific amino acid motifs. One of these residues, as previously mentioned, is Glu14, the only negatively charged residue located with the membrane which is highly conserved in the SMR family (Ninio et al., 2001). It has been identified in EmrE to be essential for substrate biding and translocation (Gutman et al., 2003). In addition to this, Glu14 has been shown to have an unusually high pK_a of 8.5, which appears to be essential for antiporter activity, whereby a hydrogen ion is exchanged for the cationic substrate (Soskine et al., 2004). Further studies determined that the environment that allows for such a high pK_a stems from an interaction with three aromatic residues found in EmrE (Tyr40, Tyr60 and Trp63), which in the absence of a substrate appear to stabilise the carboxyls of Glu14 that form the binding site. This stabilisation occurs as a result of an interaction with protons, or with parts of the six aromatic residues, thus three from each monomer making up the EmrE dimer (Adam et al., 2007). In

addition to stabilisation, recent studies on EmrE have shown that these aromatic residues and their proximity to the substrate binding site may also serve to accommodate non-polar parts of the substrate, whilst Glu14 binds to the positively charged fragment of the substrate (Padariya et al., 2015). Studies described in this thesis (Section 6.3) and elsewhere (Paulsen et al., 1995), have also demonstrated the importance of these aromatic residues, which in QacC are Phe40, Tyr60 and Trp63, for QacC activity. As shown in this study, substitutions of these three residues with Cys resulted in a significant impairment of the function of QacC, whilst substitutions of Tyr60 with Trp and Trp63 with Phe and Tyr (Paulsen et al., 1995), although conserving the aromatic side chain at these positions, resulted in significantly reduced resistance levels against a subset of dyes like ethidium and crystal violet, as well as against a number of QACs such as benzalkonimum and dequalinium.

Another QacC aromatic residue that when replaced with Cys abolished ethidium efflux and significantly reduced resistance levels against all three compounds tested is Tyr4, located in TMS1. Belonging to a conserved motif, studies in EmrE show that when mutated to Trp or Phe, it was still able to act as a drug transporter; however, replacement with Cys generated an uncoupled protein, able to bind a substrate or proton, but unable to confer resistance (Rotem et al., 2006). These results show that the need for an aromatic residue in position 4 of EmrE is vital for proper coupling, and explains why aromatic residues are so highly conserved in other SMR proteins (Figure 6.12). Correspondingly, the QacC Tyr4Cys mutant is unable to mediate ethidium efflux (Figure 6.1), or able to confer resistance to the compounds tested (Figure 6.4).

Looking at the distribution of QacC residues which were found to be significant for QacC function (Table 6.2), it is interesting to note that out of 14 residues which did not mediate ethidium efflux and conferred no resistance against any of the compounds tested, 11 are putatively located to TMSs. Of these, seven are highly conserved and an additional three showed strong similarity between residue groups when aligned with other SMR proteins (Figure 1.9). In contrast to exposed residues, which are not exposed to significant steric constraints and have been

shown to be evolutionarily variable, buried residues are commonly seen to be highly conserved. Typically hydrophobic, buried residues are thought to be under steric constrains which limit variation, and as such are more conserved (Donnelly et al., 1993; Fleishman and Ben-Tal, 2006; Hurwitz et al., 2006). These residues may thus be contributing to the structural integrity of QacC, with a replacement to Cys disrupting this stability leading to reduced function. This may be the case with residues such as Gly17, Pro32, Gly65 and Gly67, which as discussed in Section 6.5, may play a role in protein stability. Of particular interest is Gly17, which when replaced with Cys appears to completely destabilise QacC, as indicated by the inability of the QacC Gly17Cys mutant to efflux ethidium out of the cell nor confer resistance against the substrates tested to levels achieved by the negative control, as discussed in Section 6.5. It is possible that when inserted into the bacterial membrane, the Gly17Cys mutant may make cells more permeable, allowing for antimicrobials to more readily penetrate the cell resulting in cell death at very low antimicrobial concentrations. Although clearly functionally impairing, a Cys substitution at this position does not affect expression of this mutant, as Glu17Cys was shown to be expressed at levels on par with WT.

The functional analysis of all QacC residues not only provides us with information regarding those residues which were seen to be absolutely essential for QacC activity, as discussed above, but also reveals that when substituted with Cys, many display compromised levels of ethidium efflux and/or low levels of resistance against the compounds assessed. Given the small size of the QacC protein when compared to larger transporters consisting of 12 to 14 TMS such as MdfA (Edgar and Bibi, 1997) and QacA (Section 1.3.2.3) (Paulsen et al., 1996a), residues making up QacC may represent the “bare essentials” and as such Cys substitutions, although permissible in certain areas, disrupt or compromise optimal QacC function. Whether this occurs as a result of a disruption to the interaction between QacC and its substrates, as is the case with Glu14 and other residues involved in substrate binding and/or translocation, or due to structural perturbations which may result in QacC instability, such as may be the case with Pro32, these substitutions ultimately affect function. Given extensive analysis, all

QacC mutations that influence function as found in this study, can be explained, especially when the intrinsic property of each amino acid and its putative location within QacC is taken into account. As described (Bay et al., 2008; Bay and Turner, 2009; Paulsen et al., 1996c), sequence conservation, at all levels, is highly apparent throughout members of the SMR family, indicating that although some disparity is allowed between residues at a particular location, it may still be limited to a certain amino acid property, such as charge or side chain. Further work to analyse this would involve substituting functionally critical amino acids, with those which have similar properties in order to further evaluate this idea.

9.5 SepA – the quest continues

Given its current classification as a *S. aureus* SMR protein, the SepA membrane transporter was evaluated in this study (Chapter 8). Chromosomally encoded, the SepA protein has 157 amino acids and is purported to be made up of four α-helices (Narui et al., 2002). The *sepA* gene was originally identified as a multidrug resistant gene which encodes a drug efflux protein. This protein was shown to confer resistance to acriflavine and contribute to ethidium efflux from *E. coli* cells (Narui et al., 2002). To date, no other studies have been carried out on this protein, although various reviews mention it as an SMR multidrug transporter (Costa et al., 2013; Schindler et al., 2013a).

Although shown to be highly conserved in all currently sequenced *S. aureus* genomes (Section 8.2), the exact function of the SepA protein is presently unknown. Fluorimetric transport assays using ethidium and acriflavine showed that SepA could not facilitate the transport of either compounds out the cell (Section 8.4 and 8.5). Evaluation of resistance, utilising two different methods (Section 8.5.1 and 8.5.2) also failed to confirm the role of SepA as a multidrug transporter, as no resistance was observed against any of the compounds tested (Figure 8.7 and 8.8).

As initial analysis failed to reproduce the findings presented previously (Narui et al., 2002), the sequence of SepA was further analysed. A multisequence alignment revealed that when aligned with other SMR proteins such as QacC and EmrE, the location of the essential SMR residues such as Glu14 and Trp63 differed (Figure

8.1). Found in TMS1 and TMS3 in QacC and EmrE respectively, in SepA these residues are located in TMS2 and TMS4. These, and other conserved SMR residues are offset by a large first TMS which shows no sequence similarity to any SMR proteins. To further evaluate the possible function of the TMS1 of SepA, a SepAΔ1 construct was generated containing the last three TMSs only. This construct was then used to generate a fusion protein consisting of SepAΔ1 and QacC TMS4 (SepA Δ1QacCTMS4), which repositioned all SMR essential residues to positions found in other SMR proteins (Figure 8.4). Nonetheless, evaluation of expression showed that although SepA and SepAΔ1 were expressed at levels similar to WT QacC, the SepAΔ1QacCTMS4 construct exhibited significantly lower expression levels, indicating that this construct was possibly unstable (Figure 8.5). Functional assessment of all SepA derivatives revealed that none could confer resistance against acriflavine or pyronin Y (Section 8.6), nor mediate ethidium or acriflavine efflux (Section 8.5.1).

The possibility that SepA belonged to a new family of multidrug efflux proteins, the PCE family, was also evaluated. Consisting of four TMS and possessing a conserved negatively charged residue in TMS1 (Glu15) shown to be essential for transport (Hassan et al., 2013), the SepA amino acid sequence was aligned with members of this new family to evaluate possible homology. Nonetheless, as seen in the alignment shown in Figure 8.3, SepA does not share any sequence conservation with PCE proteins, even after the PCE Glu15 was aligned with the SepA Glu81 residue. As to whether this residue is actually essential for SepA function is currently unknown, making sequence analysis using Glu81 as a point of reference questionable.

Upon closer inspection of the initial SepA study, a number of inconsistencies can be observed, as discussed in Section 8.1. Analysis into the five ORF analysed revealed that the resistance and efflux activity were mediated by the product of the third ORF, which has since been identified as the *S. aureus* SdrM multidrug transporter, belonging to the MFS family (Yamada et al., 2006a). Given the results obtained in this study it is difficult to elucidate the possible role of the SepA protein in resistance. Its high conservation amongst *S. aureus* genome would indicate

importance; as such the construction of a *sepA* knockout could be used to determine whether this protein is essential for growth and/or resistance. The possibility that SepA is a multidrug transporter exists, as an extensive substrate range was not evaluated. In an effort to further identify a possible SepA substrate, high-throughput experiments in which a large number of compounds could be assessed would need to be carried out. However, given that it does not confer resistance against ethidium, a common substrate amongst a large number of multidrug transporters, including those of the SMR family, its substrate profile may not include antimicrobials, but instead it may be involved in the transport of metabolites. Therefore, until a specific profile has been determined its identification as a multidrug transporter appears to be somewhat questionable.

9.6 Concluding remarks

The emergence and continuous spread of multidrug resistant bacteria is becoming a major problem for human health. As such, obtaining a detailed understanding of the molecular mechanisms behind bacterial resistance is paramount for the development of new therapeutical agents. Multidrug efflux transporters, which constitute one of the resistance mechanisms bacteria use to evade antimicrobials, are capable of extruding a wide variety of structurally dissimilar substrates from the bacterial cell. This ability gives many bacteria the capacity to resist commonly used antimicrobials and antibiotics. As many of the drug transporter genes are carried on mobile genetic elements such as plasmids or transposons, they have been disseminated throughout bacterial populations conferring resistance to otherwise susceptible bacteria. Given the ease with which bacteria can acquire such resistance genes, and the current overuse of antimicrobials and antibiotics, we are creating an environment which selects for resistant bacteria, ultimately leading to the proliferation of strains which are untreatable.

The adaptive power of *S. aureus* has led to the emergence of MRSA which continues to be a worldwide problem in hospital environments. The staphylococcus genome encodes a number of membrane transporters which are capable of

effluxing macrolides, quinolones, tetracycline and streptomycin antibiotics as well as biocides such as biguanadines, diamidines and QACs (Hassan et al., 2007a). Although becoming increasingly resistant to the currently available antimicrobial agents, the development of new antibiotics has stagnated (White and Need, 2011), with pharmaceutical companies abandoning the quest for new antimicrobials. With this decline the race to find alternative treatments or to improve the efficacy of the currently employed drugs continues. However, treatment is complicated as the genes encoding transporter proteins and other resistance mechanisms are widespread, conferring resistance against a number of antimicrobials. As such, combination therapy, including inhibitors that target MPs preventing the expulsion of antimicrobials is being continually explored (Lomovskaya et al., 2001; Schindler et al., 2013a). Although there are currently no inhibitors found that can be used to block drug efflux by SMR proteins, docking analysis has revealed that quercetin, a common flavonoid of honey, can potentially be used as a non-antibiotic adjuvant for treatment of tuberculosis. Simulation studies have revealed that this compound can interact with the Mycobacterium drug efflux pump Mmr, a homologue of the *E. coli* EmrE membrane transporter, and as such could possibly be used to inhibit drug efflux (Suriyanarayanan and Sarojini Santhosh, 2014).

The development of new antimicrobials and inhibitors has, nonetheless, become increasingly challenging and more scientifically complex, requiring in depth understanding of pathways and mechanisms which can be targeted. This highlights the importance of structural and functional analysis of multidrug transporters such as QacC, as it is through the detailed understanding of their mechanism that effective inhibitors can be developed. As such, the extensive analysis undertaken in this study, examining each residue within QacC and its possible role in function and/or structure may ultimately aid in elucidating the differences and common features that multidrug transporters possess, which may in turn assist in the development of clinically useful drugs and treatments which are much needed to combat the growing resistance presented by bacteria all over the world.

Chapter 10 – APPENDICES

Appendix I

List of abbreviations

~	approximately
%	percent
α	alpha
ABC	superfamily, adenosine triphosphate binding cassette superfamily
Amp	ampicillin
ATP	adenosine triphosphate
BLAST	basic local alignment tool
BLASTp	protein basic local alignment tool
Bp	base pairs
BTP	Bacterial Transmembrane Pair family
C^{α}	alpha carbon
°C	degrees Celsius
$CaCl_2$	calcium chloride
CA-MRSA	community acquired methicillin-resistant *S. aureus*
CCCP	carbonyl-cyanide m-chlorophenolhydrazone
C-terminal	carboxy-terminal
Cryo-EM	electron cryomicroscopy
DDM	n-dodecyl β-D-maltoside
dH_2O	distilled water
DHA1	familydrug proton antiporter 1 family
DHA2	family drug proton antiporter 2 family
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
DNase	deoxyribonuclease
dNTP	2'-deoxynucleotide 5'-triphosphate
ds	double stranded
EDTA	ethylenediaminetetra-acetic acid
FU s^{-1}	fluorescence units per second
H^+	proton
H_2O_2	hydrogen peroxide
Hisx6	tag encoding six consecutive histidine residues
Hr	hour(s)
IPTG	isopropyl-β-D-thiogalactopyranoside
kb	kilobase(s)
kDa	kilodalton(s)
KPO_4	potassium phosphate
L	litre(s)
LB	Luria-Bertani
M	molar

µg	microgram(s)
µl	microlitre(s)
µM	micromolar(s)
MATE	multidrug and toxic compound extrusion family
MDR	multidrug resistance
MFS	major facilitator superfamily
mg	milligram(s)
$MgCl_2$	magnesium chloride
MH	muller-hinton
MIC	minimum inhibitory concentration
min	minute
ml	milliliter(s)
mM	millimolar
MPs	membrane proteins
MRSA	methicillin-resistant *Staphylococcus aureus*
mV	millivolt(s)
MW	molecular weight
Na^+	sodium
NaCl	sodium chloride
NCBI	national centre for biotechnology information
NEM	N-ethylmaleimide
ng	nanogram(s)
nm	nanometer(s)
NMR	nuclear magnetic resonance
nt	nucleotide
N-terminal	amino-terminal
OD	optical density
O/N	overnight
ORF	open reading frame
PAGE	polyacrylamide gel electrophoresis
PCR	polymerase chain reaction
PBD	protein data bank
pmol	picomole(s)
PMF	proton motive force
PSI	pounds per square inch
PSMR	paired small multidrug resistance proteins
PVDF	polyvinylidine difluoride
QACs	quaternary ammonium compounds
R	resistance
RMSD	root-mean-square deviation
RNase	ribonuclease
RND	resistance nodulation-cell division family
rpm	revs per minute
RT	room temperature
SDS	sodium dodecyl sulphate
SDS-PAGE	sodium dodecyl sulphate-polyacrylamide gel electrophoresis

sec	second(s)
SMR	small multidrug family
SUG	suppressor of *groEL* mutation proteins
TAE	Tris-acetate EDTA buffer
TBS	tris-buffered saline
TCDB	transporter classification database
TEMED	*N,N,N', N'* – tetramethylethylene diamine
TPP	tetraphenylphosphonium
TM-score	template modelling score
TMD	transmembrane domain
TMS	transmembrane spanning segment(s)
Tris-HCl	Tris-hydromethylaminomethane-HCl
V	volt(s)
v / v	volume per volume
W	watts
w / v	weight per volume

Amino acids

A	Ala	Alanine
C	Cys	Cysteine
D	Asp	Aspartic acid
E	Glu	Glutamic acid
F	Phe	Phenylalanine
G	Gly	Glycine
H	His	Histidine
I	Ile	Isoleucine
K	Lys	Lysine
L	Leu	Leucine
M	Met	Methionine
N	Asn	Asparagine
P	Pro	Proline
Q	Gln	Glutamine
R	Arg	Arginine
S	Ser	Serine
T	The	Threonine
V	Val	Valine
W	Trp	Tryptophan
Y	Tyr	Tyrosine

Chapter 11 – REFERENCES

Abramson, J., Smirnova, I., Kasho, V., Verner, G., Iwata, S., and Kaback, H.R. (2003a). The lactose permease of *Escherichia coli*: overall structure, the sugar-binding site and the alternating access model for transport. *FEBS Lett* **555**, 96-101.

Abramson, J., Smirnova, I., Kasho, V., Verner, G., Kaback, H.R., and Iwata, S. (2003b). Structure and mechanism of the lactose permease of *Escherichia coli*. *Science* **301**, 610-615.

Adam, Y., Tayer, N., Rotem, D., Schreiber, G., and Schuldiner, S. (2007). The fast release of sticky protons: kinetics of substrate binding and proton release in a multidrug transporter. *Proc Natl Acad Sci USA* **104**, 17989-17994.

Adamian, L., and Liang, J. (2001). Helix-helix packing and interfacial pairwise interactions of residues in membrane proteins. *J Mol Biol* **311**, 891-907.

Adler, J., and Bibi, E. (2004). Determinants of substrate recognition by the *Escherichia coli* multidrug transporter MdfA identified on both sides of the membrane. *J Bio Chem* **279**, 8957-8965.

Adler, J., and Bibi, E. (2005). Promiscuity in the geometry of electrostatic interactions between the *Escherichia coli* multidrug resistance transporter MdfA and cationic substrates. *J Biol Chem* **280**, 2721-2729.

Akama, H., Matsuura, T., Kashiwagi, S., Yoneyama, H., Narita, S., Tsukihara, T., Nakagawa, A., and Nakae, T. (2004). Crystal structure of the membrane fusion protein, MexA, of the multidrug transporter in *Pseudomonas aeruginosa*. *J Biol Chem* **279**, 25939-25942.

Alekshun, M.N., and Levy, S.B. (2007). Molecular mechanisms of antibacterial multidrug resistance. *Cell* **128**, 1037-1050.

Alibayov, B., Baba-Moussa, L., Sina, H., Zdenkova, K., and Demnerova, K. (2014). *Staphylococcus aureus* mobile genetic elements. *Mol Biol Rep* **41**, 5005-5018.

Aller, S.G., Yu, J., Ward, A., Weng, Y., Chittaboina, S., Zhuo, R., Harrell, P.M., Trinh, Y.T., Zhang, Q., Urbatsch, I.L.*, et al.* (2009). Structure of P-glycoprotein reveals a molecular basis for poly-specific drug binding. *Science* **323**, 1718-1722.

Alvarez-Ortega, C., Olivares, J., and Martinez, J.L. (2013). RND multidrug efflux pumps: what are they good for? *Front Microbiol* **4**.

Archer, G.L. (1998). *Staphylococcus aureus*: a well-armed pathogen. *Clin Infect Dis* **26**, 1179-1181.

Arselin, G., Giraud, M.F., Dautant, A., Vaillier, J., Brethes, D., Coulary-Salin, B., Schaeffer, J., and Velours, J. (2003). The GxxxG motif of the transmembrane domain of subunit e is involved in the dimerization/oligomerization of the yeast ATP synthase complex in the mitochondrial membrane. *Eur J Biochem* **270**, 1875-1884.

Balganesh, M., Dinesh, N., Sharma, S., Kuruppath, S., Nair, A.V., and Sharma, U. (2012). Efflux pumps of *Mycobacterium tuberculosis* play a significant role in antituberculosis activity of potential drug candidates. *Antimicrob Agents Chemother* **56**, 2643-2651.

Ball, P.R., Shales, S.W., and Chopra, I. (1980). Plasmid-mediated tetracycline resistance in *Escherichia coli* involves increased efflux of the antibiotic. *Biochem Biophys Res Commun* **93**, 74-81.

Banchs, C., Poulos, S., Nimjareansuk, W.S., Joo, Y.E., and Faham, S. (2014). Substrate binding to the multidrug transporter MepA. *Biochim Biophys Acta* **1838**, 2539-2546.

Baneyx, F. (1999). Recombinant protein expression in *Escherichia coli*. *Curr Opin Biotechnol* **10**, 411-421.

Baneyx, F., and Mujacic, M. (2004). Recombinant protein folding and misfolding in *Escherichia coli*. *Nat Biotechnol* **22**, 1399-1408.

Bay, D.C., Rommens, K.L., and Turner, R.J. (2008). Small multidrug resistance proteins: a multidrug transporter family that continues to grow. *Biochim Biophys Acta* **1778**, 1814-1838.

Bay, D.C., and Turner, R.J. (2009). Diversity and evolution of the small multidrug resistance protein family. *BMC Evol Biol* **9**, 140.

Bernaudat, F., Frelet-Barrand, A., Pochon, N., Dementin, S., Hivin, P., Boutigny, S., Rioux, J.B., Salvi, D., Seigneurin-Berny, D., Richaud, P.*, et al.* (2011). Heterologous expression of membrane proteins: choosing the appropriate host. *PLoS One* **6**, e29191.

Bernsel, A., Viklund, H., Hennerdal, A., and Elofsson, A. (2009). TOPCONS: consensus prediction of membrane protein topology. *Nucleic Acids Res* **37**, W465-468.

Biemans-Oldehinkel, E., Doeven, M.K., and Poolman, B. (2006). ABC transporter architecture and regulatory roles of accessory domains. *FEBS Lett* **580**, 1023-1035.

Bill, R.M., Henderson, P.J.F., Iwata, S., Kunji, E.R.S., Michel, H., Neutze, R., Newstead, S., Poolman, B., Tate, C.G., and Vogel, H. (2011). Overcoming barriers to membrane protein structure determination. *Nat Biotechnol* **29**, 335-340.

Bjorland, J., Steinum, T., Kvitle, B., Waage, S., Sunde, M., and Heir, E. (2005). Widespread distribution of disinfectant resistance genes among staphylococci of bovine and caprine origin in Norway. *J Clin Microbiol* **43**, 4363-4368.

Bjorland, J., Steinum, T., Sunde, M., Waage, S., and Heir, E. (2003). Novel plasmid-borne gene *qacJ* mediates resistance to quaternary ammonium compounds in equine *Staphylococcus aureus*, *Staphylococcus simulans*, and *Staphylococcus intermedius*. *Antimicrob Agents Chemother* **47**, 3046-3052.

Blair, J.M., and Piddock, L.J. (2009). Structure, function and inhibition of RND efflux pumps in Gram-negative bacteria: an update. *Curr Opin Microbiol* **12**, 512-519.

Blaszczyk, M., Jamroz, M., Kmiecik, S., and Kolinski, A. (2013). CABS-fold: Server for the de novo and consensus-based prediction of protein structure. *Nucleic Acids Res* **41**, W406-411.

Bogdanov, M., Heacock, P.N., and Dowhan, W. (2002). A polytopic membrane protein displays a reversible topology dependent on membrane lipid composition. *EMBO J* **21**, 2107-2116.

Boyd, D., and Beckwith, J. (1989). Positively charged amino acid residues can act as topogenic determinants in membrane proteins. *Proc Natl Acad Sci USA* **86**, 9446-9450.

Bradley, P., Misura, K.M., and Baker, D. (2005). Toward high-resolution de novo structure prediction for small proteins. *Science* **309**, 1868-1871.

Brandl, C.J., and Deber, C.M. (1986). Hypothesis about the function of membrane-buried proline residues in transport proteins. *Proc Natl Acad Sci USA* **83**, 917-921.

Brooks, B.E., Piro, K.M., and Brennan, R.G. (2007). Multidrug-binding transcription factor QacR binds the bivalent aromatic diamidines DB75 and DB359 in multiple positions. *J Am Chem Soc* **129**, 8389-8395.

Brown, M.H., Paulsen, I.T., and Skurray, R.A. (1999). The multidrug efflux protein NorM is a prototype of a new family of transporters. *Mol Microbiol* **31**, 394-395.

Brown, M.H., and Skurray, R.A. (2001). Staphylococcal multidrug efflux protein QacA. *J Mol Microbiol Biotechnol* **3**, 163-170.

Budavari, S., ed. (1996). The Merck index: an encyclopedia of chemicals, drugs, and biologicals. (Whitehouse Station, NJ).

Butler, P.J., Ubarretxena-Belandia, I., Warne, T., and Tate, C.G. (2004). The *Escherichia coli* multidrug transporter EmrE is a dimer in the detergent-solubilised state. *J Mol Biol* **340**, 797-808.

Callaghan, R., Luk, F., and Bebawy, M. (2014). Inhibition of the multidrug resistance P-glycoprotein: time for a change of strategy? *Drug Metab Dispos* **42**, 623-631.

Carlson, E.D., Gan, R., Hodgman, C.E., and Jewett, M.C. (2012). Cell-free protein synthesis: Applications come of age. *Biotechnol Adv* **30**, 1185-1194.

Casey, A.L., Lambert, P.A., and Elliott, T.S. (2007). Staphylococci. *Int J Antimicrob Ag* **29 Suppl 3**, S23-32.

Cervinkova, D., Babak, V., Marosevic, D., Kubikova, I., and Jaglic, Z. (2013). The role of the *qacA* gene in mediating resistance to quaternary ammonium compounds. *Microb Drug Resist* **19**, 160-167.

Chambers, H.F., and Deleo, F.R. (2009). Waves of resistance: *Staphylococcus aureus* in the antibiotic era. *Nat Rev Microbiol* **7**, 629-641.

Chang, G. (2003). Multidrug resistance ABC transporters. *FEBS Lett* **555**, 102-105.

Chang, G., Roth, C.B., Reyes, C.L., Pornillos, O., Chen, Y.J., and Chen, A.P. (2006). Retraction. *Science* **314**, 1875.

Charalambous, K., Miller, D., Curnow, P., and Booth, P.J. (2008). Lipid bilayer composition influences small multidrug transporters. *BMC Biochem* **9**, 31.

Chen, J., Morita, Y., Huda, M.N., Kuroda, T., Mizushima, T., and Tsuchiya, T. (2002). VmrA, a member of a novel class of Na(+)-coupled multidrug efflux pumps from *Vibrio parahaemolyticus*. *J Bacteriol* **184**, 572-576.

Chen, Y.J., Pornillos, O., Lieu, S., Ma, C., Chen, A.P., and Chang, G. (2007). X-ray structure of EmrE supports dual topology model. *Proc Natl Acad Sci USA* **104**, 18999-19004.

Cho, M.K., Gayen, A., Banigan, J.R., Leninger, M., and Traaseth, N.J. (2014). Intrinsic conformational plasticity of native EmrE provides a pathway for multidrug resistance. *J Am Chem Soc* **136**, 8072-8080.

Chouchani, E.T., James, A.M., Fearnley, I.M., Lilley, K.S., and Murphy, M.P. (2011). Proteomic approaches to the characterization of protein thiol modification. *Curr Opin Chem Biol* **15**, 120-128.

Chung, Y.J., and Saier, M.H. (2002). Overexpression of the *Escherichia coli sugE* gene confers resistance to a narrow range of quaternary ammonium compounds. *J Bacteriol* **184**, 2543-2545.

Claros, M.G., and von Heijne, G. (1994). TopPred II: an improved software for membrane protein structure predictions. *Comp Appl Biosci* **10**, 685-686.

Cole, C., Barber, J.D., and Barton, G.J. (2008). The Jpred 3 secondary structure prediction server. *Nucleic Acids Res* **36**, W197-201.

Collu, F., and Cascella, M. (2013). Multidrug resistance and efflux pumps: insights from molecular dynamics simulations. *Curr Top Med Chem* **13**, 3165-3183.

Colovos, C., and Yeates, T.O. (1993). Verification of protein structures: patterns of nonbonded atomic interactions. *Protein Sci* **2**, 1511-1519.

Correa, J.E., De Paulis, A., Predari, S., Sordelli, D.O., and Jeric, P.E. (2008). First report of *qacG*, *qacH* and *qacJ* genes in *Staphylococcus haemolyticus* human clinical isolates. *J Antimicrob Chemother* **62**, 956-960.

Costa, S.S., Viveiros, M., Amaral, L., and Couto, I. (2013). Multidrug efflux pumps in *Staphylococcus aureus*: an update. *Open Microbiol J* **7**, 59-71.

Cozzetto, D., and Tramontano, A. (2008). Advances and pitfalls in protein structure prediction. *Curr Protein Pept Sc* **9**, 567-577.

Crooks, G.E., Hon, G., Chandonia, J.M., and Brenner, S.E. (2004). WebLogo: a sequence logo generator. *Genome Res* **14**, 1188-1190.

Cruz, A., Micaelo, N., Felix, V., Song, J.Y., Kitamura, S., Suzuki, S., and Mendo, S. (2013). *sugE*: A gene involved in tributyltin (TBT) resistance of *Aeromonas molluscorum Av27*. *J Gen Appl Microbiol* **59**, 39-47.

Cuff, J.A., Clamp, M.E., Siddiqui, A.S., Finlay, M., and Barton, G.J. (1998). JPred: a consensus secondary structure prediction server. *Bioinformatics* **14**, 892-893.

Curran, A.R., and Engelman, D.M. (2003). Sequence motifs, polar interactions and conformational changes in helical membrane proteins. *Curr Opin Struct Biol* **13**, 412-417.

Dalbey, R.E., Wang, P., and Kuhn, A. (2011). Assembly of bacterial inner membrane proteins. *Annu Rev Biochem* **80**, 161-187.

Daniels, C., and Morona, R. (1999). Analysis of *Shigella flexneri* Wzz (Rol) function by mutagenesis and cross-linking: Wzz is able to oligomerize. *Mol Microbiol* **34**, 181-194.

Das, D., Xu, Q.S., Lee, J.Y., Ankoudinova, I., Huang, C., Lou, Y., DeGiovanni, A., Kim, R., and Kim, S.H. (2007). Crystal structure of the multidrug efflux transporter AcrB at 3.1Å resolution reveals the N-terminal region with conserved amino acids. *J Struct Biol* **158**, 494-502.

Dawson, R.J., and Locher, K.P. (2006). Structure of a bacterial multidrug ABC transporter. *Nature* **443**, 180-185.

Dawson, R.J., and Locher, K.P. (2007). Structure of the multidrug ABC transporter Sav1866 from *Staphylococcus aureus* in complex with AMP-PNP. *FEBS Lett* **581**, 935-938.

de Lencastre, H., Oliveira, D., and Tomasz, A. (2007). Antibiotic resistant *Staphylococcus aureus*: a paradigm of adaptive power. *Curr Opin Microbiol* **10**, 428-435.

De Marothy, M.T., and Elofsson, A. (2015). Marginally hydrophobic transmembrane α-helices shaping membrane protein folding. . *Protein Sci* **24**, 1057-1074.

De Rossi, E., Arrigo, P., Bellinzoni, M., Silva, P.A., Martin, C., Ainsa, J.A., Guglierame, P., and Riccardi, G. (2002). The multidrug transporters belonging to major facilitator superfamily in *Mycobacterium tuberculosis*. *Mol Med* **8**, 714-724.

De Rossi, E., Branzoni, M., Cantoni, R., Milano, A., Riccardi, G., and Ciferri, O. (1998). mmr, a *Mycobacterium tuberculosis* gene conferring resistance to small cationic dyes and inhibitors. *J Bacteriol* **180**, 6068-6071.

DeLano, W.L. (2002). The PyMOL molecular graphics system D. Scientific, ed. (Palo Alto, CA).

Deurenberg, R.H., and Stobberingh, E.E. (2008). The evolution of *Staphylococcus aureus*. *Infect Genet Evol* **8**, 747-763.

Donnelly, D., Overington, J.P., Ruffle, S.V., Nugent, J.H., and Blundell, T.L. (1993). Modeling alpha-helical transmembrane domains: the calculation and use of substitution tables for lipid-facing residues. *Protein Sci* **2**, 55-70.

Du, D., Wang, Z., James, N.R., Voss, J.E., Klimont, E., Ohene-Agyei, T., Venter, H., Chiu, W., and Luisi, B.F. (2014). Structure of the AcrAB-TolC multidrug efflux pump. *Nature* **509**, 512-515.

Duarte, J.M., Biyani, N., Baskaran, K., and Capitani, G. (2013). An analysis of oligomerization interfaces in transmembrane proteins. *BMC Struct Biol* **13**, 21.

Dunham, T., Schumacher, M.A., and Brennan, R.G. (2001). Unpublished Work.

Duran, N., Temiz, M., Duran, G.G., Eryilmaz, N., and Jenedi, K. (2014). Relationship between the resistance genes to quaternary ammonium compounds and antibiotic resistance in staphylococci isolated from surgical site infections. *Med Sci Monit* **20**, 544-550.

Dutta, S., Morrison, E.A., and Henzler-Wildman, K.A. (2014). EmrE dimerization depends on membrane environment. *Biochim Biophys Acta* **1838**, 1817-1822.

Eddy, S.R. (2004). Where did the BLOSUM62 alignment score matrix come from? *Nat Biotechnol* **22**, 1035-1036.

Edgar, R., and Bibi, E. (1997). MdfA, an *Escherichia coli* multidrug resistance protein with an extraordinarily broad spectrum of drug recognition. *J Bacteriol* **179**, 2274-2280.

Elbaz, Y., Salomon, T., and Schuldiner, S. (2008). Identification of a glycine motif required for packing in EmrE, a multidrug transporter from *Escherichia coli*. *J Biol Chem* **283**, 12276-12283.

Elbaz, Y., Steiner-Mordoch, S., Danieli, T., and Schuldiner, S. (2004). In vitro synthesis of fully functional EmrE, a multidrug transporter, and study of its oligomeric state. *Proc Natl Acad Sci USA* **101**, 1519-1524.

Elkins, C.A., and Nikaido, H. (2002). Substrate specificity of the RND-type multidrug efflux pumps AcrB and AcrD of *Escherichia coli* is determined predominantly by two large periplasmic loops. *J Bacteriol* **184**, 6490-6498.

Engelman, D.M., Steiz, T.A., and Goldman, A. (1986). Identifying nonpolar transbilayer helices in amino acids sequences of membrane proteins. *Annu Rev Biophys Bio* **15**, 321-353.

Eswar, N., Eramian, D., Webb, B., Shen, M.Y., and Sali, A. (2008). Protein structure modeling with MODELLER. *Methods Mol Biol* **426**, 145-159.

Fink, A., Sal-Man, N., Gerber, D., and Shai, Y. (2012). Transmembrane domains interactions within the membrane milieu: principles, advances and challenges. *Biochim Biophys Acta* **1818**, 974-983.

Fitzgerald, J.R., Sturdevant, D.E., Mackie, S.M., Gill, S.R., and Musser, J.M. (2001). Evolutionary genomics of *Staphylococcus aureus*: insights into the origin of methicillin-resistant strains and the toxic shock syndrome epidemic. *Proc Natl Acad Sci USA* **98**, 8821-8826.

Fleishman, S.J., and Ben-Tal, N. (2006). Progress in structure prediction of alpha-helical membrane proteins. *Curr Opin Struct Biol* **16**, 496-504.

Fleishman, S.J., Harrington, S.E., Enosh, A., Halperin, D., Tate, C.G., and Ben-Tal, N. (2006). Quasi-symmetry in the cryo-EM structure of EmrE provides the key to modeling its transmembrane domain. *J Mol Biol* **364**, 54-67.

Floudas, C.A., Fung, H.K., McAllister, S.R., Monnigmann, M., and Rajgaria, R. (2006). Advances in protein structure prediction and de novo protein design: A review. *Chem Eng Sci* **61**, 966-988.

Floyd, J.L., Smith, K.P., Kumar, S.H., Floyd, J.T., and Varela, M.F. (2010). LmrS is a multidrug efflux pump of the major facilitator superfamily from *Staphylococcus aureus*. *Antimicrob Agents Chemother* **54**, 5406-5412.

Fluman, N., Ryan, C.M., Whitelegge, J.P., and Bibi, E. (2012). Dissection of mechanistic principles of a secondary multidrug efflux protein. *Mol Cell* **47**, 777-787.

Fojo, A.T., Ueda, K., Slamon, D.J., Poplack, D.G., Gottesman, M.M., and Pastan, I. (1987). Expression of a multidrug-resistance gene in human tumors and tissues. *Proc Natl Acad Sci USA* **84**, 265-269.

Forrest, L.R., Tang, C.L., and Honig, B. (2006). On the accuracy of homology modeling and sequence alignment methods applied to membrane proteins. *Biophys J* **91**, 508-517.

Friesen, R.H., Knol, J., and Poolman, B. (2000). Quaternary structure of the lactose transport protein of *Streptococcus thermophilus* in the detergent-solubilized and membrane-reconstituted state. *J Biol Chem* **275**, 40658.

Frillingos, S., Sahin-Toth, M., Wu, J., and Kaback, H.R. (1998). Cys-scanning mutagenesis: a novel approach to structure function relationships in polytopic membrane proteins. *FASEB J* **12**, 1281-1299.

Froshauer, S., and Beckwith, J. (1984). The nucleotide sequence of the gene for *malF* protein, an inner membrane component of the maltose transport system of *Escherichia coli*. Repeated DNA sequences are found in the *malE-malF* intercistronic region. *J Biol Chem* **259**, 10896-10903.

Froshauer, S., Green, G.N., Boyd, D., McGovern, K., and Beckwith, J. (1988). Genetic analysis of the membrane insertion and topology of MalF, a cytoplasmic membrane protein of *Escherichia coli*. *J Mol Biol* **200**, 501-511.

Ganas, P., Igloi, G.L., and Brandsch, R., eds. (2009). The megaplasmid pAO1 of *Arthrobacter nicotinovorans* and nicotine catabolism (Springer Berlin Heidelberg, Heidelberg).

Ganas, P., Mihasan, M., Igloi, G.L., and Brandsch, R. (2007). A two-component small multidrug resistance pump functions as a metabolic valve during nicotine catabolism by *Arthrobacter nicotinovorans*. *Microbiology* **153**, 1546-1555.

Gautier, R., Douguet, D., Antonny, B., and Drin, G. (2008). HELIQUEST: a web server to screen sequences with specific alpha-helical properties. *Bioinformatics* **24**, 2101-2102.

Ghasemzadeh-Moghaddam, H., van Belkum, A., Hamat, R.A., van Wamel, W., and Neela, V. (2014). Methicillin-susceptible and-resistant *Staphylococcus aureus* with high-level antiseptic and low-level mupirocin resistance in malaysia. *Microb Drug Resist* **20**, 472-477.

Gibbons, S., and Udo, E.E. (2000). The effect of reserpine, a modulator of multidrug efflux pumps, on the in vitro activity of tetracycline against clinical isolates of methicillin resistant *Staphylococcus aureus* (MRSA) possessing the tet(K) determinant. *Phytother Res* **14**, 139-140.

Ginn, S.L., Brown, M.H., and Skurray, R.A. (2000). The TetA(K) tetracycline/H(+) antiporter from *Staphylococcus aureus*: mutagenesis and functional analysis of motif C. *J Bacteriol* **182**, 1492-1498.

Giorgetti, A. (2014). Membrane proteins: Insights from computational biology *Bio Med J* **6**, 101-102.

Goldberg, M., Pribyl, T., Juhnke, S., and Nies, D.H. (1999). Energetics and topology of CzcA, a cation/proton antiporter of the resistance-nodulation-cell division protein family. *J Biol Chem* **274**, 26065-26070.

Gottesman, M.M., Fojo, T., and Bates, S.E. (2002). Multidrug resistance in cancer: role of ATP-dependent transporters. *Nat Rev Cancer* **2**, 48-58.

Greener, T., Govezensky, D., and Zamir, A. (1993). A novel multicopy suppressor of a *groEL* mutation includes two nested open reading frames transcribed from different promoters. *EMBO J* **12**, 889-896.

Grinius, L., Dreguniene, G., Goldberg, E.B., Liao, C.H., and Projan, S.J. (1992). A Staphylococcal multidrug resistance gene-product is a member of a new-protein family. *Plasmid* **27**, 119-129.

Grinius, L.L., and Goldberg, E.B. (1994). Bacterial multidrug-resistance is due to a single membrane-protein which functions as a drug pump. *J Biol Chem* **269**, 29998-30004.

Grisshammer, R., and Tate, C.G. (1995). Overexpression of integral membrane proteins for structural studies. *Q Rev Biophys* **28**, 315-422.

Grkovic, S., Brown, M.H., Roberts, M.J., Paulsen, I.T., and Skurray, R.A. (1998). QacR is a repressor protein that regulates expression of the *Staphylococcus aureus* multidrug efflux pump QacA. *J Biol Chem* **273**, 18665-18673.

Grkovic, S., Brown, M.H., Schumacher, M.A., Brennan, R.G., and Skurray, R.A. (2001). The staphylococcal QacR multidrug regulator binds a correctly spaced operator as a pair of dimers. *J Bacteriol* **183**, 7102-7109.

Grkovic, S., Hardie, K.M., Brown, M.H., and Skurray, R.A. (2003). Interactions of the QacR multidrug-binding protein with structurally diverse ligands: implications for the evolution of the binding pocket. *Biochemistry* **42**, 15226-15236.

Grosdidier, A., Zoete, V., and Michielin, O. (2011). SwissDock, a protein-small molecule docking web service based on EADock DSS. *Nucleic Acids Res* **39**, W270-277.

Guan, L., and Nakae, T. (2001). Identification of essential charged residues in transmembrane segments of the multidrug transporter MexB of *Pseudomonas aeruginosa*. *J Bacteriol* **183**, 1734-1739.

Gubellini, F., Verdon, G., Karpowich, N.K., Luff, J.D., Boel, G., Gauthier, N., Handelman, S.K., Ades, S.E., and Hunt, J.F. (2011). Physiological response to membrane protein overexpression in *E. coli*. *Mol Cell Proteomics* **10**, M111 007930.

Gutman, N., Steiner-Mordoch, S., and Schuldiner, S. (2003). An amino acid cluster around the essential Glu-14 is part of the substrate- and proton-binding domain of EmrE, a multidrug transporter from *Escherichia coli*. *J Biol Chem* **278**, 16082-16087.

Guzman, L.M., Belin, D., Carson, M.J., and Beckwith, J. (1995). Tight regulation, modulation, and high-level expression by vectors containing the arabinose pBAD promoter. *J Bacteriol* **177**, 4121-4130.

Hanahan, D. (1983). Studies on transformation of *Escherichia coli* with plasmids. *J Mol Biol* **166**, 557-580.

Hannig, G., and Makrides, S.C. (1998). Strategies for optimizing heterologous protein expression in *Escherichia coli*. *Trends Biotechnol* **16**, 54-60.

Hassan, K.A. (2007). Major facilitator superfamily drug transport proteins of the Staphylococci. School of Biological Sciences (Sydney, The University of Sydney), pp. 205.

Hassan, K.A., Galea, M., Wu, J., Mitchell, B.A., Skurray, R.A., and Brown, M.H. (2006a). Functional effects of intramembranous proline substitutions in the staphylococcal multidrug transporter QacA. *FEMS Microbiol Lett* **263**, 76-85.

Hassan, K.A., Jackson, S.M., Penesyan, A., Patching, S.G., Tetu, S.G., Eijkelkamp, B.A., Brown, M.H., Henderson, P.J., and Paulsen, I.T. (2013). Transcriptomic and biochemical analyses identify a family of chlorhexidine efflux proteins. *Proc Natl Acad Sci USA* **110**, 20254-20259.

Hassan, K.A., Liu, Q., Henderson, P.J., and Paulsen, I.T. (2015). Homologs of the *Acinetobacter baumannii* AceI transporter represent a new family of bacterial multidrug efflux systems. *MBio* **6**.

Hassan, K.A., Robinson, K.L., Smith, A.N., Gibson, J.H., Skurray, R.A., and Brown, M.H. (2006b). Glycine-rich transmembrane helix 10 in the staphylococcal tetracycline transporter TetA(K) lines a solvent-accessible channel. *Biochemistry* **45**, 15661-15669.

Hassan, K.A., Skurray, R.A., and Brown, M.H. (2007a). Active export proteins mediating drug resistance in staphylococci. *J Mol Microbiol Biotechnol* **12**, 180-196.

Hassan, K.A., Skurray, R.A., and Brown, M.H. (2007b). Transmembrane helix 12 of the *Staphylococcus aureus* multidrug transporter QacA lines the bivalent cationic drug binding pocket. *J Bacteriol* **189**, 9131-9134.

Hassan, K.A., Souhani, T., Skurray, R.A., and Brown, M.H. (2008). Analysis of tryptophan residues in the staphylococcal multidrug transporter QacA reveals long-distance functional associations of residues on the opposite sides of the membrane. *J Bacteriol* **190**, 2441-2449.

Hassan, K.A., Xu, Z.Q., Watkins, R.E., Brennan, R.G., Skurray, R.A., and Brown, M.H. (2009). Optimized production and analysis of the staphylococcal multidrug efflux protein QacA. *Protein Expres Purif* **64**, 118-124.

He, G.X., Zhang, C., Crow, R.R., Thorpe, C., Chen, H., Kumar, S., Tsuchiya, T., and Varela, M.F. (2011). SugE, a new member of the SMR family of transporters, contributes to antimicrobial resistance in *Enterobacter cloacae*. *Antimicrob Agents Chemother* **55**, 3954-3957.

He, X., Szewczyk, P., Karyakin, A., Evin, M., Hong, W.X., Zhang, Q., and Chang, G. (2010). Structure of a cation-bound multidrug and toxic compound extrusion transporter. *Nature* **467**, 991-994.

Heijne, G. (1986). The distribution of positively charged residues in bacterial inner membrane proteins correlates with the trans-membrane topology. *EMBO J* **5**, 3021-3027.

Heir, E., Sundheim, G., and Holck, A.L. (1998). The Staphylococcus *qacH* gene product: a new member of the SMR family encoding multidrug resistance. *FEMS Microbiol Lett* **163**, 49-56.

Heir, E., Sundheim, G., and Holck, A.L. (1999a). Identification and characterization of quaternary ammonium compound resistant staphylococci from the food industry. *Int J Food Microbiol* **48**, 211-219.

Heir, E., Sundheim, G., and Holck, A.L. (1999b). The *qacG* gene on plasmid pST94 confers resistance to quaternary ammonium compounds in staphylococci isolated from the food industry. *J Appl Microbiol* **86**, 378-388.

Henikoff, S., and Henikoff, J.G. (1992). Amino acid substitution matrices from protein blocks. *Proc Natl Acad Sci USA* **89**, 10915-10919.

Henzler-Wildman, K. (2012). Analyzing conformational changes in the transport cycle of EmrE. *Curr Opin Struct Biol* **22**, 38-43.

Heuberger, E.H., Veenhoff, L.M., Duurkens, R.H., Friesen, R.H., and Poolman, B. (2002). Oligomeric state of membrane transport proteins analyzed with blue native electrophoresis and analytical ultracentrifugation. *J Mol Biol* **317**, 591-600.

Higgins, C.F. (1992). ABC transporters: from microorganisms to man. *Annu Rev Cell Biol* **8**, 67-113.

Higgins, C.F. (2001). ABC transporters: physiology, structure and mechanism--an overview. *Res Microbiol* **152**, 205-210.

Hiramatsu, K., Aritaka, N., Hanaki, H., Kawasaki, S., Hosoda, Y., Hori, S., Fukuchi, Y., and Kobayashi, I. (1997). Dissemination in Japanese hospitals of strains of *Staphylococcus aureus* heterogeneously resistant to vancomycin. *Lancet* **350**, 1670-1673.

Hirata, T., Wakatabe, R., Nielsen, J., Someya, Y., Fujihira, E., Kimura, T., and Yamaguchi, A. (1997). A novel compound, 1,1-dimethyl-5(1-hydroxypropyl)-4,6,7-trimethylindan, is an effective inhibitor of the tet(K) gene-encoded metal-tetracycline/H+ antiporter of *Staphylococcus aureus*. *FEBS Lett* **412**, 337-340.

Holden, M.T., Feil, E.J., Lindsay, J.A., Peacock, S.J., Day, N.P., Enright, M.C., Foster, T.J., Moore, C.E., Hurst, L., Atkin, R.*, et al.* (2004). Complete genomes of two clinical *Staphylococcus aureus* strains: evidence for the rapid evolution of virulence and drug resistance. *Proc Natl Acad Sci USA* **101**, 9786-9791.

Hollenstein, K., Frei, D.C., and Locher, K.P. (2007). Structure of an ABC transporter in complex with its binding protein. *Nature* **446**, 213-216.

Hooft, R.W., Sander, C., and Vriend, G. (1997). Objectively judging the quality of a protein structure from a Ramachandran plot. *Comp Appl Biosci* **13**, 425-430.

Hopf, T.A., Colwell, L.J., Sheridan, R., Rost, B., Sander, C., and Marks, D.S. (2012). Three-dimensional structures of membrane proteins from genomic sequencing. *Cell* **149**, 1607-1621.

Hruz, P.W., and Mueckler, M.M. (1999). Cysteine-scanning mutagenesis of transmembrane segment 7 of the GLUT1 glucose transporter. *J Biol Chem* **274**, 36176-36180.

Huang, J., O'Toole, P.W., Shen, W., Amrine-Madsen, H., Jiang, X., Lobo, N., Palmer, L.M., Voelker, L., Fan, F., Gwynn, M.N.*, et al.* (2004). Novel chromosomally encoded multidrug efflux transporter MdeA in *Staphylococcus aureus*. *Antimicrob Agents Chemother* **48**, 909-917.

Huet, A.A., Raygada, J.L., Mendiratta, K., Seo, S.M., and Kaatz, G.W. (2008). Multidrug efflux pump overexpression in *Staphylococcus aureus* after single and multiple in vitro exposures to biocides and dyes. *Microbiology* **154**, 3144-3153.

Hujer, K.M., Hujer, A.M., Hulten, E.A., Bajaksouzian, S., Adams, J.M., Donskey, C.J., Ecker, D.J., Massire, C., Eshoo, M.W., Sampath, R.*, et al.* (2006). Analysis of antibiotic resistance genes in multidrug-resistant *Acinetobacter* sp isolates from military and civilian patients treated at the Walter Reed Army Medical Center. *Antimicrob Agents Chemother* **50**, 4114-4123.

Hurwitz, N., Pellegrini-Calace, M., and Jones, D.T. (2006). Towards genome-scale structure prediction for transmembrane proteins. *Philos Trans R Soc Lond B Biol Sci* **361**, 465-475.

Husain, F., and Nikaido, H. (2010). Substrate path in the AcrB multidrug efflux pump of *Escherichia coli*. *Mol Microbiol* **78**, 320-330.

Hvorup, R.N., Goetz, B.A., Niederer, M., Hollenstein, K., Perozo, E., and Locher, K.P. (2007). Asymmetry in the structure of the ABC transporter-binding protein complex BtuCD-BtuF. *Science* **317**, 1387-1390.

Iscla, I., and Blount, P. (2012). Sensing and responding to membrane tension: the bacterial MscL channel as a model system. *Biophys J* **103**, 169-174.

Jack, D.L., Storms, M.L., Tchieu, J.H., Paulsen, I.T., and Saier, M.H., Jr. (2000). A broad-specificity multidrug efflux pump requiring a pair of homologous SMR-type proteins. *J Bacteriol* **182**, 2311-2313.

Jameson, D.M., and Ross, J.A. (2010). Fluorescence polarization/anisotropy in diagnostics and imaging. *Chem Rev* **110**, 2685-2708.

Jayaram, B. (2013). Some recent progresses and perspectives in predictions of protein structure, dynamics and function. *J Protein Proteom* **4**, 65-66.

Jevons, M.P., Rolinson, G.N., and Knox, R. (1961). Celbenin-resistant staphylococci. *Brit Med J* **1**, 124-&.

Jones, P.M., and George, A.M. (2013). Mechanism of the ABC transporter ATPase domains: catalytic models and the biochemical and biophysical record. *Crit Rev Biochem Mol Biol* **48**, 39-50.

Jones, P.M., O'Mara, M.L., and George, A.M. (2009). ABC transporters: a riddle wrapped in a mystery inside an enigma. *Trends Biochem Sci* **34**, 520-531.

Junge, F., Haberstock, S., Roos, C., Stefer, S., Proverbio, D., Dotsch, V., and Bernhard, F. (2011). Advances in cell-free protein synthesis for the functional and structural analysis of membrane proteins. *New Biotechnol* **28**, 262-271.

Kaatz, G.W., DeMarco, C.E., and Seo, S.M. (2006). MepR, a repressor of the *Staphylococcus aureus* MATE family multidrug efflux pump MepA, is a substrate-responsive regulatory protein. *Antimicrob Agents Chemother* **50**, 1276-1281.

Kaatz, G.W., McAleese, F., and Seo, S.M. (2005). Multidrug resistance in *Staphylococcus aureus* due to overexpression of a novel multidrug and toxin extrusion (MATE) transport protein. *Antimicrob Agents Chemother* **49**, 1857-1864.

Kaback, H.R., Smirnova, I., Kasho, V., Nie, Y., and Zhou, Y. (2011). The alternating access transport mechanism in LacY. *J Membrane Biol* **239**, 85-93.

Karpowich, N.K., and Wang, D.N. (2008). Structural biology. Symmetric transporters for asymmetric transport. *Science* **321**, 781-782.

Kelley, L.A., and Sternberg, M.J.E. (2009). Protein structure prediction on the Web: a case study using the Phyre server. *Nat Protocols* **4**, 363-371.

Khoury, G.A., Smadbeck, J., Kieslich, C.A., and Floudas, C.A. (2014). Protein folding and de novo protein design for biotechnological applications. *Trends Biotechnol* **32**, 99-109.

Kikukawa, T., Nara, T., Araiso, T., Miyauchi, S., and Kamo, N. (2006). Two-component bacterial multidrug transporter, EbrAB: Mutations making each component solely functional. *Biochim Biophys Acta* **1758**, 673-679.

Kim, E.H., Nies, D.H., McEvoy, M.M., and Rensing, C. (2011). Switch or funnel: how RND-type transport systems control periplasmic metal homeostasis. *J Bacteriol* **193**, 2381-2387.

Kim, Y., Ho, S.O., Gassman, N.R., Korlann, Y., Landorf, E.V., Collart, F.R., and Weiss, S. (2008). Efficient site-specific labeling of proteins via cysteines. *Bioconjugate Chem* **19**, 786-791.

Kleywegt, G.J., and Jones, T.A. (1996). Phi/psi-chology: Ramachandran revisited. *Structure* **4**, 1395-1400.

Klockenbusch, C., and Kast, J. (2010). Optimization of formaldehyde cross-linking for protein interaction analysis of non-tagged integrin beta1. *J Biomed Biotechnol* **2010**, 927585.

Koch, S., Fritsch, M.J., Buchanan, G., and Palmer, T. (2012). *Escherichia coli* TatA and TatB proteins have N-out, C-in topology in intact cells. *J Biol Chem* **287**, 14420-14431.

Korkhov, V.M., and Tate, C.G. (2009). An emerging consensus for the structure of EmrE. *Acta Crystallogr D Biol Crystallogr* **65**, 186-192.

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

Kozma, D., Simon, I., and Tusnady, G.E. (2013). PDBTM: Protein Data Bank of transmembrane proteins after 8 years. *Nucleic Acids Res* **41**, D524-529.

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

Kumar, A., and Schweizer, H.P. (2005). Bacterial resistance to antibiotics: active efflux and reduced uptake. *Adv Drug Deliver Rev* **57**, 1486-1513.

Kuroda, M., Ohta, T., Uchiyama, I., Baba, T., Yuzawa, H., Kobayashi, I., Cui, L., Oguchi, A., Aoki, K., Nagai, Y.*, et al.* (2001). Whole genome sequencing of meticillin-resistant *Staphylococcus aureus*. *Lancet* **357**, 1225-1240.

Kyte, J., and Doolittle, R. (1982). A simple method for displaying the hydropathic character of a protein. *J Mol Biol* **157**, 105.

Larkin, M.A., Blackshields, G., Brown, N.P., Chenna, R., McGettigan, P.A., McWilliam, H., Valentin, F., Wallace, I.M., Wilm, A., Lopez, R.*, et al.* (2007). Clustal W and Clustal X version 2.0. *Bioinformatics* **23**, 2947-2948.

Laskowski, R.A., Rullmannn, J.A., MacArthur, M.W., Kaptein, R., and Thornton, J.M. (1996). AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. *J Biomol NMR* **8**, 477-486.

Law, C.J., Maloney, P.C., and Wang, D.N. (2008). Ins and outs of major facilitator superfamily antiporters. *Annu Rev Microbiol* **62**, 289-305.

Lee, S.F., Shah, S., Yu, C., Wigley, W.C., Li, H., Lim, M., Pedersen, K., Han, W., Thomas, P., Lundkvist, J.*, et al.* (2004). A conserved GXXXG motif in APH-1 is critical for assembly and activity of the gamma-secretase complex. *J Biol Chem* **279**, 4144-4152.

Lemieux, M.J., Huang, Y., and Wang, D.N. (2004). The structural basis of substrate translocation by the *Escherichia coli* glycerol-3-phosphate transporter: a member of the major facilitator superfamily. *Curr Opin Struct Biol* **14**, 405-412.

Lewinson, O., Adler, J., Poelarends, G.J., Mazurkiewicz, P., Driessen, A.J.M., and Bibi, E. (2003). The *Escherichia coli* multidrug transporter MdfA catalyzes both electrogenic and electroneutral transport reactions. *Proc Natl Acad Sci USA* **100**, 1667-1672.

Li, X.Z., and Nikaido, H. (2004). Efflux-mediated drug resistance in bacteria. *Drugs* **64**, 159-204.

Li, X.Z., and Nikaido, H. (2009). Efflux-mediated drug resistance in bacteria: an update. *Drugs* **69**, 1555-1623.

Liguori, L., Marques, B., and Lenormand, J.L. (2008). A bacterial cell-free expression system to produce membrane proteins and proteoliposomes: from cDNA to functional assay. *Curr Protoc Protein Sci* **chap 5**, unit 5 22.

Limbago, B.M., Kallen, A.J., Zhu, W., Eggers, P., McDougal, L.K., and Albrecht, V.S. (2014). Report of the 13th vancomycin-resistant *Staphylococcus aureus* isolate from the United States. *J Clin Microbiol* **52**, 998-1002.

Lindsay, J.A., and Holden, M.T. (2004). *Staphylococcus aureus*: superbug, super genome? *Trends Microbiol* **12**, 378-385.

Linton, K.J. (2007). Structure and function of ABC transporters. *Physiology* **22**, 122-130.

Linton, K.J., and Higgins, C.F. (1998). The *Escherichia coli* ATP-binding cassette (ABC) proteins. *Mol Microbiol* **28**, 5-13.

Littlejohn, T.G., DiBerardino, D., Messerotti, L.J., Spiers, S.J., and Skurray, R.A. (1991). Structure and evolution of a family of genes encoding antiseptic and disinfectant resistance in *Staphylococcus aureus*. *Gene* **101**, 59-66.

Liu, R., and Sharom, F.J. (1996). Site-directed fluorescence labeling of P-glycoprotein on cysteine residues in the nucleotide binding domains. *Biochemistry* **35**, 11865-11873.

Liu, Z., Madej, M.G., and Kaback, H.R. (2010). Helix dynamics in LacY: helices II and IV. *J Mol Biol* **396**, 617-626.

Lloris-Garcera, P., Bianchi, F., Slusky, J.S.G., Seppala, S., Daley, D.O., and von Heijne, G. (2012). Antiparallel dimers of the small multidrug resistance protein EmrE are more stable than parallel dimers. *J Biol Chem* **287**, 26052-26059.

Lloris-Garcera, P., Seppala, S., Slusky, J.S., Rapp, M., and von Heijne, G. (2014). Why have small multidrug resistance proteins not evolved into fused, internally duplicated structures? *J Mol Biol* **426**, 2246-2254.

Locher, K.P. (2009). Review. Structure and mechanism of ATP-binding cassette transporters. *Philos Trans R Soc Lond B Biol Sci* **364**, 239-245.

Locher, K.P., Lee, A.T., and Rees, D.C. (2002). The *E. coli* BtuCD structure: a framework for ABC transporter architecture and mechanism. *Science* **296**, 1091-1098.

Lomovskaya, O., Warren, M.S., Lee, A., Galazzo, J., Fronko, R., Lee, M., Blais, J., Cho, D., Chamberland, S., Renau, T.*, et al.* (2001). Identification and characterization of inhibitors of multidrug resistance efflux pumps in *Pseudomonas aeruginosa*: novel agents for combination therapy. *Antimicrob Agents Chemother* **45**, 105-116.

Loo, T.W., Bartlett, M.C., and Clarke, D.M. (2008). Processing mutations disrupt interactions between the nucleotide binding and transmembrane domains of P-glycoprotein and the cystic fibrosis transmembrane conductance regulator (CFTR). *J Biol Chem* **283**, 28190-28197.

Loo, T.W., and Clarke, D.M. (1995). Membrane topology of a cysteine-less mutant of human P-glycoprotein. *J Biol Chem* **270**, 843-848.

Loo, T.W., and Clarke, D.M. (1999). Identification of residues in the drug-binding domain of human P-glycoprotein. Analysis of transmembrane segment 11 by cysteine-scanning mutagenesis and inhibition by dibromobimane. *J Biol Chem* **274**, 35388-35392.

Lowy, F.D. (2003). Antimicrobial resistance: the example of *Staphylococcus aureus*. *J Clin Invest* **111**, 1265-1273.

Lu, B., Kiessling, V., Tamm, L.K., and Cafiso, D.S. (2014). The juxtamembrane linker of full-length synaptotagmin 1 controls oligomerization and calcium-dependent membrane binding. *J Biol Chem* **289**, 22161-22171.

Lu, M., Radchenko, M., Symersky, J., Nie, R., and Guo, Y. (2013a). Structural insights into H+-coupled multidrug extrusion by a MATE transporter. *Nat Struct Mol Biol* **20**, 1310-1317.

Lu, M., Symersky, J., Radchenko, M., Koide, A., Guo, Y., Nie, R., and Koide, S. (2013b). Structures of a Na+-coupled, substrate-bound MATE multidrug transporter. *Proc Natl Acad Sci USA* **110**, 2099-2104.

Lundbald, R.L. (2015). Chemical Reagents for Protein Modification, Vol 4 (CRC Press, Boca Raton, FL).

Luthy, R., Bowie, J.U., and Eisenberg, D. (1992). Assessment of protein models with three-dimensional profiles. *Nature* **356**, 83-85.

Lyon, B.R., and Skurray, R. (1987). Antimicrobial resistance of *Staphylococcus aureus*: genetic basis. *Microbiol Rev* **51**, 88-134.

MacBeath, G., and Schreiber, S.L. (2000). Printing proteins as microarrays for high-throughput function determination. *Science* **289**, 1760-1763.

Madej, M.G., Dang, S., Yan, N., and Kaback, H.R. (2013). Evolutionary mix-and-match with MFS transporters. *Proc Natl Acad Sci USA* **110**, 5870-5874.

Maiorov, V.N., and Crippen, G.M. (1994). Significance of root-mean-square deviation in comparing three-dimensional structures of globular proteins. *J Mol Biol* **235**, 625-634.

Makrides, S.C. (1996). Strategies for achieving high-level expression of genes in *Escherichia coli*. *Microbiol Rev* **60**, 512-538.

Marisch, K., Bayer, K., Scharl, T., Mairhofer, J., Krempl, P.M., Hummel, K., Razzazi-Fazeli, E., and Striedner, G. (2013). A comparative analysis of industrial *Escherichia coli* K-12 and B strains in high-glucose batch cultivations on process-, transcriptome- and proteome level. *PLoS One* **8**, e70516.

Martin, C., Berridge, G., Mistry, P., Higgins, C., Charlton, P., and Callaghan, R. (2000). Drug binding sites on P-glycoprotein are altered by ATP binding prior to nucleotide hydrolysis. *Biochemistry* **39**, 11901-11906.

Martin, J., Gibrat, J.F., and Rodolphe, F. (2006). Analysis of an optimal hidden Markov model for secondary structure prediction. *BMC Struct Biol* **6**, 25.

Martinez Molina, D., Lundback, A.K., Niegowski, D., and Eshaghi, S. (2008). Expression and purification of the recombinant membrane protein YidC: a case study for increased stability and solubility. *Protein Expr Purif* **62**, 49-52.

Masaoka, Y., Ueno, Y., Morita, Y., Kuroda, T., Mizushima, T., and Tsuchiya, T. (2000). A two-component multidrug efflux pump, EbrAB, in *Bacillus subtilis*. *J Bacteriol* **182**, 2307-2310.

McAleese, F., Petersen, P., Ruzin, A., Dunman, P.M., Murphy, E., Projan, S.J., and Bradford, P.A. (2005). A novel MATE family efflux pump contributes to the reduced susceptibility of laboratory-derived *Staphylococcus aureus* mutants to tigecycline. *Antimicrob Agents Chemother* **49**, 1865-1871.

McDougal, L.K., Fosheim, G.E., Nicholson, A., Bulens, S.N., Limbago, B.M., Shearer, J.E., Summers, A.O., and Patel, J.B. (2010). Emergence of resistance among USA300 methicillin-resistant *Staphylococcus aureus* isolates causing invasive disease in the United States. *Antimicrob Agents Chemother* **54**, 3804-3811.

McLachlin, D.T., and Dunn, S.D. (1996). A method of screening for mutant proteins containing cysteine residues using fluorescein-5-maleimide. *Protein Expr Purif* **7**, 275-280.

McMurry, L., Petrucci, R.E., Jr., and Levy, S.B. (1980). Active efflux of tetracycline encoded by four genetically different tetracycline resistance determinants in *Escherichia coli*. *Proc Natl Acad Sci USA* **77**, 3974-3977.

Melnyk, R.A., Kim, S., Curran, A.R., Engelman, D.M., Bowie, J.U., and Deber, C.M. (2004). The affinity of GXXXG motifs in transmembrane helix-helix interactions is modulated by long-range communication. *J Biol Chem* **279**, 16591-16597.

Mendrola, J.M., Berger, M.B., King, M.C., and Lemmon, M.A. (2002). The single transmembrane domains of ErbB receptors self-associate in cell membranes. *J Biol Chem* **277**, 4704-4712.

Michel, H., Weyer, K.A., Gruenberg, H., Dunger, I., Oesterhelt, D., and Lottspeich, F. (1986). The 'light' and 'medium' subunits of the photosynthetic reaction centre from *Rhodopseudomonas viridis*: isolation of the genes, nucleotide and amino acid sequence. *EMBO J* **5**, 1149-1158.

Miroux, B., and Walker, J.E. (1996). Over-production of proteins in *Escherichia coli*: mutant hosts that allow synthesis of some membrane proteins and globular proteins at high levels. *J Mol Biol* **260**, 289-298.

Mitchell, B.A., Brown, M.H., and Skurray, R.A. (1998). QacA multidrug efflux pump from *Staphylococcus aureus*: comparative analysis of resistance to diamidines, biguanidines, and guanylhydrazones. *Antimicrob Agents Chemother* **42**, 475-477.

Mitchell, B.A., Paulsen, I.T., Brown, M.H., and Skurray, R.A. (1999). Bioenergetics of the staphylococcal multidrug export protein QacA. Identification of distinct binding sites for monovalent and divalent cations. *J Biol Chem* **274**, 3541-3548.

Mitchell, D.C. (2012). Progress in understanding the role of lipids in membrane protein folding. *Biochim Biophys Acta* **1818**, 951-956.

Molday, L.L., and Molday, R.S. (2014). 1D4: a versatile epitope tag for the purification and characterization of expressed membrane and soluble proteins. *Methods Mol Biol* **1177**, 1-15.

Mordoch, S.S., Granot, D., Lebendiker, M., and Schuldiner, S. (1999). Scanning cysteine accessibility of EmrE, an H+-coupled multidrug transporter from *Escherichia coli*, reveals a hydrophobic pathway for solutes. *J Biol Chem* **274**, 19480-19486.

Morimyo, M., Hongo, E., Hama-Inaba, H., and Machida, I. (1992). Cloning and characterization of the *mvrC* gene of *Escherichia coli* K-12 which confers resistance against methyl viologen toxicity. *Nucleic Acids Res* **20**, 3159-3165.

Morita, Y., Kodama, K., Shiota, S., Mine, T., Kataoka, A., Mizushima, T., and Tsuchiya, T. (1998). NorM, a putative multidrug efflux protein, of *Vibrio parahaemolyticus* and its homolog in *Escherichia coli*. *Antimicrob Agents Chemother* **42**, 1778-1782.

Moriyama, Y., Hiasa, M., Matsumoto, T., and Omote, H. (2008). Multidrug and toxic compound extrusion (MATE)-type proteins as anchor transporters for the excretion of metabolic waste products and xenobiotics. *Xenobiotica* **38**, 1107-1118.

Morris, A.L., MacArthur, M.W., Hutchinson, E.G., and Thornton, J.M. (1992). Stereochemical quality of protein structure coordinates. *Proteins* **12**, 345-364.

Morrison, E.A., DeKoster, G.T., Dutta, S., Vafabakhsh, R., Clarkson, M.W., Bahl, A., Kern, D., Ha, T., and Henzler-Wildman, K.A. (2012). Antiparallel EmrE exports drugs by exchanging between asymmetric structures. *Nature* **481**, 45-U50.

Morrison, E.A., and Henzler-Wildman, K.A. (2014). Transported Substrate Determines Exchange Rate in the Multidrug Resistance Transporter EmrE. *J Biol Chem* **289**, 6825-6836.

Moult, J. (2005). A decade of CASP: progress, bottlenecks and prognosis in protein structure prediction. *Curr Opin Struct Biol* **15**, 285-289.

Moult, J., Fidelis, K., Kryshtafovych, A., Schwede, T., and Tramontano, A. (2014). Critical assessment of methods of protein structure prediction (CASP)--round x. *Proteins* **82 Suppl 2**, 1-6.

Mueckler, M., and Makepeace, C. (2002). Analysis of transmembrane segment 10 of the Glut1 glucose transporter by cysteine-scanning mutagenesis and substituted cysteine accessibility. *J Biol Chem* **277**, 3498-3503.

Mueckler, M., and Makepeace, C. (2005). Cysteine-scanning mutagenesis and substituted cysteine accessibility analysis of transmembrane segment 4 of the Glut1 glucose transporter. *J Biol Chem* **280**, 39562-39568.

Murakami, S., Nakashima, R., Yamashita, E., Matsumoto, T., and Yamaguchi, A. (2006). Crystal structures of a multidrug transporter reveal a functionally rotating mechanism. *Nature* **443**, 173-179.

Murakami, S., Nakashima, R., Yamashita, E., and Yamaguchi, A. (2002). Crystal structure of bacterial multidrug efflux transporter AcrB. *Nature* **419**, 587-593.

Murakami, S., Tamura, N., Saito, A., Hirata, T., and Yamaguchi, A. (2004). Extramembrane central pore of multidrug exporter AcrB in *Escherichia coli* plays an important role in drug transport. *J Biol Chem* **279**, 3743-3748.

Murray, D.S., Schumacher, M.A., and Brennan, R.G. (2004). Crystal structures of QacR-diamidine complexes reveal additional multidrug-binding modes and a novel mechanism of drug charge neutralization. *J Biol Chem* **279**, 14365-14371.

Muth, T.R., and Schuldiner, S. (2000). A membrane-embedded glutamate is required for ligand binding to the multidrug transporter EmrE. *EMBO J* **19**, 234-240.

Nadeau, O.W., and Carlson, G.M. (2007). Protein interactions captured by chemical cross-linking: One-step cross-linking with formaldehyde. *CSH protocols* **2007**, pdb prot4634.

Nara, T., Kouyama, T., Kurata, Y., Kikukawa, T., Miyauchi, S., and Kamo, N. (2007). Anti-parallel membrane topology of a homo-dimeric multidrug transporter, EmrE. *J Biochem* **142**, 621-625.

Narui, K., Noguchi, N., Wakasugi, K., and Sasatsu, M. (2002). Cloning and characterization of a novel chromosomal drug efflux gene in *Staphylococcus aureus*. *Biol Pharm Bull* **25**, 1533-1536.

Nelson, M.L., and Levy, S.B. (1999). Reversal of tetracycline resistance mediated by different bacterial tetracycline resistance determinants by an inhibitor of the Tet(B) antiport protein. *Antimicrob Agents Chemother* **43**, 1719-1724.

Neyfakh, A.A., Borsch, C.M., and Kaatz, G.W. (1993). Fluoroquinolone resistance protein NorA of *Staphylococcus aureus* is a multidrug efflux transporter. *Antimicrob Agents Chemother* **37**, 128-129.

Niegowski, D., Hedren, M., Nordlund, P., and Eshaghi, S. (2006). A simple strategy towards membrane protein purification and crystallization. *Int J Biol Macromol* **39**, 83-87.

Nikaido, H. (1994). Prevention of drug access to bacterial targets: permeability barriers and active efflux. *Science* **264**, 382-388.

Nikaido, H. (1996). Multidrug efflux pumps of gram-negative bacteria. *J Bacteriol* **178**, 5853-5859.

Nikaido, H. (1998). Multiple antibiotic resistance and efflux. *Curr Opin Microbiol* **1**, 516-523.

Nikaido, H., and Takatsuka, Y. (2009). Mechanisms of RND multidrug efflux pumps. *Biochim Biophys Acta* **1794**, 769-781.

Ninio, S., Rotem, D., and Schuldiner, S. (2001). Functional analysis of novel multidrug transporters from human pathogens. *J Biol Chem* **276**, 48250-48256.

Ninio, S., and Schuldiner, S. (2003). Characterization of an archaeal multidrug transporter with a unique amino acid composition. *J Biol Chem* **278**, 12000-12005.

Noguchi, N., Hase, M., Kitta, M., Sasatsu, M., Deguchi, K., and Kono, M. (1999). Antiseptic susceptibility and distribution of antiseptic-resistance genes in methicillin-resistant *Staphylococcus aureus*. *FEMS Microbiol Lett* **172**, 247-253.

Oldham, M.L., Khare, D., Quiocho, F.A., Davidson, A.L., and Chen, J. (2007). Crystal structure of a catalytic intermediate of the maltose transporter. *Nature* **450**, 515-521.

Oliveira, D.C., Tomasz, A., and de Lencastre, H. (2002). Secrets of success of a human pathogen: molecular evolution of pandemic clones of meticillin-resistant *Staphylococcus aureus*. *Lancet Infect Dis* **2**, 180-189.

Omote, H., Hiasa, M., Matsumoto, T., Otsuka, M., and Moriyama, Y. (2006). The MATE proteins as fundamental transporters of metabolic and xenobiotic organic cations. *Trends Pharmacol Sci* **27**, 587-593.

Ong, Y.S., Lakatos, A., Becker-Baldus, J., Pos, K.M., and Glaubitz, C. (2013). Detecting substrates bound to the secondary multidrug efflux pump EmrE by DNP-enhanced solid-state NMR. *J Am Chem Soc* **135**, 15754-15762.

Otsuka, M., Yasuda, M., Morita, Y., Otsuka, C., Tsuchiya, T., Omote, H., and Moriyama, Y. (2005). Identification of essential amino acid residues of the NorM Na+/multidrug antiporter in *Vibrio parahaemolyticus*. *J Bacteriol* **187**, 1552-1558.

Padariya, M., Kalathiya, U., and Baginski, M. (2015). Structural and dynamic changes adopted by EmrE, multidrug transporter protein-Studies by molecular dynamics simulation. *Biochim Biophys Acta*.

Pakula, A.A., and Simon, M.I. (1992). Determination of transmembrane protein structure by disulfide cross-linking: the *Escherichia coli* Tar receptor. *Proc Natl Acad Sci USA* **89**, 4144-4148.

Pantosti, A., Sanchini, A., and Monaco, M. (2007). Mechanisms of antibiotic resistance in *Staphylococcus aureus*. *Future Microbiol* **2**, 323-334.

Papadopoulos, M., and Morona, R. (2010). Mutagenesis and chemical cross-linking suggest that Wzz dimer stability and oligomerization affect lipopolysaccharide O-antigen modal chain length control. *J Bacteriol* **192**, 3385-3393.

Patel, D., Kosmidis, C., Seo, S.M., and Kaatz, G.W. (2010). Ethidium bromide MIC screening for enhanced efflux pump gene expression or efflux activity in *Staphylococcus aureus*. *Antimicrob Agents Chemother* **54**, 5070-5073.

Paulsen, I.T., Brown, M.H., Dunstan, S.J., and Skurray, R.A. (1995). Molecular characterization of the staphylococcal multidrug resistance export protein QacC. *J Bacteriol* **177**, 2827-2833.

Paulsen, I.T., Brown, M.H., Littlejohn, T.G., Mitchell, B.A., and Skurray, R.A. (1996a). Multidrug resistance proteins QacA and QacB from *Staphylococcus aureus*: membrane topology and identification of residues involved in substrate specificity. *Proc Natl Acad Sci USA* **93**, 3630-3635.

Paulsen, I.T., Brown, M.H., and Skurray, R.A. (1996b). Proton-dependent multidrug efflux systems. *Microbiol Rev* **60**, 575-608.

Paulsen, I.T., Brown, M.H., and Skurray, R.A. (1998). Characterization of the earliest known *Staphylococcus aureus* plasmid encoding a multidrug efflux system. *J Bacteriol* **180**, 3477-3479.

Paulsen, I.T., Chen, J., Nelson, K.E., and Saier, M.H. (2001). Comparative genomics of microbial drug efflux systems. *J Mol Microb Biotech* **3**, 145-150.

Paulsen, I.T., Littlejohn, T.G., Radstrom, P., Sundstrom, L., Skold, O., Swedberg, G., and Skurray, R.A. (1993). The 3' conserved segment of integrons contains a gene associated with multidrug resistance to antiseptics and disinfectants. *Antimicrob Agents Chemother* **37**, 761-768.

Paulsen, I.T., Skurray, R.A., Tam, R., Saler, M.H., Turner, R.J., Weiner, J.H., Goldberg, E.B., and Grinius, L.L. (1996c). The SMR family: A novel family of multidrug efflux proteins involved with the efflux of lipophilic drugs. *Mol Microbiol* **19**, 1167-1175.

Pe'er, I., Felder, C.E., Man, O., Silman, I., Sussman, J.L., and Beckmann, J.S. (2004). Proteomic signatures: amino acid and oligopeptide compositions differentiate among phyla. *Proteins* **54**, 20-40.

Peters, K.M., Brooks, B.E., Schumacher, M.A., Skurray, R.A., Brennan, R.G., and Brown, M.H. (2011). A single acidic residue can guide binding site selection but does not govern QacR cationic-drug affinity. *PLoS One* **6**, e15974.

Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C., and Ferrin, T.E. (2004). UCSF Chimera--a visualization system for exploratory research and analysis. *J Comput Chem* **25**, 1605-1612.

Piddock, L.J. (2006). Multidrug-resistance efflux pumps - not just for resistance. *Nat Rev Microbiol* **4**, 629-636.

Pluchino, K.M., Hall, M.D., Goldsborough, A.S., Callaghan, R., and Gottesman, M.M. (2012). Collateral sensitivity as a strategy against cancer multidrug resistance. *Drug Resist Updat* **15**, 98-105.

Poget, S.F., Harris, R., Cahill, S.M., and Girvin, M.E. (2010). ^{1}H, ^{13}C, ^{15}N backbone NMR assignments of the *Staphylococcus aureus* small multidrug-resistance pump (Smr) in a functionally active conformation. *Biomol NMR Assign* **4**, 139-142.

Poole, K. (2002). Mechanisms of bacterial biocide and antibiotic resistance. *J Appl Microbiol* **92 Suppl**, 55S-64S.

Poulsen, B.E., Rath, A., and Deber, C.M. (2009). The assembly motif of a bacterial small multidrug resistance protein. *J Biol Chem* **284**, 9870-9875.

Prax, M., Lee, C.Y., and Bertram, R. (2013). An update on the molecular genetics toolbox for staphylococci. *Microbiology* **159**, 421-435.

Prive, G.G. (2007). Detergents for the stabilization and crystallization of membrane proteins. *Methods* **41**, 388-397.

Procko, E., O'Mara, M.L., Bennett, W.F., Tieleman, D.P., and Gaudet, R. (2009). The mechanism of ABC transporters: general lessons from structural and functional studies of an antigenic peptide transporter. *FASEB J* **23**, 1287-1302.

Punta, M., Forrest, L.R., Bigelow, H., Kernytsky, A., Liu, J., and Rost, B. (2007). Membrane protein prediction methods. *Methods* **41**, 460-474.

Purewal, A.S. (1991). Nucleotide sequence of the ethidium efflux gene from *Escherichia coli*. *FEMS Microbiol Lett* **66**, 229-231.

Radford, S.E. (2006). GroEL: More than Just a folding cage. *Cell* **125**, 831-833.

Raine, A., Ullers, R., Pavlov, M., Luirink, J., Wikberg, J.E., and Ehrenberg, M. (2003). Targeting and insertion of heterologous membrane proteins in *E. coli*. *Biochimie* **85**, 659-668.

Ramachandran, G.N., Ramakrishnan, C., and Sasisekharan, V. (1963). Stereochemistry of polypeptide chain configurations. *J Mol Biol* **7**, 95-99.

Rapp, M., Granseth, E., Seppala, S., and von Heijne, G. (2006). Identification and evolution of dual-topology membrane proteins. *Nat Struct Mol Biol* **13**, 112-116.

Rapp, M., Seppala, S., Granseth, E., and von Heijne, G. (2007). Emulating membrane protein evolution by rational design. *Science* **315**, 1282-1284.

Rath, A., Melnyk, R.A., and Deber, C.M. (2006). Evidence for assembly of small multidrug resistance proteins by a "two-faced" transmembrane helix. *J Biol Chem* **281**, 15546-15553.

Raviv, Y., Pollard, H.B., Bruggemann, E.P., Pastan, I., and Gottesman, M.M. (1990). Photosensitized labeling of a functional multidrug transporter in living drug-resistant tumor cells. *J Biol Chem* **265**, 3975-3980.

Reddy Ch, S., Vijayasarathy, K., Srinivas, E., Sastry, G.M., and Sastry, G.N. (2006). Homology modeling of membrane proteins: a critical assessment. *Comput Biol Chem* **30**, 120-126.

Reddy, V.S., Shlykov, M.A., Castillo, R., Sun, E.I., and Saier, M.H., Jr. (2012). The major facilitator superfamily (MFS) revisited. *FEBS J* **279**, 2022-2035.

Rees, D.C., Johnson, E., and Lewinson, O. (2009). ABC transporters: the power to change. *Nat Rev Mol Cell Biol* **10**, 218-227.

Rost, B., Casadio, R., and Fariselli, P. (1996a). Refining neural network predictions for helical transmembrane proteins by dynamic programming. *Proc Int Conf Intell Syst Mol Biol* **4**, 192-200.

Rost, B., Fariselli, P., and Casadio, R. (1996b). Topology prediction for helical transmembrane proteins at 86% accuracy. *Protein Sci* **5**, 1704-1718.

Rotem, D., Sal-man, N., and Schuldiner, S. (2001). In vitro monomer swapping in EmrE, a multidrug transporter from *Escherichia coli*, reveals that the oligomer is the functional unit. *J Biol Chem* **276**, 48243-48249.

Rotem, D., and Schuldiner, S. (2004). EmrE, a multidrug transporter from *Escherichia coli,* transports monovalent and divalent substrates with the same stoichiometry. *J Biol Chem* **279**, 48787-48793.

Rotem, D., Steiner-Mordoch, S., and Schuldiner, S. (2006). Identification of tyrosine residues critical for the function of an ion-coupled multidrug transporter. *J Biol Chem* **281**, 18715-18722.

Sahin-Tóth, M., Persson, B., Schwieger, J., Cohan, P., and Kaback, H.R. (1994). Cysteine scanning mutagenesis of the N-terminal 32 amino acid residues in the lactose permease of *Escherichia coli*. *Protein Sci* **3**, 240-247.

Saier, M.H., Jr. (2003). Tracing pathways of transport protein evolution. *Mol Microbiol* **48**, 1145-1156.

Saier, M.H., Jr., Beatty, J.T., Goffeau, A., Harley, K.T., Heijne, W.H., Huang, S.C., Jack, D.L., Jahn, P.S., Lew, K., Liu, J.*, et al.* (1999). The major facilitator superfamily. *J Mol Microbiol Biotechnol* **1**, 257-279.

Saier, M.H., Jr., and Paulsen, I.T. (2001). Phylogeny of multidrug transporters. *Semin Cell Dev Biol* **12**, 205-213.

Saiful, A.J., Mastura, M., Zarizal, S., Mazurah, M.I., Shuhaimi, M., and Ali, A.M. (2008). Efflux genes and active efflux activity detection in Malaysian clinical isolates of methicillin-resistant *Staphylococcus aureus* (MRSA). *J Basic Microbiol* **48**, 245-251.

Sali, A., and Blundell, T.L. (1993). Comparative protein modelling by satisfaction of spatial restraints. *J Mol Biol* **234**, 779-815.

Sambrook, J., and Russell, D.W. (2001). Molecular cloning: a lab manual (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York).

Sanchez, R., and Sali, A. (2000). Comparative protein structure modeling. Introduction and practical examples with modeller. *Method Mol Biol* **143**, 97-129.

Sandercock, A. (2010). Isoaltion and characteriation of the novel *Staphylococcus aureus* SepA protein: A mutlidrug transporter. School of Biological Sciences (Adelaide, SA. Flinders University), pp. 71.

Saravolatz, L.D., Pawlak, J., and Johnson, L.B. (2012). In vitro susceptibilities and molecular analysis of vancomycin-intermediate and vancomycin-resistant *Staphylococcus aureus* isolates. *Clin Infect Dis* **55**, 582-586.

Sasatsu, M., Shima, K., Shibata, Y., and Kono, M. (1989). Nucleotide sequence of a gene that encodes resistance to ethidium bromide from a transferable plasmid in *Staphylococcus aureus*. *Nucleic Acids Res* **17**, 10103.

Sasatsu, M., Shirai, Y., Hase, M., Noguchi, N., Kono, M., Behr, H., Freney, J., and Arai, T. (1995). The origin of the antiseptic-resistance gene *ebr* in *Staphylococcus aureus*. *Microbios* **84**, 161-169.

Schindler, B.D., Jacinto, P., and Kaatz, G.W. (2013a). Inhibition of drug efflux pumps in *Staphylococcus aureus*: current status of potentiating existing antibiotics. *Future Microbiol* **8**, 491-507.

Schindler, B.D., Patel, D., Seo, S.M., and Kaatz, G.W. (2013b). Mutagenesis and modeling to predict structural and functional characteristics of the *Staphylococcus aureus* MepA multidrug efflux pump. *J Bacteriol* **195**, 523-533.

Schindler, B.D., Seo, S.M., Birukou, I., Brennan, R.G., and Kaatz, G.W. (2015). Mutations within the mepA operator affect binding of the MepR regulatory protein and its induction by MepA substrates in *Staphylococcus aureus*. *J Bacteriol* **197**, 1104-1114.

Schito, G.C. (2006). The importance of the development of antibiotic resistance in *Staphylococcus aureus*. *Clin Microbiol Infect* **12 Suppl 1**, 3-8.

Schuldiner, S. (2007). When biochemistry meets structural biology: the cautionary tale of EmrE. *Trends Biochem Sci* **32**, 252-258.

Schuldiner, S. (2009). EmrE, a model for studying evolution and mechanism of ion-coupled transporters. *Biochim Biophys Acta* **1794**, 748-762.

Schuldiner, S. (2012). Undecided membrane proteins insert in random topologies. Up, down and sideways: it does not really matter. *Trends Biochem Sci* **37**, 215-219.

Schuldiner, S., Granot, D., Mordoch, S.S., Ninio, S., Rotem, D., Soskin, M., Tate, C.G., and Yerushalmi, H. (2001a). Small is mighty: EmrE, a multidrug transporter as an experimental paradigm. *News Physiol Sci* **16**, 130-134.

Schuldiner, S., Granot, D., Steiner, S., Ninio, S., Rotem, D., Soskin, M., and Yerushalmi, H. (2001b). Precious things come in little packages. *J Mol Microb Biotech* **3**, 155-162.

Schuldiner, S., Lebendiker, M., and Yerushalmi, H. (1997). EmrE, the smallest ion-coupled transporter, provides a unique paradigm for structure-function studies. *J Exp Biol* **200**, 335-341.

Schumacher, A., Trittler, R., Bohnert, J.A., Kummerer, K., Pages, J.M., and Kern, W.V. (2007). Intracellular accumulation of linezolid in *Escherichia coli*, *Citrobacter freundii* and *Enterobacter aerogene*s: role of enhanced efflux pump activity and inactivation. *J Antimicrob Chemother* **59**, 1261-1264.

Schumacher, M.A., and Brennan, R.G. (2003). Deciphering the molecular basis of multidrug recognition: crystal structures of the *Staphylococcus aureus* multidrug binding transcription regulator QacR. *Res Microbiol* **154**, 69-77.

Schumacher, M.A., Miller, M.C., and Brennan, R.G. (2004). Structural mechanism of the simultaneous binding of two drugs to a multidrug-binding protein. *EMBO J* **23**, 2923-2930.

Schumacher, M.A., Miller, M.C., Grkovic, S., Brown, M.H., Skurray, R.A., and Brennan, R.G. (2001). Structural mechanisms of QacR induction and multidrug recognition. *Science* **294**, 2158-2163.

Schwaiger, M., Lebendiker, M., Yerushalmi, H., Coles, M., Groger, A., Schwarz, C., Schuldiner, S., and Kessler, H. (1998). NMR investigation of the multidrug transporter EmrE, an integral membrane protein. *Eur J Biochem* **254**, 610-619.

Scott, D.J., Kummer, L., Tremmel, D., and Pluckthun, A. (2013). Stabilizing membrane proteins through protein engineering. *Curr Opin Chem Biol* **17**, 427-435.

Seddon, A.M., Curnow, P., and Booth, P.J. (2004). Membrane proteins, lipids and detergents: not just a soap opera. *Biochim Biophys Acta* **1666**, 105-117.

Seeger, M.A., Diederichs, K., Eicher, T., Brandstatter, L., Schiefner, A., Verrey, F., and Pos, K.M. (2008). The AcrB efflux pump: conformational cycling and peristalsis lead to multidrug resistance. *Curr Drug Targets* **9**, 729-749.

Seeger, M.A., Schiefner, A., Eicher, T., Verrey, F., Diederichs, K., and Pos, K.M. (2006). Structural asymmetry of AcrB trimer suggests a peristaltic pump mechanism. *Science* **313**, 1295-1298.

Senes, A., Engel, D.E., and DeGrado, W.F. (2004). Folding of helical membrane proteins: the role of polar, GxxxG-like and proline motifs. *Curr Opin Struct Biol* **14**, 465-479.

Senes, A., Ubarretxena-Belandia, I., and Engelman, D.M. (2001). The Calpha ---H...O hydrogen bond: a determinant of stability and specificity in transmembrane helix interactions. *Proc Natl Acad Sci USA* **98**, 9056-9061.

Sennhauser, G., Bukowska, M.A., Briand, C., and Grutter, M.G. (2009). Crystal structure of the multidrug exporter MexB from *Pseudomonas aeruginosa*. *J Mol Biol* **389**, 134-145.

Seppala, S., Slusky, J.S., Lloris-Garcera, P., Rapp, M., and von Heijne, G. (2010). Control of membrane protein topology by a single C-terminal residue. *Science* **328**, 1698-1700.

Sharoni, M., Steiner-Mordoch, S., and Schuldiner, S. (2005). Exploring the binding domain of EmrE, the smallest multidrug transporter. *J Biol Chem* **280**, 32849-32855.

Shimizu, T., Mitsuke, H., Noto, K., and Arai, M. (2004). Internal gene duplication in the evolution of prokaryotic transmembrane proteins. *J Mol Biol* **339**, 1-15.

Shinefield, H.R., and Ruff, N.L. (2009). Staphylococcal infections: a historical perspective. *Infect Dis Clin N Am* **23**, 1-15.

Sigal, N., Vardy, E., Molshanski-Mor, S., Eitan, A., Pilpel, Y., Schuldiner, S., and Bibi, E. (2005). 3D model of the *Escherichia coli* multidrug transporter MdfA reveals an essential membrane-embedded positive charge. *Biochemistry* **44**, 14870-14880.

Sikora, C.W., and Turner, R.J. (2005). SMR proteins SugE and EmrE bind ligand with similar affinity and stoichiometry. *Biochem Biophys Res Commun* **335**, 105-111.

Skare, J.T., Ahmer, B.M., Seachord, C.L., Darveau, R.P., and Postle, K. (1993). Energy transduction between membranes. TonB, a cytoplasmic membrane protein, can be chemically cross-linked in vivo to the outer membrane receptor FepA. *J Biol Chem* **268**, 16302-16308.

Skinner, D., and Keefer, C.S. (1941). Significance of bacteremia caused by *Staphylococcus aureus* - A study of one hundred and twenty-two cases and a review of the literature concerned wiih experimental infection in animals. *Arch Intern Med* **68**, 851-875.

Son, M.S., Del Castilho, C., Duncalf, K.A., Carney, D., Weiner, J.H., and Turner, R.J. (2003). Mutagenesis of SugE, a small multidrug resistance protein. *Biochem Biophy Res Co* **312**, 914-921.

Sorensen, H.P., and Mortensen, K.K. (2005). Advanced genetic strategies for recombinant protein expression in *Escherichia coli*. *J Biotechnol* **115**, 113-128.

Soskine, M., Adam, Y., and Schuldiner, S. (2004). Direct evidence for substrate-induced proton release in detergent-solubilized EmrE, a multidrug transporter. *J Biol Chem* **279**, 9951-9955.

Soskine, M., Mark, S., Tayer, N., Mizrachi, R., and Schuldiner, S. (2006). On parallel and antiparallel topology of a homodimeric multidrug transporter. *J Biol Chem* **281**, 36205-36212.

Soskine, M., Steiner-Mordoch, S., and Schuldiner, S. (2002). Crosslinking of membrane-embedded cysteines reveals contact points in the EmrE oligomer. *Proc Natl Acad Sci USA* **99**, 12043-12048.

Soubias, O., and Gawrisch, K. (2012). The role of the lipid matrix for structure and function of the GPCR rhodopsin. *Biochim Biophys Acta* **1818**, 234-240.

Srinivasan, V.B., and Rajamohan, G. (2013). KpnEF, a new member of the *Klebsiella pneumoniae* cell envelope stress response regulon, is an SMR-type efflux pump involved in broad-spectrum antimicrobial resistance. *Antimicrob Agents Chemother* **57**, 4449-4462.

Srinivasan, V.B., Rajamohan, G., and Gebreyes, W.A. (2009). Role of AbeS, a novel efflux pump of the SMR family of transporters, in resistance to antimicrobial agents in *Acinetobacter baumannii*. *Antimicrob Agents Chemother* **53**, 5312-5316.

Steiner-Mordoch, S., Soskine, M., Solomon, D., Rotem, D., Gold, A., Yechieli, M., Adam, Y., and Schuldiner, S. (2008). Parallel topology of genetically fused EmrE homodimers. *EMBO J* **27**, 17-26.

Stemmer, W.P., Crameri, A., Ha, K.D., Brennan, T.M., and Heyneker, H.L. (1995). Single-step assembly of a gene and entire plasmid from large numbers of oligodeoxyribonucleotides. *Gene* **164**, 49-53.

Strop, P., and Brunger, A.T. (2005). Refractive index-based determination of detergent concentration and its application to the study of membrane proteins. *Protein Sci* **14**, 2207-2211.

Studier, F.W., and Moffatt, B.A. (1986). Use of bacteriophage-T7 RNA-polymerase to direct selective high-level expression of cloned genes. *J Mol Biol* **189**, 113-130.

Sun, J., Deng, Z., and Yan, A. (2014). Bacterial multidrug efflux pumps: mechanisms, physiology and pharmacological exploitations. *Biochem Biophys Res Commun* **453**, 254-267.

Suriyanarayanan, B., and Sarojini Santhosh, R. (2014). Docking analysis insights quercetin can be a non-antibiotic adjuvant by inhibiting Mmr drug efflux pump in Mycobacterium sp. and its homologue EmrE in *Escherichia coli*. *J Biomol Struct Dyn*, 1-16.

Sutherland, B.W., Toews, J., and Kast, J. (2008). Utility of formaldehyde cross-linking and mass spectrometry in the study of protein-protein interactions. *J Mass Spectrom* **43**, 699-715.

Takatsuka, Y., and Nikaido, H. (2006). Threonine-978 in the transmembrane segment of the multidrug efflux pump AcrB of *Escherichia coli* is crucial for drug transport as a probable component of the proton relay network. *J Bacteriol* **188**, 7284-7289.

Takatsuka, Y., and Nikaido, H. (2009). Covalently linked trimer of the AcrB multidrug efflux pump provides support for the functional rotating mechanism. *J Bacteriol* **191**, 1729-1737.

Tal, N., and Schuldiner, S. (2009). A coordinated network of transporters with overlapping specificities provides a robust survival strategy. *Proc Natl Acad Sci USA* **106**, 9051-9056.

Tanaka, Y., Hipolito, C.J., Maturana, A.D., Ito, K., Kuroda, T., Higuchi, T., Katoh, T., Kato, H.E., Hattori, M., Kumazaki, K.*, et al.* (2013). Structural basis for the drug extrusion mechanism by a MATE multidrug transporter. *Nature* **496**, 247-251.

Tartof, K., and Hobbs, C. (1987). Improved media for growing plasmid and cosmid clones. *Focus* **9**, 12.

Tate, C.G., Kunji, E.R., Lebendiker, M., and Schuldiner, S. (2001). The projection structure of EmrE, a proton-linked multidrug transporter from *Escherichia coli*, at 7 A resolution. *EMBO J* **20**, 77-81.

Tate, C.G., Ubarretxena-Belandia, I., and Baldwin, J.M. (2003). Conformational changes in the multidrug transporter EmrE associated with substrate binding. *J Mol Biol* **332**, 229-242.

Tennent, J.M., Lyon, B.R., Gillespie, M.T., May, J.W., and Skurray, R.A. (1985). Cloning and expression of *Staphylococcus aureus* plasmid-mediated quaternary ammonium resistance in *Escherichia coli*. *Antimicrob Agents Chemother* **27**, 79-83.

Thompson, J.D., Gibson, T.J., and Higgins, D.G. (2002). Multiple sequence alignment using ClustalW and ClustalX. *Curr Prot Bioinfo* **chap 2**, unit 2 3.

Tirosh, O., Sigal, N., Gelman, A., Sahar, N., Fluman, N., Siemion, S., and Bibi, E. (2012). Manipulating the drug/proton antiport stoichiometry of the secondary multidrug transporter MdfA. *Proc Natl Acad Sci USA* **109**, 12473-12478.

Tong, S.Y., Holden, M.T., Nickerson, E.K., Cooper, B.S., Koser, C.U., Cori, A., Jombart, T., Cauchemez, S., Fraser, C., Wuthiekanun, V.*, et al.* (2015). Genome sequencing defines phylogeny and spread of methicillin-resistant *Staphylococcus aureus* in a high transmission setting. *Genome Res* **25**, 111-118.

Tornroth-Horsefield, S., Gourdon, P., Horsefield, R., Brive, L., Yamamoto, N., Mori, H., Snijder, A., and Neutze, R. (2007). Crystal structure of AcrB in complex with a single transmembrane subunit reveals another twist. *Structure* **15**, 1663-1673.

Treutlein, H.R., Lemmon, M.A., Engelman, D.M., and Brunger, A.T. (1992). The glycophorin A transmembrane domain dimer: sequence-specific propensity for a right-handed supercoil of helices. *Biochemistry* **31**, 12726-12732.

Truong-Bolduc, Q.C., Dunman, P.M., Strahilevitz, J., Projan, S.J., and Hooper, D.C. (2005). MgrA is a multiple regulator of two new efflux pumps in *Staphylococcus aureus*. *J Bacteriol* **187**, 2395-2405.

Truong-Bolduc, Q.C., Strahilevitz, J., and Hooper, D.C. (2006). NorC, a new efflux pump regulated by MgrA of *Staphylococcus aureus*. *Antimicrob Agents Chemother* **50**, 1104-1107.

Tseng, T.T., Gratwick, K.S., Kollman, J., Park, D., Nies, D.H., Goffeau, A., and Saier, M.H., Jr. (1999). The RND permease superfamily: an ancient, ubiquitous and diverse family that includes human disease and development proteins. *J Mol Microbiol Biotechnol* **1**, 107-125.

Tusnady, G.E., Dosztanyi, Z., and Simon, I. (2004). Transmembrane proteins in the Protein Data Bank: identification and classification. *Bioinformatics* **20**, 2964-2972.

Tusnady, G.E., Dosztanyi, Z., and Simon, I. (2005). PDB_TM: selection and membrane localization of transmembrane proteins in the protein data bank. *Nucleic Acids Res* **33**, D275-278.

Ubarretxena-Belandia, I., Baldwin, J.M., Schuldiner, S., and Tate, C.G. (2003). Three-dimensional structure of the bacterial multidrug transporter EmrE shows it is an asymmetric homodimer. *EMBO J* **22**, 6175-6181.

Ubarretxena-Belandia, I., and Tate, C.G. (2004). New insights into the structure and oligomeric state of the bacterial multidrug transporter EmrE: an unusual asymmetric homo-dimer. *FEBS Lett* **564**, 234-238.

Ulmschneider, M.B., and Sansom, M.S. (2001). Amino acid distributions in integral membrane protein structures. *Biochim Biophys Acta* **1512**, 1-14.

Van Bambeke, F., Balzi, E., and Tulkens, P.M. (2000). Antibiotic efflux pumps. *Biochem Pharmacol* **60**, 457-470.

Vandenesch, F., Naimi, T., Enright, M.C., Lina, G., Nimmo, G.R., Heffernan, H., Liassine, N., Bes, M., Greenland, T., Reverdy, M.E.*, et al.* (2003). Community-acquired methicillin-resistant *Staphylococcus aureus* carrying Panton-Valentine leukocidin genes: worldwide emergence. *Emerg Infect Dis* **9**, 978-984.

Varela, M.F., Sansom, C.E., and Griffith, J.K. (1995). Mutational analysis and molecular modelling of an amino acid sequence motif conserved in antiporters but not symporters in a transporter superfamily. *Mol Membr Biol* **12**, 313-319.

Velamakanni, S., Yao, Y., Gutmann, D.A., and van Veen, H.W. (2008). Multidrug transport by the ABC transporter Sav1866 from *Staphylococcus aureus*. *Biochemistry* **47**, 9300-9308.

Vetrivel, U., and Subramanian, G. (2014). Importance of ABC transporters in different tissues. *Drug Metabol Drug Interact* **29**, 65-66.

Vinothkumar, K.R., Raunser, S., Jung, H., and Kuhlbrandt, W. (2006). Oligomeric structure of the carnitine transporter CaiT from *Escherichia coli*. *J Biol Chem* **281**, 4795-4801.

Vitrac, H., Bogdanov, M., and Dowhan, W. (2013). In vitro reconstitution of lipid-dependent dual topology and postassembly topological switching of a membrane protein. *Proc Natl Acad Sci USA* **110**, 9338-9343.

von Heijne, G. (1986). The distribution of positively charged residues in bacterial inner membrane proteins correlates with th trans-membrane topology. *EMBO J* **5**, 3021-3027.

von Heijne, G. (1991). Proline kinks in transmembrane alpha-helices. *J Mol Biol* **218**, 499-503.

von Heijne, G. (1992). Membrane Protein Structure Prediction: Hydrophobicity Analysis and the 'Positive Inside' Rule. *J Mol Biol* **225**, 487-494.

von Heijne, G. (2006). Membrane-protein topology. *Nat Rev Mol Cell Biol* **7**, 909-918.

Wagner, S., Baars, L., Ytterberg, A.J., Klussmeier, A., Wagner, C.S., Nord, O., Nygren, P.A., van Wijk, K.J., and de Gier, J.W. (2007). Consequences of membrane protein overexpression in *Escherichia coli*. *Mol Cell Proteomics* **6**, 1527-1550.

Wagner, S., Bader, M.L., Drew, D., and de Gier, J.W. (2006). Rationalizing membrane protein overexpression. *Trends Biotechnol* **24**, 364-371.

Wang, J., Rath, A., and Deber, C.M. (2014). Functional response of the small multidrug resistance protein EmrE to mutations in transmembrane helix 2. *FEBS Lett* **588**, 3720-3725.

Wassenaar, T.M., Ussery, D., Nielsen, L.N., and Ingmer, H. (2015). Review and phylogenetic analysis of qac genes that reduce susceptibility to quaternary ammonium compounds in *Staphylococcus* species. *Eur J Microbiol Immunol* **5**, 44-61.

Weber, G. (1953). Rotational Brownian motion and polarization of the fluorescence of solutions. *Adv Protein Chem* **8**, 415-459.

Weinglass, A.B., Soskine, M., Vazquez-Ibar, J.L., Whitelegge, J.P., Faull, K.F., Kaback, H.R., and Schuldiner, S. (2005). Exploring the role of a unique carboxyl residue in EmrE by mass spectrometry. *J Biol Chem* **280**, 7487-7492.

White, A.R., and Need, B.W.P.U. (2011). Effective antibacterials: at what cost? The economics of antibacterial resistance and its control. *J Antimicrob Chemother* **66**, 1948-1953.

White, S.H., and Wimley, W.C. (1999). Membrane protein folding and stability: physical principles. *Annu Rev Biophys Biomol Struct* **28**, 319-365.

Wilson, M.R., Hou, Z., and Matherly, L.H. (2014). Substituted cysteine accessibility reveals a novel transmembrane 2-3 reentrant loop and functional role for transmembrane domain 2 in the human proton-coupled folate transporter. *J Biol Chem* **289**, 25287-25295.

Winstone, T.L., Jidenko, M., le Maire, M., Ebel, C., Duncalf, K.A., and Turner, R.J. (2005). Organic solvent extracted EmrE solubilized in dodecyl maltoside is monomeric and binds drug ligand. *Biochem Bioph Res Co* **327**, 437-445.

Wittig, I., Beckhaus, T., Wumaier, Z., Karas, M., and Schagger, H. (2010). Mass estimation of native proteins by blue native electrophoresis: principles and practical hints. *Mol Cell Proteomics* **9**, 2149-2161.

Wittig, I., Karas, M., and Schagger, H. (2007). High resolution clear native electrophoresis for in-gel functional assays and fluorescence studies of membrane protein complexes. *Mol Cell Proteomics* **6**, 1215-1225.

Wittig, I., and Schagger, H. (2008). Features and applications of blue-native and clear-native electrophoresis. *Proteomics* **8**, 3974-3990.

Wong, T.Z., Zhang, M., O'Donoghue, M., and Boost, M. (2013). Presence of antiseptic resistance genes in porcine methicillin-resistant *Staphylococcus aureus*. *Vet Microbiol* **162**, 977-979.

Xu, D., and Zhang, Y. (2012). Ab initio protein structure assembly using continuous structure fragments and optimized knowledge-based force field. *Proteins* **80**, 1715-1735.

Xu, Z. (2005). Molecular analysis of Staohylococcal multidrug transport protein QacA. School of Biological Sciences (Sydney, The Univeristy of Sydney), pp. 155.

Xu, Z., O'Rourke, B.A., Skurray, R.A., and Brown, M.H. (2006). Role of transmembrane segment 10 in efflux mediated by the staphylococcal multidrug transport protein QacA. *J Biol Chem* **281**, 792-799.

Yamada, Y., Hideka, K., Shiota, S., Kuroda, T., and Tsuchiya, T. (2006a). Gene cloning and characterization of SdrM, a chromosomally-encoded multidrug efflux pump, from *Staphylococcus aureus*. *Biol Pharm Bull* **29**, 554-556.

Yamada, Y., Shiota, S., Mizushima, T., Kuroda, T., and Tsuchiya, T. (2006b). Functional gene cloning and characterization of MdeA, a multidrug efflux pump from *Staphylococcus aureus*. *Biol Pharm Bull* **29**, 801-804.

Yan, B.X., and Sun, Y.Q. (1997). Glycine residues provide flexibility for enzyme active sites. *J Biol Chem* **272**, 3190-3194.

Yan, N. (2013). Structural advances for the major facilitator superfamily (MFS) transporters. *Trends Biochem Sci* **38**, 151-159.

Yerushalmi, H., Lebendiker, M., and Schuldiner, S. (1995). EmrE, an *Escherichia coli* 12-kDa multidrug transporter, exchanges toxic cations and H+ and is soluble in organic solvents. *J Biol Chem* **270**, 6856-6863.

Yerushalmi, H., Lebendiker, M., and Schuldiner, S. (1996). Negative dominance studies demonstrate the oligomeric structure of EmrE, a multidrug antiporter from Escherichia coli. *J Biol Chem* **271**, 31044-31048.

Yerushalmi, H., Mordoch, S.S., and Schuldiner, S. (2001). A single carboxyl mutant of the multidrug transporter EmrE is fully functional. *J Biol Chem* **276**, 12744-12748.

Yerushalmi, H., and Schuldiner, S. (2000). An essential glutamyl residue in EmrE, a multidrug antiporter from *Escherichia coli*. *J Biol Chem* **275**, 5264-5269.

Yerushalmi, H., and Schuldiner, S. (2000a). A common binding site for substrates and protons in EmrE, an ion-coupled multidrug transporter. *FEBS Lett* **476**, 93-97.

Yerushalmi, H., and Schuldiner, S. (2000b). A model for coupling of H^+ and substrate fluxes based on "time-sharing" of a common binding site. *Biochemistry* **39**, 14711-14719.

Yin, Y., He, X., Szewczyk, P., Nguyen, T., and Chang, G. (2006). Structure of the multidrug transporter EmrD from *Escherichia coli*. *Science* **312**, 741-744.

Yoshida, H., Bogaki, M., Nakamura, S., Ubukata, K., and Konno, M. (1990). Nucleotide sequence and characterization of the *Staphylococcus aureus norA* gene, which confers resistance to quinolones. *J Bacteriol* **172**, 6942-6949.

Zemla, A. (2003). LGA: A method for finding 3D similarities in protein structures. *Nucleic Acids Res* **31**, 3370-3374.

Zgurskaya, H.I., and Nikaido, H. (2000). Multidrug resistance mechanisms: drug efflux across two membranes. *Mol Microbiol* **37**, 219-225.

Zhang, W., Bogdanov, M., Pi, J., Pittard, A.J., and Dowhan, W. (2003). Reversible topological organization within a polytopic membrane protein is governed by a change in membrane phospholipid composition. *J Biol Chem* **278**, 50128-50135.

Zhang, W., Guan, L., and Kaback, H.R. (2002). Helices VII and X in the lactose permease of *Escherichia coli*: proximity and ligand-induced distance changes. *J Mol Biol* **315**, 53-62.

Zhang, Y., and Skolnick, J. (2004). Scoring function for automated assessment of protein structure template quality. *Proteins* **57**, 702-710.

Zhang, Y., and Skolnick, J. (2005). TM-align: a protein structure alignment algorithm based on the TM-score. *Nucleic Acids Res* **33**, 2302-2309.

Zhang, Z., Ma, C., Pornillos, O., Xiu, X., Chang, G., and Saier, M.H., Jr. (2007). Functional characterization of the heterooligomeric EbrAB multidrug efflux transporter of *Bacillus subtilis*. *Biochemistry* **46**, 5218-5225.

Zhou, Y., Jiang, X., and Kaback, H.R. (2012). Role of the irreplaceable residues in the LacY alternating access mechanism. *Proc Natl Acad Sci USA* **109**, 12438-12442.

Zhu, H., Klemic, J.F., Chang, S., Bertone, P., Casamayor, A., Klemic, K.G., Smith, D., Gerstein, M., Reed, M.A., and Snyder, M. (2000). Analysis of yeast protein kinases using protein chips. *Nat Genet* **26**, 283-289.

Zhu, H., and Snyder, M. (2001). Protein arrays and microarrays. *Curr Opin Chem Biol* **5**, 40-45.

Zolnerciks, J.K., Wooding, C., and Linton, K.J. (2007). Evidence for a Sav1866-like architecture for the human multidrug transporter P-glycoprotein. *FASEB J* **21**, 3937-3948.

Zvelebil, M.J., Barton, G.J., Taylor, W.R., and Sternberg, M.J.E. (1987). Prediction of protein secondary structures and active sites using the alignment of homologous sequences. *J Mol Biol* **195**, 957-961.