Elucidation of the structure and molecular mechanism of the tripartite multidrug efflux pumps in the Gram-negative pathogens: Vibrio cholerae and Neisseria gonorrhoeae

Dijun Du

University of Durham

Elucidation of the structure and molecular mechanism of the tripartite multidrug efflux pumps in the Gram-negative pathogens *Vibrio cholerae* and *Neisseria gonorrhoeae*

By

Dijun Du

School of Biological and Biomedical Sciences

University of Durham

A thesis submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy at University of Durham

December 2006

Abstract

In bacteria, multidrug efflux systems have been identified as significant determinants of resistance recently. These resistance pumps are widely distributed in bacterial species and many pathogenic bacteria posses them, which play an important role in their intrinsic and acquired multidrug resistances. The RND and MATE family transporters have also been shown to be involved in the pathogenicity of bacteria. Knowledge of the structure and mechanism of these transporter proteins would be exceedingly useful in the design of inhibitors.

In Gram-negative bacteria, multidrug resistance is conferred in part by the tripartite multidrug efflux pumps that are composed of an inner membrane transport protein, a membrane fusion protein and an outer membrane protein. One such tripartite pump, VceCAB of *Vibrio cholerae*, is composed of an inner membrane H^+-antiporter VceB, a membrane fusion protein VceA and an outer membrane channel VceC. To investigate the role of this pump in the multidrug resistance of *Vibrio cholerae*, we have characterized functionally and structurally the three components of the VceCAB pump and the regulator VceR.

The crystal structure of VceC was determined at 1.8 Å resolutions. Despite the very low degree of sequence identity between them, VceC shares the same overall architecture as TolC, consisting of three domains: the β-domain, the α-domain and the equatorial domain. The trimeric VceC packs in laminar sheets in the crystal that resemble membranes. Like TolC, the α-barrel of the VceC channel at the periplasmic end is closed through the packing interactions of coiled-coil helices, but the residues that maintain the closed state of the channels of VceC and TolC are different. The β-barrel region of VceC is also closed, whereas the β-barrel region of TolC is open to the extracellular medium. The channel interior of VceC is generally electronegative and contains twó rings of clusters negative charge. The ring made by residues Glu^{397} and Glu^{303} is conserved in OprM, but is not in TolC. Mutagenesis assay of this

negative charged ring indicated its functional role during transport. The optimal desolvation area (ODA) on the surface of VceC is different from that of TolC, suggesting distinct architectures of VceC-based and TolC-based tripartite pumps.

Sub-cellular fractionation of cells expressing full length and truncated VceA suggested that VceA is anchored to the IM via a transmembrane helix. Analytical gel filtration chromatography experiments revealed that the periplasmic domain of VceA that was expressed in the periplasm of *E.coil* forms a trimer, which could represent its oligomeric state in the VceCAB pump.

The three components of tripartite pumps are easy to dissociate *in vitro*, making it difficult to co-crystallize them. We overproduced the protein complex in which VceA (12-406) is in complex with the VceB-VceA fusion protein. This complex was stable during purification, which could provide an invaluable way for co-crystallization of these two components of the VceCAB pump.

An analytical gel filtration and DLS experiments indicated that the basic functional unit of VceR is a dimer; the binding of substrate CCCP to VceR has been determined to occur with a Hill coefficient of about four, and thus each VceR dimer binds four CCCP molecules. This stoichiometry of drug/VceR-subunit is different from that of other transcriptional regulators in the TetR/CamR family.

There are differences between MtrD and other RND family multidrug efflux pumps. The knowledge of difference will be important for understanding the mechanism of these family transporters. In this study, we successfully overexpressed, purified and crystallized MtrD. The resolution of MtrD crystals was optimised to 7-10 Å at present.

Acknowledgements

I would like to give my greatest thanks to Prof. Adrian R. Walmsley for his supervision, constant support and encouragement throughout this project and for the proof reading of this thesis.

I would also like to thank Wolfson Research Institute, University of Durham and The Wellcome Trust for the provision of the funds that sustained me throughout the project.

This work would not have been possible without the additional assistance I received from people of our research group, such as Mr. Hong Tin Lin, Ms Teresa Massam-Wu, and Miss Emma Cunningham etc. My particular thanks go to Dr. Ines Borges-Walmsley, Dr. Daliang Chen, and Dr. Kenny McKeegen for their knowledge, advice and assistance, especially at the early stage of this project. Dr. M. Ines Borges-Walmsley and Dr. Kenny McKeegen made the constructs pET-VceR and pET-MtrD, respectively; Prof. Adrian R. Walmsley provided the dynamic light-scattering analysis for VceR; Dr. Kenny McKeegen and M. Li Zhang developed a procedure to overexpress MtrD; Ms Li Zhang also prepared cellular membrane of MtrD. Further thanks go to the technical staffs Ms Diane Hart and Mr. Simon Padbury.

I wish to give thanks to Dr. Ben Luisi and his group at Cambridge University for the 3D crystallography and data analysis of VceC. Ben Luisi's group also provided the MALDI MS analysis for VceB-VceA fusion protein. I would like to thank Dr. Rick Lewis at University of Newcastle upon Tyne for his help to collect X-ray diffraction data of MtrD crystals at ESRF synchrotron source.

This thesis is dedicated to my wife and daughter. I owed them a lot because I could not take care of them during the several months I wrote my thesis.

Table of contents

Abbreviations

ABC	ATP-binding cassette
ADP.Vi	adenosine 5'-diphosphate and inorganic vanadate
Be	berberine
CAMP	cationic antimicrobial peptide
CCCP	carbonyl cyanide *m*-chlorophenylhydrazone
Cip	ciprofloxacin
CMC	critical micelles concentration
Cryo-EM	cryo-electron microscopy
CV	crystal violet/column volume
CYMAL-6	cyclohexyl-*n*-hexyl-β-D-maltoside
DDM	*n*-dodecyl-ß-D-maltoside
DLS	dynamic light-scattering
Dq	dequalinium
Et	ethidium
FRET	fluorescence resonance energy transfer
HTH	helix-turn-helix
ICD	intracellular domain
IMAC	Immobilized Metal Affinity Chromatography
IMP	Inner membrane protein
IPTG	isopropyl-D-thiogalactopyranoside
ITC	isothermal titration calorimetry
kDa	Kilo dalton
LacY	lactose permease
MALDI MS	Matrix Assisted Laser Desorption Ionization Time-of-flight Mass Spectrometry
MATE	multidrug and toxic compounds extrusion
MDR	Multidrug resistance

MFP	membrane fusion protein
MFS	major facilitator superfamily
MG	malachite green
MPD	2-methyl-2,4-pentanediol
mtr	multiple transferable resistance
MWCO	Molecular weight cut-off
NBD	nucleotide-binding domain
GlpT	glycerol-3-phosphate transporter
OG	octyl-β-glucoside
OMP	outer membrane protein
OxlT	oxalate transporter
PCR	polymerase chain reaction
pmf	electrochemical proton gradient
Pt	3,6-diaminoacridine
PVDF	polyvinglidene difluoride
R6G	rhodamine 6G
Ra LPS	rough-chemotype lipopolysaccharide
RMS	root-mean-square
RND	resistance/nodulation/cell division
SDS-PAGE	Sodium Dodacyl Sulphate-polyacrylamide Gel Electrophoresis
SMR	small multidrug resistance
TCEP	*tris*(2-Carboxyethyl)phosphine
TDG	β-D-galactopyranosyl-1-thio-b-D-galactopyranoside
THP	Tris(hydroxypropyl)phosphine
TM	transmembrane
TMD	transmembrane domain
TMS	transmembrane segment
UVCD	electronic (ultraviolet) circular dichroism
X-Gal	5-bromo-4-chloro-3-indolyl-β-D-galactoside

Chapter 1 General introduction

Structure and mechanism of multidrug efflux pumps in bacteria

Microorganisms have developed various strategies to resist the toxic effects of antibiotics and other drugs (Putman *et al.*, 2001). One of these mechanisms involves the inactivation of antibiotics by enzymatic modification. A well-known example of resistance in bacteria is the hydrolytic deactivation of the β-lactam ring in the penicillins and cephalosporins by the hydrolytic enzyme β-lactamase (Bush *et al.*, 1995). The aminoglycoside antibiotics can be modified by aminoglycoside-resistance enzymes to an inactive form (Shaw *et al.*, 1993). A second mechanism of resistance focuses not on destruction of the antibiotic but on a reprogramming of the cellular target of the antibiotic. This may occur by mutation or enzymatic modification of the target in such a way that the affinity of the antibiotic for the target is reduced (Bussiere *et al.*, 1998). A third resistance strategy of microorganisms is the inhibition of drug entry into the cell (Nikaido, 2001). The peptidoglycan cell wall in both Gram-positive and Gram-negative bacteria, and the outer membrane and layers of lipopolysaccharide of Gram-negative bacteria can prevent drugs from diffusion across the cell envelope. The permeability of the outer membrane of Gram-negative bacteria can be further decreased by the loss or mutation of genes encoding porins (Nikaido, 1994; Nikaido, 2001). These barriers, however, cannot prevent the toxic effects of drugs once they have entered the cell. A fourth mechanism of resistance is the active efflux of antibiotics and other drugs from the cell by multidrug resistance proteins (MDRs) to ensure significant levels of drug resistance (Putman *et al.*, 2001; Borges-Walmsley *et al.*, 2003).

Five families of transporters have been found to be involved in multidrug resistance in prokaryotic/eukáryotic organisms, which can be categorized into primary transporters and secondary transporters. ATP-binding cassette (ABC) superfamily transporters are

the primary transport proteins that exploit the free energy by ATP hydrolysis to catalyze the extrusion of drugs (Davidson *et al.*, 2004). The secondary multidrug transporters are proton or sodium motive force-driven antiporters in which the drug efflux process is coupled to the influx of protons or sodium ions, respectively. These multidrug transporters comprise of four protein families: the major facilitator superfamily (MFS), the small multidrug resistance (SMR) family, the resistance/nodulation/cell division (RND) family, and the multidrug and toxic compounds extrusion (MATE) family (Putman *et al.*, 2001; Borges-Walmsley *et al.*, 2003).

The ABC family of multidrug transporters plays an important role in drug resistance in eukaryotic cells, including human cells. However, in prokaryotes, drug extrusion is largely conferred by H^+-antiporters that belong to the MF superfamily, the SMR family or RND family. Hundreds of multidrug transporters have been characterized in bacteria (reviewed by Markham *et al.*, 2001; Borges-Walmsley *et al.*, 2003; Poole, 2004). These resistance pumps are widely distributed in bacterial species and many pathogenic bacteria posses them. Some of them can extrude a very large range of compounds, including antibiotics, antiseptics, dyes, detergents and solvents, as well as virulence determinants of bacteria, such as adhesins, toxins or other proteins, which are important for the colonization and infection of human and animal cells (Piddock *et al.*, 2006). The RND and MATE family transporters have been shown to be involved in the pathogenicity of bacteria (Piddock *et al.*, 2006). Knowledge of the structure and mechanism of these multidrug transporters would be exceedingly useful in the design of new drugs.

1.1 Major facilitator superfamily

The major facilitator superfamily (MFS) comprises of membrane transport proteins which are found in prokaryotic and eukaryotic cells. To date, more than 3600 members of MFS have been identified (Ren *et al.*, 2004). The MFS consists of five subfamilies, which are involved in drug resistance, phosphate ester/phosphate antiport, sugar and oligosaccharide uptake as well as uptake of Krebs cycle intermediates, respectively (Marger *et al.*, 1993).

MFS transporters are typically composed of approximately 400 amino acid residues. Hydropathy analysis and alignment of conserved motifs of MFS transport proteins reveal that the drug resistance subfamily proteins can be divided into two separate clusters with either 12 or 14 transmembrane segments (TMS) (Paulsen *et al.*, 1993), while the remaining four subfamily proteins contain 12 TMS.

The structures of four MFS transporters have been determined: the X-ray crystallographic structures of the multidrug transporter EmrD (Yin *et al.*, 2006), the lactose/proton symporter (lactose permease, LacY) (Abramson *et al.*, 2003) and the inorganic phosphate/glycerol-3-phosphate antiporter (glycerol-3-phosphate transporter, GlpT) (Huang *et al.*, 2003), as well as the cryo-EM map of oxalate/formate antiporter (OxlT) (Heymann *et al.*, 2003; Hirai *et al.*, 2004).

1.1.1 Crystal structure of EmrD

EmrD is a proton-dependent MFS transporter from *E.coli.* The crystal structure of EmrD has been determined at 3.5 Å resolution, which is illustrated in figure 1.1 (Yin *et al.*, 2006). The functional unit of EmrD is a monomer that is composed of 12 transmembrane helices. The 12 transmembrane helices are divided into two distinct domains, forming N- and C-terminal six-helix bundles connected by a loop between

helices 6 and 7. The transmembrane helix topology of the two domains, which are related by an intramolecular pseudo-twofold axis normal to the membrane plane, is well conserved. The superimposition of the N- and C-terminal domains yields an RMS deviation of 0.78 Å for 116 Cα atoms, indicating that the N- and C-terminal domains are likely to have the same genetic origin.

At the interface between the two domains, a hydrophobic cavity is formed from helices 1, 2, 4 and 5 of the N-terminal domain and helices 7, 8, 10 and 11 of the C-terminal domain. The helices 3, 6, 9 and 12 face toward the lipid bilayer and are not exposed to solvent. The internal cavity is rich in bulky and aromatic amino-acid residues (Ile28, Ile217, Ile253, Tyr52, Tyr56, Trp300 and Phe249). These aromatic side chains may interact with the drug molecule by hydrophobic or aromatic–aromatic interactions, which may play an important role in dictating a level of drug specificity (Figure 1.2).

The two long helical regions (Helix 4, Loop 4-5, Helix 5 and Helix 10, Loop 10-11 and Helix 11), the so-called selectivity filter, could provide additional substrate specificity. Functional studies of EmrD homologues, such as MdfA and LmrP (Adler *et al.*, 2004; Adler *et al.*, 2005; Mazurkiewicz *et al.*, 2002), indicated that several residues in this region are important for substrate recognition (Figure 1.2). The charged residues (Arg118, Arg122, Asp123, Glu126, Arg127 and Arg131) at the cytoplasmic end of Helix 4 may play a role both in defining the topology of the transporter and in substrate recognition.

The residues Thr25, Asp33 and Glu227 on the periplasmic side could easily reorientate into the cavity during the transport cycle, which may participate in H^+ translocation.

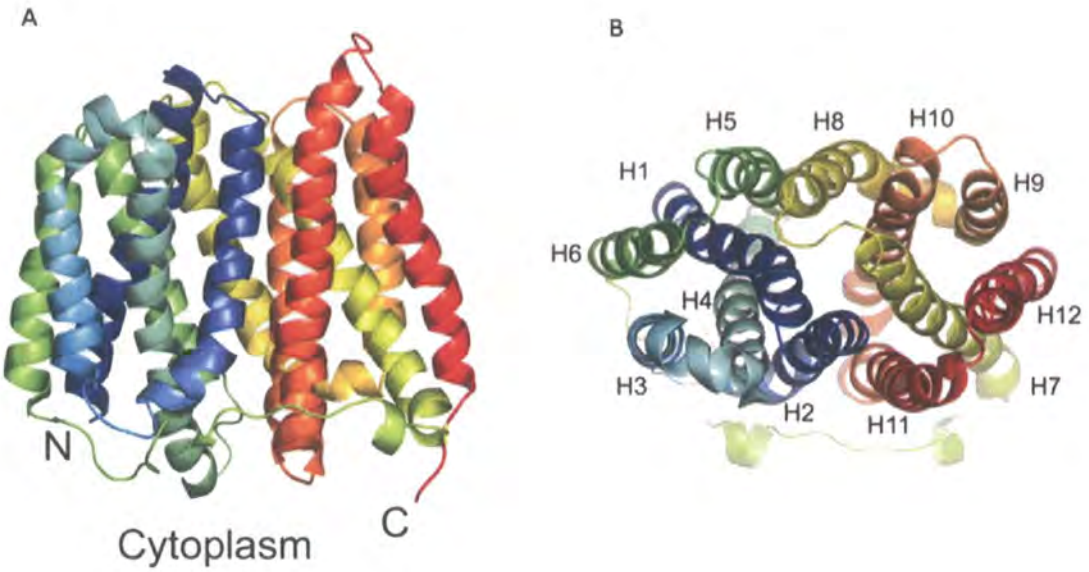


Figure 1.1 Ribbon representation of EmrD. (A) Side view of EmrD along the membrane plane. (B) Top view of EmrD perpendicular to the membrane plane from extracellular side. Transmembrane helices are indicated (Adapted from Yin *et al.*, 2006).

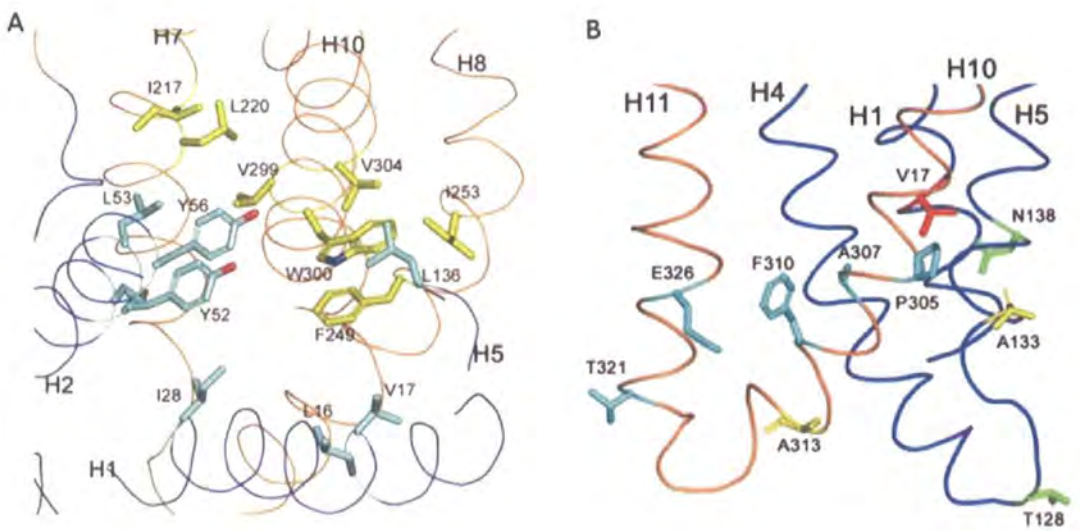


Figure 1.2 (A) The hydrophobic internal cavity of EmrD. The N- and C-terminal halves of EmrD and the corresponding hydrophobic residues are colored blue and orange, respectively. (B) The selectivity filter region of EmrD. The residues displayed are supposed to be involved in substrate recognition according to the mutagenesis assays of the EmrD homolog proteins (Adapted from Yin *et al.*, 2006).

1.1.2 Proposed transport mechanism of EmrD

EmrD is supposed to undergo structural changes between inward and outward-facing conformations during the transport cycle (Yin *et al.*, 2006). In the inward-facing conformation of EmrD, the hydrophobic substrates in the inner leaflet of the cytoplasmic membrane or in the cytoplasm could gain access to the internal cavity of the transporter. The specificity of substrate recognition and binding could be determined by the selectivity filter and the hydrophobic residues in the internal cavity. Then the drug is transported across the lipid bilayer through a rocker-switch between the N- and C-terminal domains coupled with H^+ antiport. The residues Thr25, Asp33 and Glu227 may participate in H^+ translocation (Abramson *et al.*, 2003; Huang *et al.*, 2003; Yin *et al.*, 2006).

The crystal structures of LacY and GlpT share the same overall fold as EmrD, despite the low degree of sequence identity. The major difference between them lies in that LacY and GlpT have hydrophilic internal cavities, while that of EmrD is primarily lined with hydrophobic residues (Abramson *et al.*, 2003; Huang *et al.*, 2003; Abramson *et al.*, 2004a; Abramson *et al.*, 2004b; Yin *et al.*, 2006). Such a difference is consistent with the properties of the substrates they transport. The projection structure of the bacterial oxalate transporter OxlT also suggests a similar architecture with EmrD, LacY and GlpT. Yang *et al.* (2005) devised a homology model for OxlT based upon the structure of GlpT. Despite the low sequence identity between OxlT and GlpT (with 18-20% sequence identity), the model of OxlT accords well with their Oregon green maleimide accessibility and cross-linking assays in their studies. The model also accommodates earlier biochemical information on OxlT. These evidences suggest that MFS transporters with 12-TMS maintain the same basic architecture (Abramson *et al.*, 2004a; Abramson *et al.*, 2004b).

The 14-TMS MFS multidrug efflux proteins QacA and QacB have seven nucleotide differences only. QacB confers resistance to monovalent organic cations but

characteristically differs from QacA by conferring lower or no resistance to divalent cations. These phenotypic differences between QacA and QacB are due solely to the presence of an acidic residue (Asp) at position 323 in QacA instead of an uncharged residue (Ala) in QacB (Paulsen *et al.*, 1996). It is likely that only a few key residues are involved in substrate recognition and the substitution of the corresponding residues can change the specificity of MFS transporters. This may explain the substrate diversity of the MFS.

1.2 The resistance/nodulation/division (RND) superfamily

The resistance/nodulation/division (RND) superfamily transporters are found in bacteria, archaea and eukaryotes (Saier *et al.* 2001). Characterized members of the RND superfamily are all H^+-antiporters. The RND superfamily transporters are composed typically of approximately 1000 amino acid residues and the first half of the proteins usually resembles to the second half, suggesting that the proteins arose as a result of an intragenic tandem duplication event that occurred in the primordial system prior to divergence of the family members. The RND superfamily transporters are predicted to contain 12 transmembrane helices (per monomer) and characteristically have two large periplasmic loops between transmembrane helices 1 and 2 and between 7 and 8 (Saier *et al.*, 2001; Putman *et al.*, 2000).

The RND family of transporters tend to play major roles in the intrinsic drug resistance of Gram-negative bacteria (Nikaido 1998). Some of them can extrude a very large range of compounds. The well-studied RND family multidrug efflux pump AcrB is the inner membrane component of the tripartite pump AcrAB/TolC of *Escherichia coli*. AcrB confers resistance to the antibiotics tetracycline, chloramphenicol, β-lactams, novobiocin, fusidic acid, nalidixic acid, and fluoroquinolones; the chemotherapeutic agents, SDS and Triton X-100, bile salts that

act as detergents, various cationic dyes, disinfectants, and even solvents (Tsukagoshi *et al.*, 2000; White *et al.*, 1997).

The crystal structure of AcrB has been determined in several conformational states with/without bound substrates (Murakami *et al.*, 2002; Yu *et al.*, 2003; Murakami *et al.*, 2006; Seeger *et al.*, 2006), which gave us the clues to structure-function relationships in multidrug transporters that underlie the transport mechanism of RND family proteins.

1.2.1 Crystal structure of AcrB

The first crystal structure of AcrB, belonging to R32 space group, was obtained at 3.5 Å resolution. The three protomers are fixed symmetrically by the crystal packing force (Murakami *et al.*, 2002). The new crystal forms of AcrB, belonging to C2 and P1 space group, were determined at 2.8 Å and 2.9 Å resolution. These space groups allow each protomer to take a different conformation (Murakami *et al.*, 2006; Seeger *et al.*, 2006).

The functional unit of AcrB is a trimer that has a superficial jellyfish-like appearance. The three protomers adopt different conformations, termed binding protomer, extrusion protomer and access protomer, respectively. AcrB comprises an extra-membrane (periplasmic) headpiece and a transmembrane region approximately 70Å and 50Å thick, respectively. The AcrB structure contains three domains: the transmembrane (TM) domain, the porter domain and the TolC docking domain (Figure 1.3).

The porter domain locates at the lower part of the headpiece, which is composed of four subdomains, PN1, PN2, PC1 and PC2, for each protomer. PN1 and PN2 locate between TM1 and TM2, while PC1 and PC2 locate between TM7 and TM8. Each

subdomain contains two characteristic β-strand-α-helix-β-strand motifs. These four subdomains are packed with their β-sheets surrounding a substrate binding pocket.

Three α-helices (the second α-helix of PN1 of each protomer) form a pore at the centre of the porter domain. There is a central cavity between the proximal end of the porter domain and TM domain. Three channels are open from the central cavity into the periplasm at the side of the headpiece between PN2 and PC2, which could provide access to periplasmic substrates. In the binding protomer, the channel is branched at a vestibule near the entrance and continues through the space between the β-sheets in the porter domain to the binding pocket. The helix from the extrusion protomer is inclined nearly 15° towards the binding protomer, blocking a gap between PN1 and PN2 of its binding pocket, which is a possible path from the binding pocket to the top funnel. The inclination of this helix also creates a space between PN1 and PN2 in the extrusion protomer, opening to the top funnel, however, the uptake channel for the extrusion protomer has disappeared. The movement of PC2 and partial unwinding and rewinding of the upper end of transmembrane helix 8 (TM8) result in the opening and shutting of the pathway branched at the vestibule (Figure 1.3).

The TM domain comprises 36 TM α-helices (12 from each protomer) and an additional extra-membrane α-helix for each protomer located between TM6 and TM7, which attaches to the cytoplasmic membrane surface. The TM domain seems to be loosely packed. A central hole with a diameter of about 30 Å is formed by the TM domain of the protomers, which is thought to be filled with lipid. The TM4 and TM10 form the centre of the transmembrane helix bundles, which are suggested to be the proton translocation pathway. The residues Asp407, Asp408 and Lys940 that locate in the middle of TM4 and TM10 are possibly central to gating. In the access and binding protomers, the Lys 940 residue is coordinated by salt bridges with Asp 407 and Asp 408; in the extrusion protomer, however, Lys 940 is turned nearly 45° towards Thr 978 of TM11 and the salt bridges are abolished. This movement result in the twisting of TM4 and TM10.

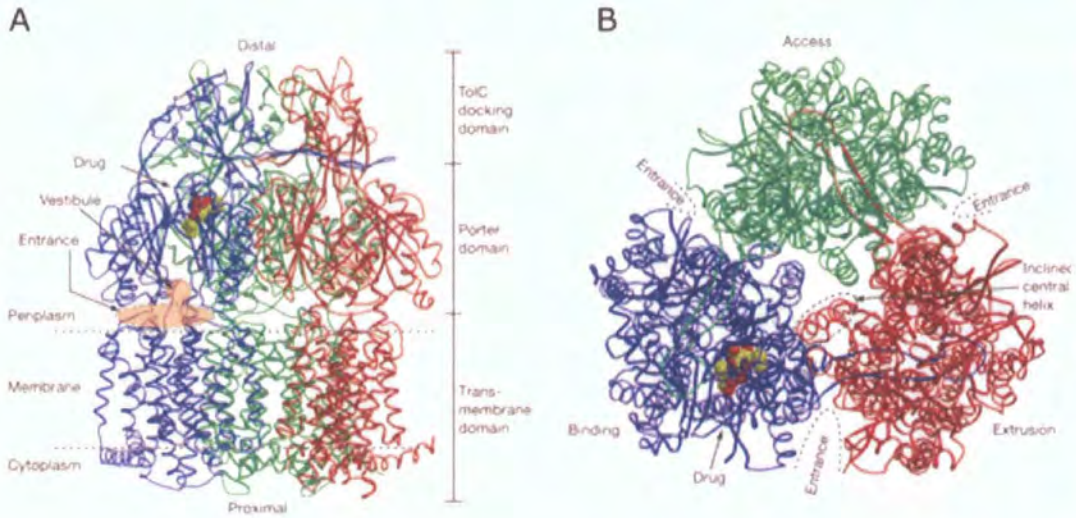


Figure 1.3 Ribbon representation of the AcrB-minocycline complex. The three AcrB protomers are colored blue, red and green, respectively.The minocycline molecule is shown in CPK. (A) Side view of AcrB along the membrane plane. (B) Top view of AcrB perpendicular to the membrane plane from extracellular side (Adapted from Murakami *et al.*, 2006).

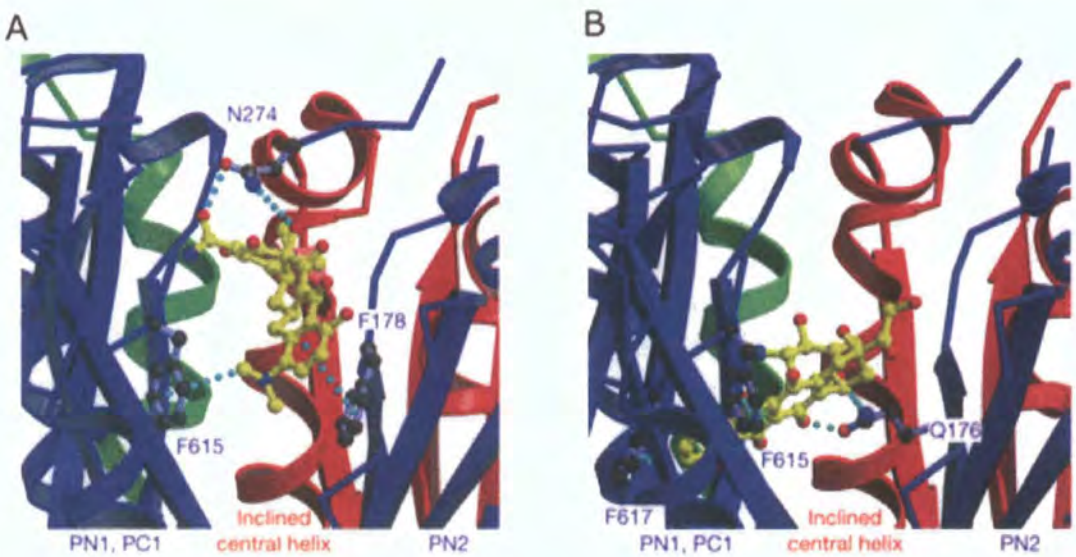


Figure 1.4 Ribbon representation of the drug binding site of AcrB in complex with (A) minocycline and (B) doxorubicin, respectively (Adapted from Murakami *et al.*, 2006).

On the membrane-exposed surface of the TM domain is a vertical groove between helices TM7 and TM8. It extends across the whole membrane-embedded surface. The base of the groove is a tilted helix TM9. This groove perhaps provides access for substrates located at the cytoplasmic side of the membrane.

The TolC docking domain locates at the upper part of the headpiece, which is composed of two subdomains, DN and DC. Each subdomain contains a four-stranded mixed β-sheet. Two antiparallel β-strands are placed parallel to a hairpin structure from the other half or from the other protomer. The TolC docking domain is open like a funnel with an internal diameter of about 30 Å, similar to the diameter of the modelled open state of the TolC entrance.

1.2.2 Molecular basis for substrate recognition

The crystal structure of AcrB indicates that it is likely to capture substrates from both the periplasm and cytoplasm. Aires *et al.* (2005) reconstituted the purified RND transporter AcrD of *E. coli* into unilamellar proteoliposomes, established that they could take up a hydrophilic substrate aminoglycosides, indicating that AcrD captures its substrate aminoglycosides from both the periplasm and cytoplasm. This activity required the presence of AcrA within the proteoliposomes.

The periplasmic headpiece of RND family transporters has been shown to play a major role in determining substrate specificity. Elkins and Nikaido (2002) swapped sections of the periplasmic loops between two RND family transporters AcrB and AcrD. AcrB conferred resistance only to the typical substrates of AcrD when the two loops of AcrB were replaced with those of AcrD, and *vice versa*. Tikhonova *et al.* (2002) produced a chimeric AcrB/MexB transporter by fusing the N-terminal regions of AcrB and the C-terminal residues of MexB from *P. aeruginosa*, which contained both loops of AcrB. This chimaeric transporter conferred resistance to all AcrB

substrates. Mao *et al.* (2002) isolated spontaneous mutants of a RND family pump MexD from *P. aeruginosa*, which allow the pump to efflux the previously non-transported antibiotic carbenicillin. These mutants were all mapped to the periplasmic loops of the pump, indicating that the large periplasmic loops of MexD from *P. aeruginosa* are involved in substrate recognition.

The crystal structure of AcrB in complex with minocycline and doxorubicin provides a structural basis of the periplasmic headpiece in substrate recognition and transport (Figure 1.4). In the C2 crystals, the four subdomains of the porter domain are packed with their β-sheets surrounding a substrate binding pocket, which is rich in aromatic amino-acid residues: Phe 136 and Phe 178 (PN2), and Phe 610, Phe615, Phe617 and Phe628 (PC1). The substrate binding pockets of the three protomers have a conformational difference, which is caused by the movement of PN2 and PC1. The pocket of the binding protomer has an expanded structure. The phenylalanine side chains in this expanded pocket are favourable for hydrophobic interaction with substrates. The bound minocycline and doxorubicin are observed at the distal end of the pocket between the β-sheets of PN2 and PC1. The vacant binding pockets of the access protomer and the extrusion protomer have shrunken structure and the phenylalanine side chains in them are moved towards each other, which are unfavourable for drug binding.

The crystal structures of AcrB in complex with four substrates, rhodamine 6G (R6G), ethidium (Et), dequalinium (Dq), and ciprofloxacin (Cip), have also been determined at 3.5 to 3.8 Å resolutions (Yu *et al.*, 2003). All four substrates are bound to the periphery of the central cavity, which is surrounded by many hydrophobic residues, including 12 well-conserved phenylalanine residues, each protomer contributing Phe386, Phe388, Phe458 and Phe459. In every case, the trimeric AcrB bound three drug molecules. These drug binding sites in the transmembrane domain of AcrB may be a snapshot of a transient point in the efflux process.

1.2.3 Proposed transport mechanism of AcrB

The asymmetric structure of the AcrB trimer provides a structural basis for considering the molecular mechanism of RND family multidrug transporters (Murakami *et al.*, 2006; Seeger *et al.*, 2006; Schuldiner, 2006).

The three protomer conformations of AcrB may represent the consecutive states of a transport cycle (Figure 1.5). The substrates could first gain access to the binding domain of AcrB when it is in the access protomer conformation. The conversion of the PN1/PC2 subdomains from the access to binding protomer conformation enables the substrates to move through the uptake channel and bind in the large internal pocket. At this point, the exit of the binding pocket is blocked by the inclined helix of the extrusion protomer. Then the PN1/PC2 subdomains converse from the binding to extrusion protomer conformation, the drug binding pocket is shrunken and the exit is opened so that the substrates are squeezed out. The PN1 subdomains play an important role to such an ordered binding change mechanism (Murakami *et al.*, 2006; Seeger *et al.*, 2006; Schuldiner, 2006).

Murakami *et al.* (2005) constructed cysteine-scanning mutants as to the twenty-one residues comprising the pore helix. Five of the twenty-one mutants, D101C, V105C, N109C, Q112C, and P116C, which are localized on the wall of the pore, showed significantly reduced drug resistance and drug-exporting activity, indicating the important role of this pore in the drug transport process. Two residues, V105C and Q112C, are close to the corresponding residues in the next protomer of AcrB trimer. The formation of disulfide cross-linking between these residues prevent any conformational change of the pore, resulting in the greatest loss of activity in all of the pore mutants. These findings suggest that a conformational change of the pore is indispensable during the transport process.

The TM4 and TM10 helices have been suggested to form the proton translocation pathway. The residues Asp407, Asp408 and Lys940 that locate in the middle of TM4 and TM10 are possible candidates for the transmembrane proton translocation site. The protonation of aspartic acids would disrupt the ion pairs between them, resulting in a conformational change of the helices TM4 and TM10. However, how this conformational change of the transmembrane helices causes the movements of subdomains (including those in the porter domain), rewinding of the upper part of TM8, and the inclination of the central α-helix is still unclear.

Two putative pathways exist for substrates to enter the drug binding pocket of AcrB. The groove at the periphery of each transmembrane domain perhaps could provide access for substrates located in the cytoplasm or inner leaflet of the membrane; substrates located on the outer surface or in the outer leaflet of the membrane could gain access to the cavity through the vestibules that open into the periplasm.

The diameter of the top of TolC docking domain is about the same as that of the modelled open state of the TolC entrance, and they fit well with each other by manual docking, suggesting the possibility of a direct interaction between TolC and AcrB. Tamura *et al.* (2005) introduced cysteine mutations at the tops of the vertical hairpins of AcrB and the bottoms of the coiled coils of TolC molecules. The AcrB-TolC complex formed through disulfide cross-linking *in vivo* when a specific pair of mutants was coexpressed in *E. coli,* indicating direct interaction between TolC and AcrB.

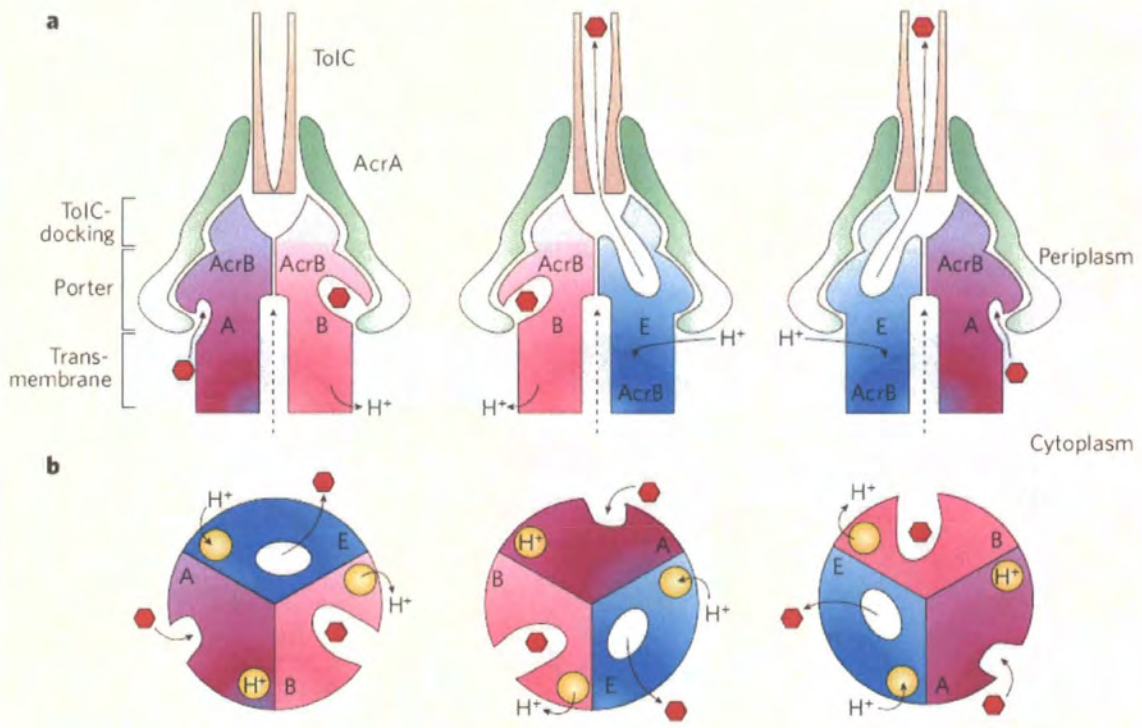


Figure 1.5 Schematic representation of the transport mechanism of AcrB. 'A', 'B' and 'E' represent the access, binding and extrusion conformation, respectively. (a) Side view of the tripartite efflux pump AcrAB-TolC. (b) The proposed consecutively conformational change of the three protomers during a transport cycle (Adapted from Murakami *et al.,* 2006; Seeger *et al.,* 2006; Schuldiner, 2006).

1.3 Small multidrug resistance family

Small multidrug resistance (SMR) family belongs to the drug/metabolite transporters superfamily. Its members constitute the smallest known secondary transporters. These proteins are typically composed of around 110 amino acid residues with four predicted TMS and function as homooligomers. Some SMR family transporters, such as EbrAB and YkkCD, are composed of two dissimilar but homologous subunits: a short subunit with 105 to 106 amino acids and a long subunit with 111 to 117 amino acids (Putman *et al.*, 2000). This difference is due to a C-terminal hydrophilic extension (Jack *et al.*, 2000).

Multiple-sequence alignment of SMR family proteins reveals a number of conserved residues, which may play essential structural or functional roles (Putman *et al.*, 2000).

EmrE from *E.coli* is one of the best-characterized SMR family transporters, which confers resistance to monovalent cations, such as ethidium, proflavine, pyronin Y, safranin O, and methyl viologen as well as to erythromycin, sulfadiazine, TPP, and tetracycline (Yerushalmi *et al.*, 1995; Paulsen *et al.*, 1996).

The structure of EmrE has been determined by cryo-electron microscopy (cryo-EM) (Ubarretxena-Belandia *et al.*, 2003).

1.3.1 Cryo-electron microscopy structure of EmrE

The three-dimensional structure of EmrE with bound TPP^+ derived from cryo-electron microscopy of two-dimensional crystal has been determined at 7 Å resolution (Ubarretxena-Belandia *et al.*, 2003). The structure suggested an EmrE antiparallel dimer consisting of eight α-helix bundles (Figure1.6).

TPP^+ was bound in a chamber formed by six tilted α-helices (helices A-C and F-H). The helices D and E were nearly perpendicular to the membrane and were separated from the chamber by two of the highly tilted helices C and F in the wall of the chamber. The chamber appears to be closed by the convergence of helices F and H just past the membrane centre. The density for TPP^+ merged to the main density for helix H. Two openings of the substrate binding chamber face the aqueous medium and the lipid bilayer, respectively, which could provide an access for hydrophobic substrates in one leaflet of the *E.coli* inner membrane to the EmrE binding site.

1.3.2 Substrate recognition by EmrE

Site-directed mutagenesis and chemical modification experiments have provided insights into the identity of the amino acid residues of EmrE in the substrate binding pocket and translocation pathway.

EmrE includes eight charged residues, seven of which are located in putative loops and can be replaced with either cysteine or other amino acids bearing the same charge without significantly impairing transport activity. A conserved residue Glu14 in the putative trasnmembrane helix-1 is essential for activity, as it constitutes the site at which both protons and cationic substrates bind in a mutually exclusive fashion (Yerushalmi *et al.*, 2000).

Gutman *et al.* (2003) constructed cysteine-scanning mutants as to the residues in the transmembrane helix-1of EmrE, revealing an amino acid cluster on the same face of helix-1 as Glu-14 that is part of the substrate- and proton-binding domain. Substitutions at most of these positions yielded mutants with reduced substrate-exporting activity or modified affinity to substrates. Substitutions at the Ala-10 position yielded mutants with modified affinity to protons and thereby impaired in the coupling of substrate and proton fluxes. These results indicate the important role of

one face of helix-1 in substrate recognition, binding, and H^+-coupled transport and the existence of a common binding site for substrates and protons.

Sharoni *et al.* (2005) explored the binding domain by alkylation of Cys replacements of EmrE, indicating that the existence of a binding cavity accessible to alkylating reagents where at least Leu-7, Ala-10, and Ile-11 from TM1, Tyr-40 from TM2, and Trp-63 in TM3 are involved in substrate binding.

Rotem *et al.* (2004) compared transport of monovalent and divalent substrates by EmrE in intact cells and in proteoliposomes reconstituted with the purified protein, demonstrating that the transport of monovalent substrates involves charge movement (*i.e.* electrogenic), while the transport of divalent substrate does not (*i.e.* electroneutral), indicating that an EmrE dimer exchanges two protons per substrate molecule during each transport cycle.

1.3.3 Transport mechanism of EmrE

Ubarretxena-Belandia *et al.* (2003) proposed a drug-efflux mechanism based on the cryo-EM structure of EmrE (Figure 1.7). The substrate binding pocket and translocation pathway are proposed to be the chamber formed by six helices (A-C and F-H). The hydrophobic substrates in the inner leaflet of the cytoplasmic membrane or in the cytoplasm could gain access to the chamber through the openings that laterally face the lipid bilayer or face the aqueous medium and bind to the drug-binding site composed of two Glu-14s from helix-1s (Figure 1.7-1). The conformational change could then re-orient the binding chamber to open towards the periplasm and the direct binding of protons to both Glu-14 residues induces the release of substrate at the periplasmic surface (Figure.1.7-2, 3). A second proton-driven conformation change (*) re-orients the binding site to face the cytoplasm (Figure 1.7-4).

Ninio *et al.* (2004) probed the membrane topology of EmrE, focusing on the carboxy-terminus of the protein, supporting a secondary structure where the carboxy-terminus faces the cytoplasm. Soskine *et al.* (2006) introduced two cysteine mutations into TM4 of EmrE: D84C and T108C. Crosslinking experiments suggested a parallel arrangement of TM4 between the EmrE monomers. These findings argue for an antiparallel topology of EmrE. Fleishman *et al.* (2006) proposed an EmrE model with parallel topology by assignment of the transmembrane segments in EmrE to the densities seen in the cryo-EM structure. This model accommodates much of the earlier biochemical information on EmrE.

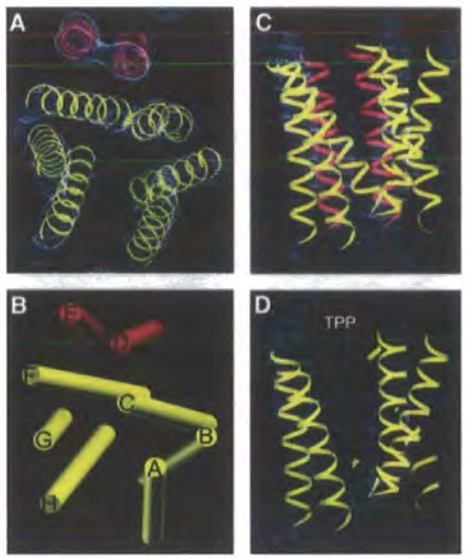


Figure 1.6. The three-dimensional structure of EmrE with bound TPP^+ derived from cryo-electron microscopy of two-dimensional crystal. (A) and (C) Top and side view of the density contoured at 1.2σ, respectively. (B) Schematic of top view of the EmrE. (D) Side view of a slice along the membrane plane of the density contoured at 0.8σ (Adapted from Ubarretxena-Belandia *et al.*, 2003).

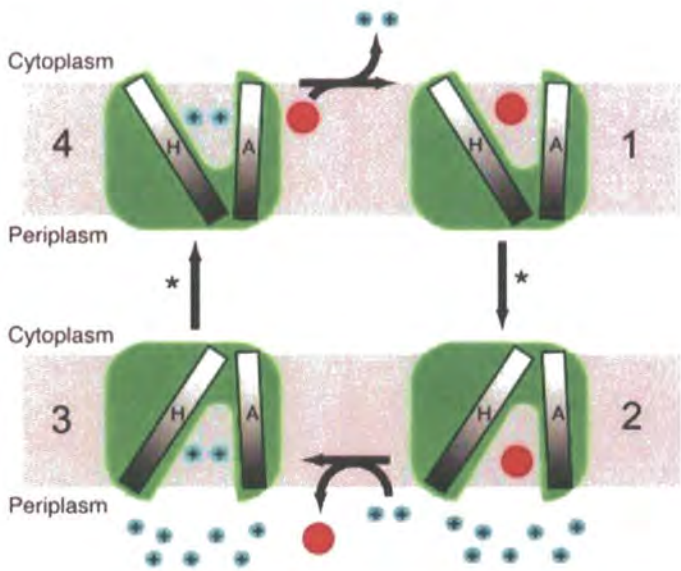


Figure 1.7. Transport mechanism of EmrE. TPP^+ is represented as a red ball and protons as blue balls, labelled '+' (Adapted from Ubarretxena-Belandia *et al.*, 2003).

1.4 The multidrug and toxic compound extrusion family transporter

The multidrug and toxic compound extrusion (MATE) family of transporters are found in bacteria and yeast. Similar to the MFS transporters, the bacterial MATE family transporters are typically composed of approximately 450 amino acid residues with 12 predicted TMS and short loops between them; however, they do not have sequence homology with the MFS transporters. The yeast MATE family transporters are larger (up to 700 amino acid residues) but they are also putatively arranged into 12 transmembrane helices. Phylogenetic analysis of the MATE family has revealed the presence of three distinct clusters (Brown *et al.*, 1999).

Most of the characterized MATE family pumps, such as NorM and VmrA of *Vibrio parahaemolyticus* (Morita *et al.*, 2000; Chen *et al.*, 2001), YdhE of *Escherichia coli* (Nishino *et al.*, 2001), VcmA and VcrM of *Vibrio cholerae non-O1* (Huda *et al.*, 2001; Huda *et al.*, 2003), CdeA of *Clostridium difficile* (Dridi *et al.*, 2004), HmrM of *Haemophilus influenzae* (Xu *et al.*, 2003) and MepR of *Staphylococcus aureus* (Kaatz *et al.*, 2005), utilize an electrochemical potential of Na^+ across membranes as the driving force; however, two MATE family pumps, PmpM of *Pseudomonas aeruginosa* (He *et al.*, 2004) and AbeM of *Acinetobacter baumannii* (Su *et al.*, 2005), were identified recently to utilize an electrochemical potential of H^+ across their membranes as the driving force.

The MATE family transporters exhibit a broad range of substrates. For instance, the VcrM confers resistance to acriflavine, 4',6-diamidino-2-phenylindole, rhodamine 6G, ethidium bromide, etc (Huda *et al.*, 2003). Mutagenesis experiments indicated that the acidic amino acid residues Asp32, Glu251, and Asp367 in the transmembrane region of NorM, which are conserved in the clusters containing NorM, are involved in the Na^+-dependent drug transport process (Otsuka *et al.*, 2005).

The MATE family transporters are the least well characterized. Structural or functional information on the molecular mechanism underlying the Na^+ (or H^+)/multidrug antiport by the MATE family is very limited at present.

1.5 ATP Binding Cassette Transporters

The ATP-binding cassette (ABC) superfamily is one of the largest superfamilies of proteins known, which couple the energy released from ATP hydrolysis to the translocation of a wide variety of substances into or out of cells and organelles (Davidson *et al.*, 2004).

The ABC transporter proteins are generally composed of four protein domains or subunits: two hydrophobic transmembrane domains (TMDs) that are supposed to provide a translocation pathway across the membrane and two hydrophilic nucleotide-binding domains (NBDs) that break down the ATP nucleotide to supply the energy for active transport (Davidson *et al.*, 2004).

The individual domains can be expressed as separate proteins or may be fused into multidomain polypeptides in a variety of ways. For example, a single gene encodes the four domains of P-glycoprotein (Gottesman *et al.*, 1996); the *lmrA* and *msbA* genes encode proteins with a single combined TMD and NBD, and the *drrA* and *drrB* genes encode the NBD and TMD as individual proteins. While the NBD of MsbA is located at the N-terminus of TMD, the NBD of LmrA is located at the C-terminus of TMD (Dawson *et al.*, 2006; van Veen *et al.*, 1996).

The predicted transmembrane helical number in the TMD between subfamilies is different. MacB and DrrB are predicted to have four and eight transmembrane helices in the TMD respectively, while MsbA and BtuC have six and ten transmembrane

helices in the TMD respectively (Kobayashi *et al.*,2003; Gandlur *et al.*, 2004; Locher *et al.*, 2002).

The homology between TMDs in different subfamilies is lower, however, the NBD sequence of the entire superfamily shares 25% to 30% identity and contains some highly conserved motifs, such as Walker A/Walker B motifs, LSGGQ signature motif, H motif, P loop and Q loop, suggesting that a similar mechanism for coupling of transport to ATP hydrolysis is employed (Diederichs *et al.*, 2000; Gaudet *et al.*, 2001; Yuan *et al.*, 2001; Karpowich *et al.*, 2001; Schmitt *et al.*, 2003; Verdon *et al.*, 2003).

The ABC superfamily contains both uptake and efflux transport systems. All ATP-dependent drug efflux proteins known to date are members of the ABC superfamily. Most bacterial ABC drug transporters mediate the export of specific antibiotics (Barrasa *et al.*, 1995; Guilfoile *et al.*, 1991; Linton *et al.*, 1994; Olano *et al.*, 1995; Podlesek *et al.*, 1995; Ross *et al.*, 1990). Some bacterial ABC transport proteins, such as LmrA and MsbA, share up to 35% sequence identity to the mammalian multidrug efflux pump P-glycoprotein. The region of homology exists in not only NBD but also TMD, indicating that these ABC proteins may recognize and extrude similar substrates (Davidson *et al.*, 2004).

The crystal structures of intact ABC family transporters Sav1866, BtuCD and HI1470/1 (Pinkett *et al.*, 2007), as well as the structures of isolated NBDs of transporters, have given an insight into their mechanism.

1.5.1 Structure of Sav1866

Sav1866 is a multidrug ABC transporter from *Staphylococcus aureus*. The crystal structure of Sav1866 in complex with ADP has been determined at 3 Å resolution (Dawson *et al.*, 2006), which is shown in Figure 1.8. The structure represents an outward-facing conformation. The functional unit of Sav1866 is a homodimer and each subunit is composed of two domains: an amino-terminal transmembrane domain (TMD, amino-acid residues 1–320) and a carboxy-terminal nucleotide-binding domain (NBD, residues 337–578). The two subunits twist and embrace each other. Both the TMDs and the NBDs interact tightly. Two ADP molecules are bound by the P-loops and the ABC signature motifs of opposing subunits.

The six transmembrane α-helices bundles of each subunit are connected by long intracellular and short extracellular loops (ICLs and ECLs, respectively). Around the middle of the membrane, bundles of transmembrane helices diverge into two discrete 'wings' that point away from one another and consist of helices TM1–TM2 from one subunit and TM3–TM6 from the other subunit. The ICLs extend the helical secondary structure beyond the lipid bilayer, which protrude approximately 25 Å into the cytoplasm.

The large cavity at the interface of the two transmembrane domains is accessible from the outer leaflet and exposed to the extracellular space. The bottom of the cavity reaches beyond the intracellular membrane boundary without connection to the cytoplasm and the inner leaflet. At the level of the inner leaflet, the cavity is formed from helices TM2–TM5 and is rich in polar and charged amino-acid residues, while at the level of the outer leaflet, the cavity is formed from helices TM1, TM3 and TM6.

The coupling helices of ICLs, which orient roughly parallel to the membrane plane, provide the bulk of the contacts to NBDs: the coupling helix 1 contacts the NBDs of both subunits, while coupling helix 2 interacts with that of the opposite subunit. The

residues around the Q-loop of NBD form the contact surface with the TMD. Two conserved residues, Tyr 391 and Glu 473, are exceptional. The Glu 473 interacts with both ICLs and is part of the x-loop motifs (TEVGERG), which precede the ABC signature motifs (Figure 1.9).

1.5.2 Proposed transport mechanism of Sav1866

Sav1866 is supposed to undergo structural changes between inward and outward-facing conformations during the transport cycle. The proposed transport mechanism of Sav1866 is based on the outward-facing crystal structure. In the outward-facing conformation, an extrusion pocket exposes to the external medium. The polar and changed amino acids in the extrusion pocket have little or no affinity for hydrophobic drugs, so that the bound substrates are released into the outer leaflet of the lipid bilayer or into the aqueous medium. Hydrolysis of ATP induces a transition to the inward-facing conformation, granting access to the binding site from the cell interior, and the transporter is ready for another cycle (Dawson *et al.*, 2006).

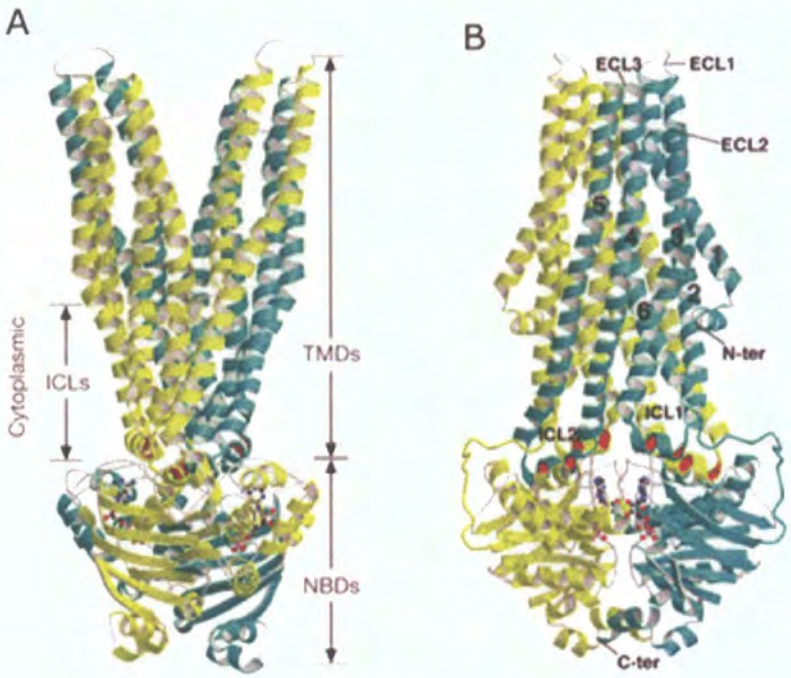


Figure 1.8. Two perpendicular views of the crystal structure of Sav1866. The two protomers are colored yellow and turquoise, respectively. The ADP molecules are shown in ball-and-stick (Adapted from Dawson *et al.,* 2006).

Figure 1.9. Ribbon representation of the transmission interface of Sav1866. The ADP molecules and the side chains of the conserved residues Tyr391 and Glu473 are shown in ball-and-stick. The 'coupling helices' are coloured red (Adapted from Dawson *et al.,* 2006).

1.6 Polyspecific substrate recognition by multidrug efflux pumps, implied by multidrug-binding regulatory proteins

The structural determination of multidrug efflux pumps could be critical for us to understand their mechanism; however, membrane transport proteins are quite difficult to purify and crystallize. In bacteria, the expression of multidrug efflux pump genes is tightly regulated by transcriptional activators and/or repressors. These cytosolic multidrug-binding transcriptional regulators often bind the same substrate as that of the pumps they regulate and thus, these proteins are themselves multidrug binding proteins. The transcriptional regulators of multidrug efflux pumps provide an alternative in studying the multidrug recognition mechanisms (Schumacher *et al.*, 2003; Paulsen, 2003).

The crystal structures of several multidrug-binding regulatory proteins in complex with their operator DNA and/or substrates have been determined, which provide insight into the molecular basis of multidrug recognition (Schumacher *et al.*, 2001; Heldwein *et al.*, 2001; Godsey *et al.*, 2001; Schumacher *et al.*, 2002; Grkovic *et al.*, 2002; Newberry *et al.*, 2004; Murray *et al.*, 2004; Schumacher *et al.*, 2004).

1.6.1 Molecular basis of substrate recognition by QacR

The *Staphylococcus aureus* protein QacR is a transcriptional repressor of the *qacA* gene, a MF family multidrug transporter. The crystal structures of QacR bound to structurally diverse cytotoxic agents, rhodamine 6G (R6G), ethidium (Et), dequalinium (Dq), crystal violet (CV), malachite green (MG), berberine (Be), pentamidine, hexamidine and 3,6-diaminoacridine (Pt) have been determined (Schumacher *et al.*, 2001; Grkovic *et al.*, 2002; Murray *et al.*, 2004; Schumacher *et al.*, 2004).

The structure of QacR comprises nine helices. The first three helices form a three helix bundle DNA binding domain, which contains a helix-turn-helix (HTH) motif (α2 and α3). Helices 4 to 9 form the drug binding/dimerization domain.

The ligand-binding to QacR triggers a coil-to-helix transition of residues Thr89–Tyr93 of the drug-bound subunit, resulting in extension of the C-terminus of α5 by 4 residues. This coil-to-helix transition expulses the Tyr92 and Tyr93 from the hydrophobic core of QacR, which is composed of the side chains of residues Trp61, Trp65, Phe79, Tyr82, Tyr92, Tyr93, Tyr123, Tyr127, Phe131, Trp140, Phe178 and Phe182, thereby increasing the volume of the binding pocket for drug binding. The coil-to-helix transition of α5 also relocates residue Glu90 into the drug binding pocket so that it could assist with drug binding. In the absence of a drug, the side chains of residues Tyr92 and Tyr93 play an essential role of structural drug surrogates (Schumacher *et al.*, 2001; Grkovic *et al.*, 2002).

The binding of six structurally diverse compounds in the extended QacR ligand-binding pocket is shown in Figure 1.10. The drug binding pocket created by tyrosine expulsion is composed of residues from α4 to α8 from one subunit and α8' from the other subunit of the dimeric QacR. Two drug binding pockets are present within the large QacR drug binding pocket: the R6G binding pocket and the Et binding pocket, which are distinct but partially overlap. The monovalent compounds R6G and ethidium bromide were found to bind to the R6G binding pocket and Et binding pocket, respectively, while the positively charged aminomethylquinolinium moieties of the bivalent compound dequalinium that have a flexible decamethylene linker was found to bind in both of these binding pocket. Crystal violet was bound at an intermediate position between the R6G binding pocket and Et binding pocket, demonstrating the versatility of the extended QacR ligand-binding pocket (Schumacher *et al.*, 2001; Grkovic *et al.*, 2002; Schumacher *et al.*, 2003).

The large number of aromatic and hydrophobic residues lining the QacR drug-binding pocket present incoming drugs with a broad range of possibilities for the formation of hydrophobic interactions or the stacking of phenyl rings. Several polar residues located within the pocket serve as hydrogen bond donors and acceptors, including asparagine, glutamine, threonine and serine residues that add versatility for drug binding by contacting the polar moieties of different drugs. For example, in the QacR–R6G complex, the R6G phenyl moiety stacks and makes hydrophobic interaction with Tyr123 and Leu54. The three-ring system of R6G makes stacking with the side chains of Trp61 from α4 and Tyr93 from α5. The polar residues Gln64 and Thr89 of QacR contact the R6G N2 and Gln96 contacts R6G N1. The R6G central ring O1 atom contacts the Oγ of residue Thr89, further securing the R6G ring system (Schumacher *et al.*, 2001; Grkovic *et al.*, 2002; Schumacher *et al.*, 2003).

Four negatively charged glutamate residues, Glu57, Glu58, Glu90 and Glu120, line the QacR ligand-binding pocket roughly equidistant from the centre of the large pocket, which are available for forming electrostatic interactions with drugs. The Glu90 neutralizes the positive charge of R6G and Glu120 is for neutralization of ethidium bromide; the positively charged nitrogen moiety of dequalinium in R6G binding pocket was neutralized by Glu57 and Glu58, while the other one in Et binding pocket was neutralized by Glu120. The single delocalized positive charge of crystal violet was neutralized by Glu90 and Glu120 (Schumacher *et al.*, 2001; Grkovic *et al.*, 2002; Schumacher *et al.*, 2003).

The structure of the QacR-Pf-Et ternary complex demonstrates the structural mechanism for simultaneous binding of two different drugs by a multidrug-binding protein (Figure 1.11 A-C). The near-UVCD binding assays together with relevant QacR–drug complex structures provide insight into the possible structural basis of competitive, noncompetitive, uncompetitive and cooperative multidrug binding (Schumacher *et al.,* 2004). For example, both R6G and Pf bind to the R6G binding pocket and sterically interfere with the binding of the other, so as to display competitive binding; Pf and Et display noncompetitive binding because of partial overlap of their binding pockets (Schumacher *et al.,* 2004).

Two bivalent diamidines, pentamidine and hexamidine, structurally differ by only one methylene carbon in the alkyl chain linker. The binding mode of hexamidine to QacR is similar to dequalinium, while pentamidine adopts a novel binding mode, with one of its benzamidine groups interacting with residue Glu63, and the other is neutralized by carbonyl and side chain oxygen atoms, demonstrating that a formal negative charge is not a prerequisite for binding positively charged drugs and underscoring the versatility of the QacR, which may also be the case for other multidrug-binding pockets (Fig.1.11 D, E. Murray *et al.,* 2004).

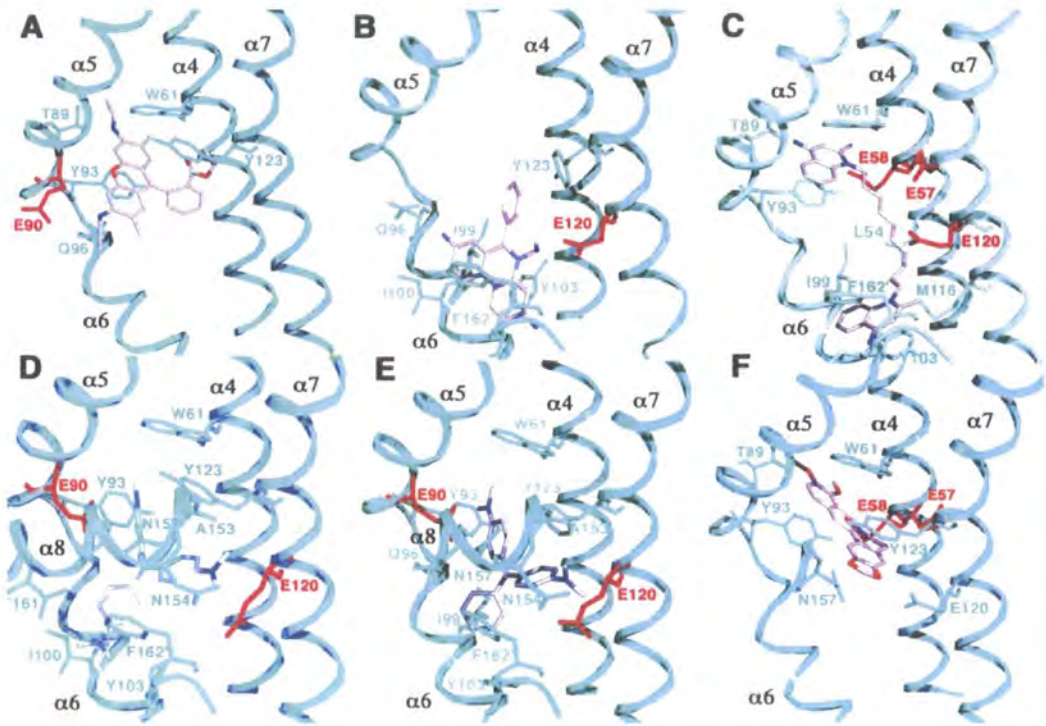


Figure 1.10. Binding of substrates in the extended QacR ligand-binding pocket. The key QacR residues and relevant α-helices in the ligand-binding site are shown. The acidic residues are colored red. A to F represent QacR in complex with R6G, Et, Dq, CV, MG and Be, respectively (Adapted from Schumacher *et al.,* 2001; Murray *et al.,* 2004 and Schumacher *et al.,* 2004).

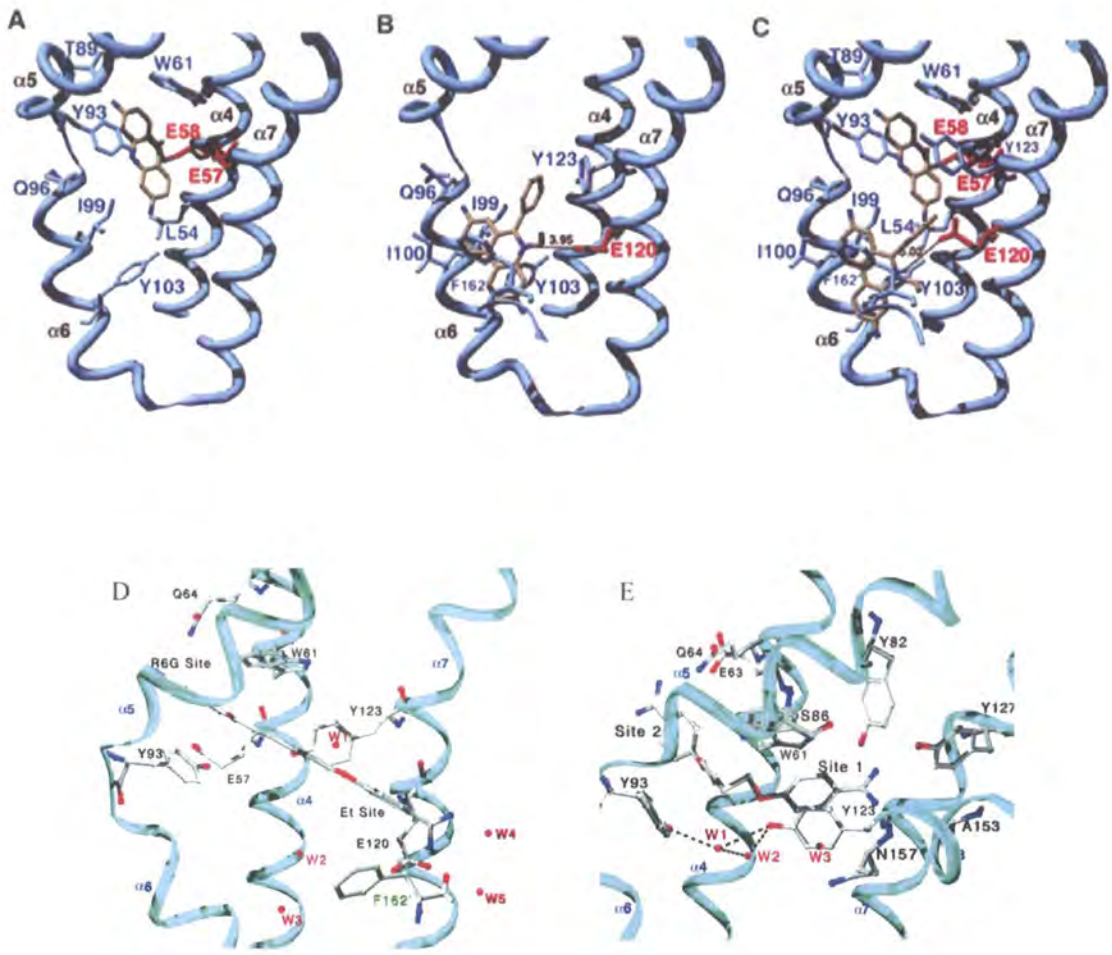


Figure 1.11. Multidrug recognition by QacR. (A) QacR–Pf binary complex. (B) QacR–Et binary complex. (C) QacR–Pf–Et ternary complex. (D) QacR-hexamidine complex. (E) QacR-pentamidine complex (Adapted from Schumacher *et al.*, 2004 and Murray *et al.*, 2004).

1.7 Tripartite multidrug efflux pumps in Gram negative bacteria

The envelope of Gram-negative bacteria contains two distinct membrane lipid bilayers. The inner membrane layer is a phospholipid bilayer, consisting largely of phosphatidylethanolamine (PE), phosphatidylglycerol, and cardiolipin; the outer membrane layer is asymmetric bilayer: the lipids on the inner and outer surface are glycerophospholipids and lipopolysaccharides, respectively. The shell between the inner and outer membranes layer, so-called periplasm, composes of peptidoglycan that confers mechanical robustness to the dual membrane system (Sharff *et al.*, 2001; Yu *et al.*, 2003).

The two membrane layers of Gram-negative bacteria can drastically slow down the influx of most antibiotics; however, these barriers cannot prevent the toxic effects of the drugs once they have entered the cell. The multidrug transporters in the inner membrane could form a complex with its partner proteins and actively transport antibiotics across the two membrane layers and the intervening periplasmic spaces in order to achieve their characteristic levels of intrinsic resistance (Nikaido, 2001). Generally these transport systems consist of an inner-membrane protein (IMP), usually a MF (Furukawa *et al.*, 1993; Lee *et al.*, 1999) or RND (Mokhonov *et al.*, 2005) or ABC (Rouquette-Loughlin *et al.*, 2005) family transporter, a membrane fusion protein (MFP), such as AcrA (Mikolosko *et al.*, 2006), MexA (Higgins *et al.*, 2004; Akama *et al.*, 2004), and an outer membrane protein (OMP), such as TolC (Koronakis *et al.*, 2000), forming a tripartite multidrug efflux pump. Hundreds of these tripartite pumps have been identified in Gram-negative bacteria (Schweizer, 2003; Poole, 2004). The structure of three components, TolC, AcrA and AcrB, provide us with clues to the function of these tripartite efflux complexes.

1.7.1 Structure and mechanism of TolC

TolC is the outer membrane component of the tripartite multidrug efflux pump from *E.coli.* The crystal structure of TolC was determined at 2.1Å resolution, which is shown in Figure 1.12 (Koronakis *et al.,* 2000).

The functional unit of TolC is a homotrimer with a long axis measuring 140Å. The TolC molecule is composed of three domains: a β-barrel domain, a α-helical barrel domain, and a mixed α/β-domain (equatorial domain). The body has a large interior cavity that is mostly solvent-filled with an average accessible interior diameter of 19.8 Å and a volume of roughly 43,000Å^3.

The 12-stranded β-barrel domain of TolC is assembled from three protomers, with each protomer contributing four β-strands. The β-barrel domain is 40 Å in length. It inserts into the outer membrane and is wide open to the external medium.

The 100 Å-long α-helical barrel domain is contiguous with the β-barrel and protrudes into the periplasm. The α-helices also form a 12-stranded antiparallel barrel, with each protomer contributing four α-helices (two continuous long helices and two pairs of shorter helices). While the 12-stranded β-barrel is right-twisted, this α-helical barrel is left twisted. The superhelical twist at upper section of the α-helical domain (near the β-barrel) tends to untwist at the distal end compared with conventional coiled coils. The α-helices in the lower (β-barrel distal) half of α-helical barrel form six pairs of conventional antiparallel coiled-coils, with one pair from each protomer folding inward, which constricts the periplasmic entrance to a resting closed state with an effective diameter of ~3.9 Å.

The equatorial domain is connected to the helical barrel at the junction of short helices. The helices and strands of this mixed α/β structure pack against the helices of the helical barrel.

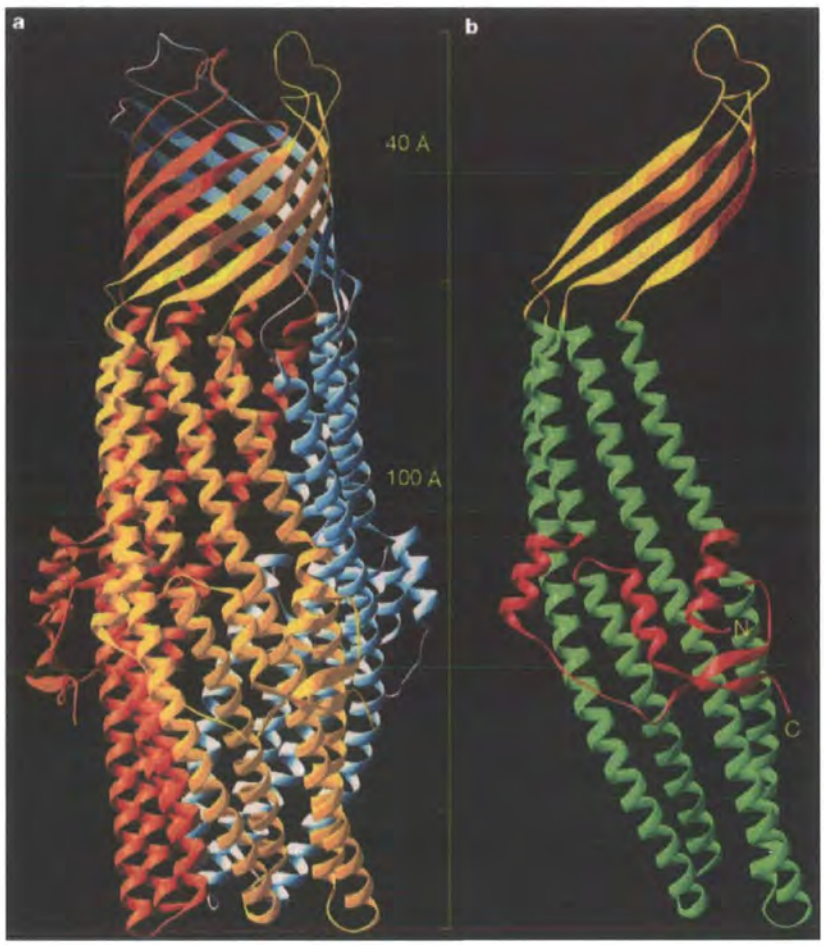


Figure 1.12. Structure of TolC. (a) Side view of a ribbon representation. Three protomers are individually colored (yellow, blue and red). (b) Ribbon representation of a TolC protomer. The β-barrel domain is yellow, the α-helix domain is green and the equatorial domain is red (Adapted from Koronakis *et al.*, 2000).

The substrates are presumed to channel through the translocase to the periplasmic entrance of TolC. TolC must undergo a conformational change to open the entrance, as it is too small for the substrates to enter the channel. An allosteric mechanism has been proposed for TolC opening (Koronakis *et al.*, 2000; Koronakis *et al.*, 2004). According to this proposed model, the inner coil (comprising helices H7 and H8) of each protomer realign relative to the outer coil, thereby enlarging the aperture diameter. The intra-monomer hydrogen bonds between residues Asp^{153}-Try^{382} and Gln^{136}-Glu^{359}, and the inter-monomer salt bridge interaction between Asp^{153} of a monomer and Arg^{367} of the adjacent one, were shown to stabilize the closed state of TolC (Andersen *et al.* 2002). By disrupting these interactions, the conductance of purified TolC proteins in black lipid bilayers increases six to ten times over wild-type conductance, which would be compatible with an aperture of 16 Å (Andersen *et al.* 2002). Eswaran *et al.* (2003) found that the function of TolC was abolished by introducing disulphide bonds to constrain the entrance coiled coils in the close state. These evidences of *in vivo* and *in vitro* experiments support the proposed mechanism of TolC opening. Vediyappan *et al.* (2006) found that the hybrid system VceAB-TolC is functional, while AcrAB-VceC is not functional; however, some point mutations at the periplasmic tip of VceC restored the functionality of the AcrAB-VceC pump. Bokma *et al.* (2006) identified several MexAB-adapted TolC mutants with amino acid substitutions in the lower α-helical barrel of TolC. These mutations converge to the MexAB partner OprM. The interaction variations between VceC and VceC mutants with AcrAB, as well as TolC and TolC variants with MexAB, may give us the clues that underlie the functional assembly and specificity of OMPs with their partner proteins IMPs-MFPs.

1.7. 2 Structure and mechanism of MexA and AcrA

MexA is a MFP component of the tripartite multidrug efflux pump MexAB-OprM from *P. aeruginosa*. The structure of residues 29–259 of the 360 residues mature

processed form of MexA has been solved at 2.4 Å and 3 Å resolution (Higgins *et al.*, 2004; Akama *et al.*, 2004). The crystal structure of MexA is illustrated in Figure 1.13.

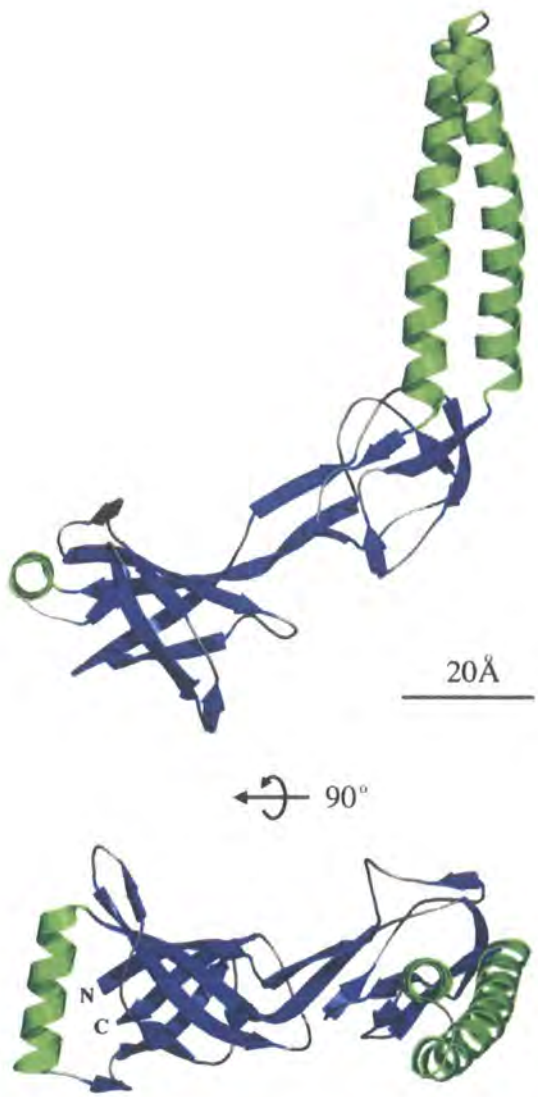


Figure 1.13. Two perpendicular views of the crystal structure of MexA monomer. The α-helices are green, the β-strands are blue. The α-helical hairpin is to the right; the lipoyl domain is central, and the β-barrel to the left (Adapted from Higgins *et al.*, 2004).

The asymmetric unit of the MexA crystal contained 13 molecules, which have the same conformation. The monomer of MexA has an elongated structure and is composed of three linearly arranged subdomains: a β-barrel domain, a lipoyl domain, and a 47-Å-long α-helical hairpin domain comprising a straight C-terminal helix and a N-terminal left handed superhelical twist.

The β-barrel domain contains six anti-parallel β-strands and a single α-helix situated at one entrance to the barrel.

The lipoyl domain consists of two interlocking lipoyl motifs of four β-strands, which is structurally homologous to the lipoyl domain of pyruvate dehydrogenase. Although these motifs are conserved throughout the family of adaptor proteins, they are separated by variable lengths of intervening sequence.

The α-helical hairpin domain forms a conventional antiparallel coiled-coil, with four heptad repeats in each of its helices. The large, hydrophobic side chains in the **a** and **d** positions, Ala residues in the **f** position and hydrophilic residues such as Ser or Glu in the **c** position, are conserved throughout the family of periplasmic adaptors.

AcrA is a MFP component of the tripartite multidrug efflux pump AcrAB-TolC form *E.coli*. The crystal structure of a 28 kDa core of AcrA (residues 45-312) has been solved at 2.7 Å resolution (Figure 1.14. Mikolosko *et al.,* 2006). The asymmetric unit of the AcrA crystal contained four molecules. The three domains of AcrA (45-312) monomer, a β-barrel domain, a lipoyl domain and a α-helical hairpin domain, share the same overall fold as that of MexA. The α-helical hairpin of the four monomers has different orientations. A hinge between the α-helical hairpin and lipoyl domain is composed of residues 99-106 in α-helix 1 and residues 169-173 in α-helix 2. The helical winding of the hinge causes such flexibility, which may play a functional role in coupling between AcrA conformations and TolC channel opening (Mikolosko *et al.,* 2006).

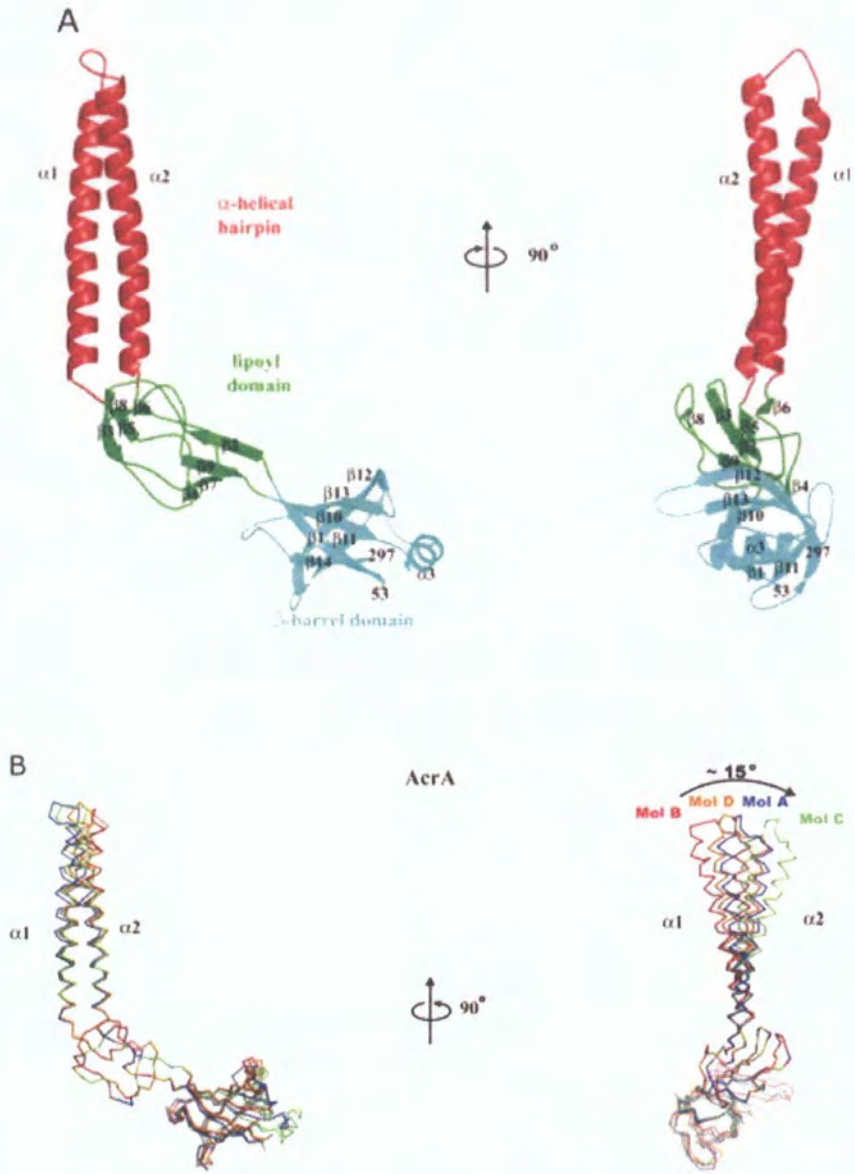


Figure 1.14. (A) Two perpendicular views of the crystal structure of AcrA (45-312) monomer, with the α-helical hairpin domain in red, lipoyl domain in green and β-barrel domain in cyan. (B) Superimposed AcrA (45-312) of four conformations in the asymmetric unit of the crystals show the orientations difference of the α-helical hairpin (Adapted from Mikolosko *et al.,* 2006).

The mechanisms of MFP have been studied intensely. Stegmeier *et al.* (2006) found that the hybrid system AcrAB-OprM is not functional; however, exchange of the hairpin domain of AcrA with that of MexA restored the function of the pump, indicating that the MexA hairpin domain is involved in the functional interaction with OprM. Elkins *et al.* (2003) swapped sections between two MFPs, AcrA and YhiU. The 290-residucs N-terminal segment and the residues from 357 to 397 of the 398-residue protein AcrA could be replaced with a sequence coding for the corresponding region of YhiU, but replacement of the region between residues 290 and 357 produced a protein incapable of functioning with AcrB, indicating that a small region of AcrA close to its C terminus is involved in the interaction with AcrB. Ip *et al.* (2003) found that AcrA undergoes reversible conformational changes induced by a narrow pH range, suggesting that the fluctuations of pH in the periplasm accompanying the AcrB-mediated drug efflux could act as a signal to trigger the action of AcrA. Borges-Walmsley *et al.* (2003) identified drug-binding domains of a MFP EmrA from *E.coli*, indicating that MFP could not only act as a channel between the IMP and OMP, but also involve in drug translocation. Yoneyama *et al.* (2000) revealed that the periplasmic domain of MexA expressed in the periplasm has the same function as that of the wide-type MexA protein. Cross-linking of *in vivo* complexes has identified MFP trimers in both drug efflux and protein export (Thanabalu *et al.*, 1998; Zgurskaya *et al.*, 2000), but the oligomeric state of the MFP in active pumps unclear yet.

1.7.3 The three components of tripartite multidrug efflux pumps form a stable intermembrane complex

The assembly of three components of type I protein translocation apparatus in Gram negative bacteria has been studied in three distinct systems: Has, Prt and Hly (Letoffe *et al.*, 1996; Thanabalu *et al.*, 1998). In these cases, the substrate triggers the stepwise

recruitment of the three components to form an intermembrane channel and the complex disintegrated after translocation.

The assembly of three components of tripartite multidrug efflux pumps has been studied in the AcrAB-TolC and MexAB-TolC systems recently.

Tikhonova *et al.* (2004) analyzed the AcrAB-TolC complex *in vivo* using sucrose gradient fractionation and pull-down assays. The sucrose gradient fractionation study revealed that in the absence of AcrA, the amount of AcrB detected in the OM fraction was far smaller then that in the presence of AcrA, indicating the important role of AcrA in complex formation. TolC is required for AcrB fractionation with the OM, suggesting tight adhesion between AcrAB and TolC. In the pull-down assay, the affinity tag is attached to only one of the three components of AcrAB-TolC and tested the assembly of complex with or without substrates and proton-motive force. All the three components co-purified when the His_6 tag is fused to AcrA, AcrB or TolC, indicating that the AcrAB-TolC form a stable intermembrane complex *in vivo*.

Touze *et al.* (2004) used *in vivo* cross-linking to detect interactions between the components of AcrAB-TolC, indicating that AcrA could be cross-linked independently to AcrB and TolC, and TolC could be cross-linked independently to AcrB. Isothermal titration calorimetry (ITC) was applied to characterize purified proteins *in vitro*, suggesting that the AcrA-AcrB and AcrA-TolC interactions occurred spontaneously as they were associated with favourable free energy changes. ITC showed no association between AcrB and TolC, implying that AcrA is required to link these two components stably in the tripartite pump.

Tamura *et al.* (2005) examined the interaction sites of AcrB and TolC using *in vivo* disulfide cross-linking between cysteine residues introduced to AcrB and TolC. The formation of spontaneous disulfide bonds between positions 255, 256, 257 in AcrB and position 147 in TolC suggests that the tops of the vertical hairpins of AcrB and

the bottoms of the coiled coils of TolC are in close proximity, supporting the direct docking model.

Gerken *et al.* (2004) found that mutant AcrA can be stabilized by TolC and *vice versa*, and the functional interaction between $TolC_{P246R, S350C}$ and AcrA are dependent on AcrB, which provide genetic evidence that a functional complex exists between TolC and AcrAB *in vivo*.

Mokhonov *et al.* (2004) also used *in vivo* cross-linking and pull-down assays to analyze the MexAB-OprM complex from *P. aeruginosa* and gave similar overall results with that of the AcrAB-TolC system.

All the interactions between the components *in vivo* exist even in the absence of substrates and proton-motive force. Taken together, these data suggest that the three components of tripartite multidrug efflux pumps form a stable inter-membrane complex.

1.7.4 The structural models for tripartite multidrug efflux pumps

The crystal structures of the three components of tripartite multidrug efflux pump, AcrB, AcrA and TolC, have been determined. The question remains as to how the three components assemble to form a functional pump. Different models based on these crystal structures have been proposed.

Higgins *et al.* (2004) proposed a sheath-like model in which nine MexA monomers form a ring. The α-helical hairpins in the ring interact by means of the stripes of conserved small residues to form an α-barrel. The ring could form a sheath around the open state model of TolC and the inner membrane transporter, generating an

intermembrane channel. The IMP:MFP:OMP ratio was predicted to be 1:3:1 in this model (Figure 1.15).

Akama *et al.* (2004) devised homology models for MexB and OprM based upon the structures of AcrB and TolC. In the proposed model, the helical hairpins of three pairs of MexA dimers are inclined with the surface of OprM with an angle of about 20°. The lipoyl domain fits within the gap between MexB-OprM contact regions. The β-barrel domain and disordered domain may have interacted with the MexB trimer. The IMP:MFP:OMP ratio was predicted to be 1:2:1 in this model (Figure 1.16).

Fernandez-Recio *et al.* (2004) devised a homology model for AcrA based upon the structure of MexA. In the model proposed by Fernandez-Recio J et al., AcrB present a canyon-like domain along the interface of the periplasmic domain which can leave space for AcrA to bind. Three AcrA fit simultaneously in the interprotomer grooves of AcrB and the open state of TolC with modest structural adjustment. The IMP:MFP:OMP ratio was predicted to be 1:1:1 in this model (Figure 1.17).

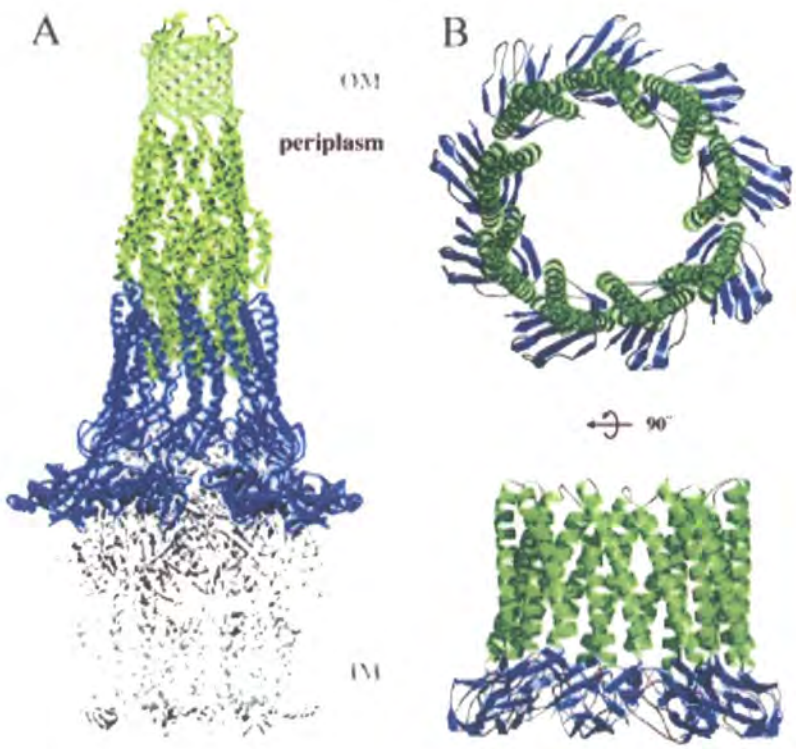


Figure 1.15. Model of the assembled tripartite drug efflux pump. (A) The open-state model of TolC (green) contacts the TolC docking domain of AcrB (white). Nine MexA molecules (blue) form a sheath around AcrB and the α-barrel of TolC. (B) Two perpendicular views of the sheath formed by nine MexA molecules (Adapted from Higgins *et al.*, 2004).

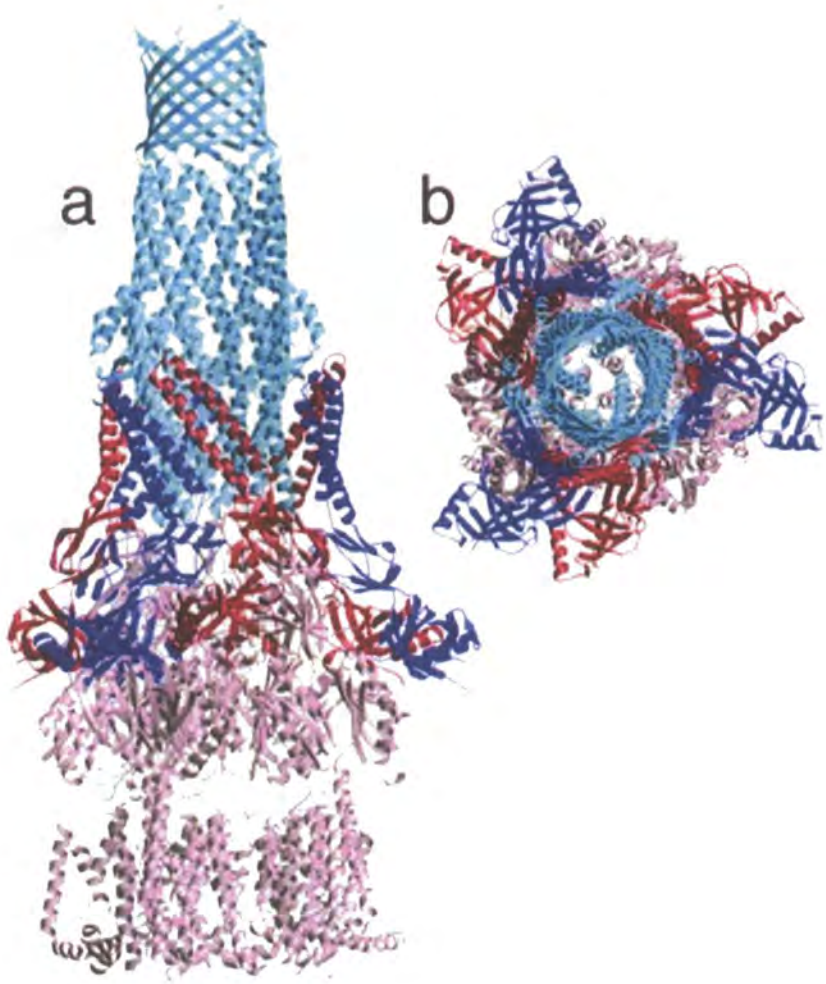


Figure 1.16. Side and top view of the model of assembled tripartite drug efflux pump. The helical hairpins of three pairs of MexA dimers (blue and red) are inclined with the surface of OprM (light sky blue). The lipoyl domain fits with the gap between MexB-OprM contact regions. The β-barrel domain and disordered domain interact with MexB (pink) (Adapted from Akama *et al.*, 2004).

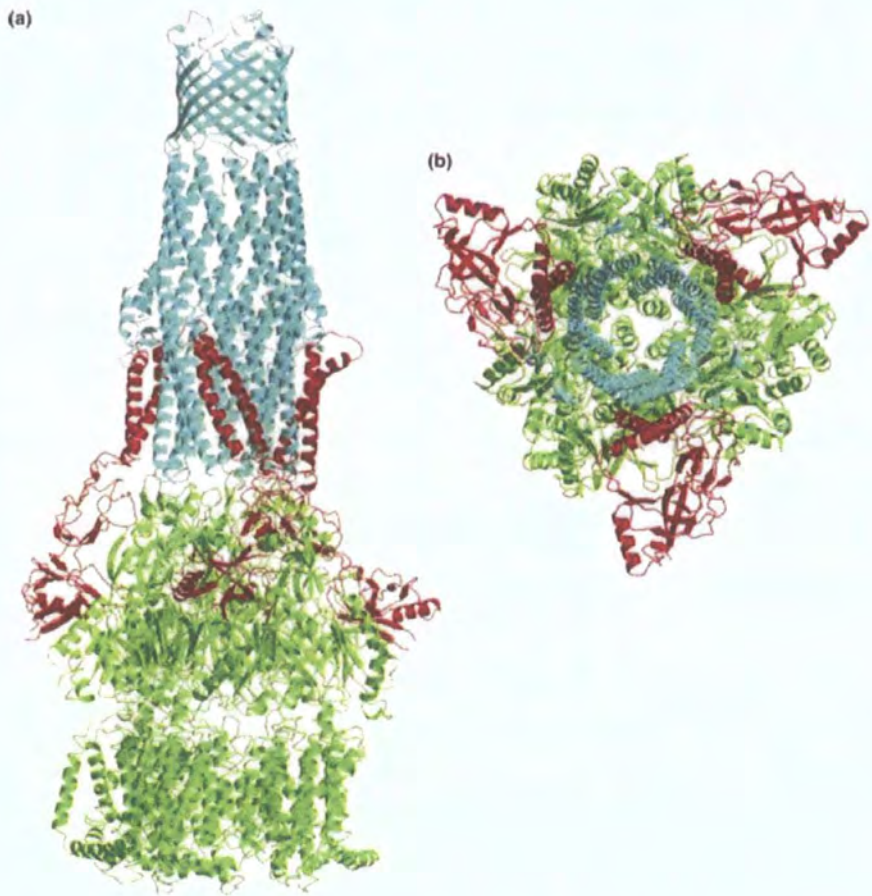


Figure 1.17. Side and top view of the model of assembled tripartite drug efflux pump. Three AcrA (red) molecules fit simultaneously in the interprotomer grooves of AcrB (green) and the open state of TolC (light sky blue) (Adapted from Fernandez-Recio *et al.*, 2004).

1.8 Perspective

The lack of structure of multidrug efflux pumps has hindered our understanding of these transporters. The structural determination of multidrug efflux pumps EmrD, AcrB, Sav1866, and the outer membrane protein TolC, as well as the membrane fusion protein AcrA and MexA etc, probably represent the most significant progress that has been made recently. The combined biochemical and structural data have provided us a closer vision to structure-function relationships in these multidrug transporters.

The structure of EmrD is likely to act as a paradigm for other MFS transporters with 12-TMS; however, the structure of MFS transporters with 14-TMS is not available. The structure of EmrD in complex with substrate will also be essential for us to understand the mechanism of substrate recognition and translocation of these family transporters.

The structural asymmetry of AcrB trimer suggests an ordered binding change mechanism (Seeger *et al.*, 2006). Biochemical evidence is necessary to support such a proposed mechanism. Multiple alignments of amino acid sequences and phylogenetic analysis indicate that there are structural and/or functional differences between AcrB and other RND family transporter, such as MtrD (Paulsen *et al.*, 1996). The knowledge of the difference between them will be critical for understanding the mechanism of these family transporters.

The crystal structure of EmrE is not yet solved. There are still debates as to the membrane topology and functional unit of EmrE (Tate, 2006; Soskine *et al.*, 2006). The crystal structure of SMR family transporters with two dissimilar but homologous subunits, such as EbrAB and YkkCD, will be helpful for solving these debates.

The homology between TMDs in different subfamilies of ABC transporters is lower, indicating that a different transport mechanism is employed. The crystal structures of ABC transporters in different subfamilies will be critical for understanding their mechanism.

The MATE family transporters are the least well characterized. Structural or functional information on the molecular mechanism underlying the Na^+ (or H^+)/multidrug antiport by the MATE family is still very limited at present.

In addition, it is not clear yet how multidrug efflux pumps couple transport to proton or sodium motive force or free energy by ATP hydrolysis. Future studies to pursue will be to solve the crystal structure of the pumps that represents an intermediate point in the efflux process.

The protomer stoichiometry for the tripartite efflux pump of Gram-negative bacteria is still uncertain. It is likely that different arrangement of MFPs can be expected in MFS, RND and ABC-based tripartite pump. The crystal structure of these tripartite efflux pumps and the details underlying the dynamics (e.g. key coiled-coil interactions and channel gating) will be critical for deciphering their mechanism (Eswaran *et al.*, 2004).

Chapter 2

Materials and Methods

2. 1 Reagents and equipments

2. 1. 1 Sources of reagents

General chemical reagents such as salts with analytical purity grade were acquired from Sigma (UK) Ltd. Microbiological growth media was supplied by Oxoid, Difco or Melford Laboratories. Enzymes used in molecular biology protocols were purchased from Promega (UK) Ltd unless otherwise stated. Detergents such as *n*-dodecyl-β-D-maltoside (DDM) were purchased from Calbiochem.

The sources of additional reagents used in specialised applications are detailed in association with specific protocols.

2. 1. 2 Preparation of solutions

Gross quantities of reagents were weighed on a Sartorius digital balance whilst smaller quantities were measured using a Mettler Toledo digital analytical balance. Microbiological growth media and molecular biology reagents were prepared using milliQ distilled/deionized water where appropriate. Recipes of stock solutions are shown in table 2.1.

Table 2.1 stock solutions

Carbenicillin (disodium salt)	100 mg/ml in deionized water. Filter sterilize and store at –20°C. Use at 100 μg/ml.
Kanamycin (sulfate)	20 mg/ml in deionized water. Filter sterilize and store at –20°C. Use at 25 μg/ml for cells containing kan^R plasmids, and at 15 μg/ml for cells with a chromosomal kan^R gene.
Chloramphenicol	34 mg/ml in ethanol. Store at –20°C. Use at 34 μg/ml.
1 M IPTG (isopropyl-D-thiogalactopyranoside)	2.38 g IPTG in 10 ml deionized water. Filter sterilize and store at –20°C.
20% L-arabinose (w/vol)	L-arabinose solution in deionized water. Filter sterilize and store at –20°C.
Rifampicin	10 mg/ml in 67% methanol, 0.17 N NaOH. Use at 200 μg/ml within 5 days. Protect from light.
Tetracycline	5 mg/ml in ethanol. Store at –20°C. Use at 12.5 μg/ml.
20% Glucose (w/vol)	D-glucose solution in H_20. Autoclave. Store sterile solution at room temperature.
80% Glycerol (v/v)	Mix 80 ml glycerol with 20 ml deionized water. Autoclave. Store sterile solution at room temperature.
X-gal	50 mg/ml in DSMO, Store at –20°C. Protect from light.

2. 1. 3 Sterilization of chemical and media solution

Media and reagents to be sterilised were either autoclaved or filter sterilized depending on their stability to heating. Heat stable materials, such as growth media, were heated to 121 °C for 15 minutes at 4 atmospheres pressure in a Priorclave electrically heated autoclave set on liquid cycle. Solutions of heat labile reagents, such as antibiotics, were sterilized by filtration through 0.22 μm syringe filters. Non-autoclavable items such as centrifuge tubes were sterilized by washing thoroughly in 100% ethanol followed by a drying period in a laminar flow hood.

2. 1. 4 Centrifugation

Routine centrifugation of materials in microcentrifuge tubes was performed in a Sigma benchtop microfuge fitted with a 24 place fixed angle rotor, and a Jouan CR3 refrigerated benchtop centrifuge was used where there was a requirement for cooling. Larger volumes of liquids were processed in a Beckman-Coulter Avanti J-E refrigerated centrifuge using different fixed angle rotors. A Beckman JA-10 rotor was used to process volumes up to 3 liters in 500 ml tubes and a Beckman JA-20 rotor was used to process up to 300 ml in 40 ml tubes. High-speed ultracentrifugation was performed in a Beckman L8M ultracentrifuge using a Beckman Ti50 fixed angle rotor.

2. 2 Bacterial strains

All bacterial strains used are listed in table 2.2.

Table 2.2 – Bacterial strains used in these studies

Strains or Plasmids	**Genotype**	**Source**	**Description/Applica -tion**
V. cholerae vaccine strain CVD101	CT-A- CT-B+ deletion derivative of classical *V. cholerae* strain Ogawa 395	J. B. Kaper	Used as the source of genomic DNA for PCR
E. coli XL1-Blue	*recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac* [F′ *proAB lacI*^{q}*ZΔM15* Tn*10* (Tet^{r})].	Stratagene	General cloning, Site-directed mutagenesis
E. coli NovaBlue	*endA1 hsdR17*(*rK12*$^{-}$ *mK12*$^{+}$) *supE44 thi-1 recA1 gyrA96 relA1 lac*F'[*proA*$^{+}$*B*$^{+}$ *lacIqZΔM15* ::Tn*10*(TcR)]	Novagen	non-expression host, general purpose cloning, plasmid preparation.
E. coli C41(DE3)	F^{-} *ompT hsdSB* (*rB*$^{-}$ *mB*$^{-}$) *gal dcm* (DE3)	Avidis S.A.	Mutant strain from BL21(DE3), general purpose expression host T7 Expression systems.
E. coli C43(DE3)	F^{-} *ompT hsdSB* (*rB*$^{-}$ *mB*$^{-}$) *gal dcm* (DE3)	Avidis S.A.	Mutant strain from BL21 (DE3), general purpose expression host for T7 Expression systems.
E. coli BL21-AI	F^{-} *omp*T *hsd*S_B ($r_B^{-}m_B^{-}$) *gal dcm ara*B::T7RNAP-tetA	Invitrogen	High-stringency expression host for T7 Expression

			systems
E. coli TOP10	F⁻ *mcr*A Δ(*mrr*⁻*hsd*RMS⁻*mcr*BC) φ80*lac*ZΔM15 Δ*lac*X74 *rec*A1 *ara*D139 Δ(*araleu*)7697 *gal*U *gal*K *rps*L (StrR) *end*A1 *nup*G	Invitrogen	General cloning, High-stringency expression host for P_{BAD} Expression systems
E. coli M15(pREP4)	$Nal^S Str^S Rif^S Thi^-$ $Lac^- Ara^+ Gal^+ Mtl^-$ $F^- RecA^+ Uvr^+ Lon^+$/*pREP4*	Qiagen	High-stringency expression host for T5 Expression systems
E. coli N43	F⁻ *acrA1 lac ara xyl rpsL*	*E. coli* Genetic Stock Centre	Complementation studies

2. 2. 1 Growth media for bacteria

The growth media for liquid cultures of the *Escherichia coli* strains were listed as follows:

Luria-Bertani (LB) broth (per Liter)

10 g of NaCl
10 g of tryptone
5 g of yeast extract
Add deionized H_2O to a final volume of 1 liter, Adjust pH to 7.0 with 5 N NaOH, Autoclave.

2x YT broth (per Liter)

5g of NaCl
16g of Tryptone
10g of Yeast extract
Add deionized H_2O to a final volume of 1 liter, Adjust pH to 7.0 with 5 N NaOH, Autoclave.

SOB Medium (per Liter)

20.0 g of tryptone
5.0 g of yeast extract
0.5 g of NaCl
Add deionized H_2O to a final volume of 1 liter, autoclave. Add 10 ml of filter-sterilized 1 M $MgCl_2$ and 10 ml of filter-sterilized 1 M $MgSO_4$ prior to use.

SOC Medium (per 100 ml)

Note: This medium should be prepared immediately before use.
Add 2 to ml of filter-sterilized 20% (w/v) glucose or 1 ml of filter-sterilized 2 M glucose SOB medium (autoclaved) to a final volume of 100 ml.

NZY^+ Broth (per Liter)

10 g of NZ amine (casein hydrolysate)
5 g of yeast extract
5 g of NaCl
Add deionized H_2O to a final volume of 1 liter, adjust to pH 7.5 using NaOH, autoclave.
Add the following filer-sterilized supplements prior to use:
12.5 ml of 1 M MgCl2
12.5 ml of 1 M MgSO4
20 ml of 20% (w/v) glucose (or 10 ml of 2 M glucose)

LB agar plate was prepared by adding 1.5% w/v agar to liquid LB medium prior to autoclaving.

2.2.2 Storage of bacterial strains

Permanent stocks of bacterial strains were kept as glycerol stocks. Note that high glycerol concentrations (> 10%) may lead to plasmid instability.

2.2.2.1 Preparation stock cultures of bacterial strains

A single colony was inoculated into 20 ml LB medium containing appropriate antibiotic(s) in a 50 ml centrifuge tube. The incubation was shaken vigorously at 37°C during the day until the OD_{600} reached 0.6–0.8. Then 0.9 ml of the culture was removed and transferred to a cryovial, 0.1 ml 80% glycerol was added, mixed well and stored at –70°C.

Plasmid-bearing strains, particularly those having any tendency toward instability, were titered at the time of freezing to be sure that the vast majority of cells in the culture have the intended host-plasmid combination.

2.2.2.2 Inoculation of a culture from the frozen stock

A few microliters of the frozen stock were scraped or melted from the surface (use a sterile pipettete tip or plastic culture loop) and streaked on an agar plate or inoculated into liquid medium (containing appropriate antibiotic(s)). The remainder was returned to the –70°C freezer without thawing.

2.3 Isolation and analysis of nucleic acids

2.3.1 Plasmid DNA purification

Small-scale purification of plasmid DNA from *E.coli* host strains was performed using the QIAprep Spin Miniprep Kit (Qiagen) and the protocol provided, which is summarised below.

Ten ml of Luria-Bertani (LB) broth containing appropriate antibiotics were inoculated with a single colony of transformed bacteria. The culture was incubated at 37°C with vigorous shaking for 12–16 hours. Bacterial cells were harvested by centrifugation at 4000 ×g at 4 °C for 10 min and the culture medium was decanted away from the cell pellet. The pelleted cells were resuspended in 250 µl Buffer P1 supplied with the kit and transferred to a microcentrifuge tube. The cells were then lysed by adding 250 µl Buffer P2 and mixing thoroughly by inverting the tube 4–6 times. The mixture was then neutralised by adding 350 µl Buffer N3 and mixing immediately and thoroughly by inverting the tube 4–6 times. The insoluble and soluble components were separated by centrifuging for 10 min at 13,000 ×g in a table-top microcentrifuge. The supernatant fraction, which contains the soluble plasmid DNA, was applied to the QIAprep spin column by pipetteting. The supernatant was spun through the column for 30–60 s; the flow-through was discarded. The column contains a positively charged silica-based matrix upon which plasmid DNA becomes immobilised. The immobilised plasmid DNA was washed by adding 0.75 ml Buffer PB and centrifuging for 30–60 s, the flow-through was discarded. The column was washed again by adding 0.75 ml Buffer PE and centrifuging for 30–60 s, the flow-through was discarded and the column was centrifuged for an additional 1 min to remove residual wash buffer. The column was placed in a clean 1.5 ml microcentrifuge tube. To elute DNA, 50 µl Buffer EB (10 mM Tris·Cl, pH 8.5) or deionized water was added to the centre of each column and the column was allowed to sit at room temperature for 1

min, and then the column was centrifuged for 1 min. Plasmid DNA was then analyzed by agarose gel electrophoresis. Purified Plasmid DNA was stored at -20°C.

2.3.2 Agarose gel electrophoresis of DNA

Electrophoresis buffer:

TAE: 40 mM Tris-acetate 1 mM EDTA

Stock solution: 25×

121 g Tris base

28.6 ml of glacial acetic acid

50 ml of 0.5 M EDTA (pH: 8.0)

6× Gel-loading buffers

0.25% bromophenol blue

0.25 xylene cyanol FF

30% glycerol in dH_2O

Storage temperature: 4 °C

The protocol of agarose gel electrophoresis (Molecular Cloning, Handbook) was used to analyze the results of nucleic acid manipulations. Agarose was dissolved in 1×TAE buffer at a concentration appropriate for separating the particular size fragment expected in the DNA samples by heating the TAE/agarose mixture in a microwave oven. The molten gel was allowed to cool to about 55°C, and then ethidium bromide was added to a final concentration of 0.5μg/ml. An appropriate comb was positioned 0.5-1.0 mm above a horizontal gel-tray to form the sample slots in the gel, and the warm agarose solution was poured into the gel-tray. The gel was set for about 30 min at room temperature. After the gel had solidified, the comb was removed and the gel slab was placed in the electrophoresis tank containing 1×TAE buffer allowing for approximately 5 mm depth over the gel. Each DNA sample was mixed with 0.2 volume of 6×gel-loading buffers, and then the sample mixture was loaded into a separate well of the submerged gel. The DNA size standard (1 KB Plus DNA Ladder, Invitrogen) was loaded into one well. The lid of the gel tank was closed and the electrical leads were attached so that the DNA will migrate toward the positive anode,

and a voltage of 1-5 V/cm was applied. Progress of DNA through the gel was monitored by the migration of xylene cyanol and bromophenol blue. When the DNA samples had migrated a sufficient distance through the gel, the gel was examined by UV light and photographed using a Gene Genius Bio Imaging System (SYNGRNE).

2.3.3 Extraction and purification of DNA from agarose gels

Double-stranded DNA fragments from PCR and other enzymatic reactions need to be separated by agarose gel eletrophoresis. Extraction and purification of DNA from agarose gels was performed using a QIAquick Gel Extraction Kit (Qiagen). The manufacturer's protocol was adopted, which is summarised below.

A gel slice containing a desired DNA fragment was excised from the agarose gel slab with a clean, sharp scalpel. The size of gel slice was minimized by removing extra agarose and was weighed in a microcentrifuge tube. Three volumes of buffer QG were added to one volume of gel (100 mg ~ 100 µl). The gel slice was then dissolved in buffer QG by incubating in a 50°C water bath for about 10 min, with vortexing every 2–3 min during the incubation to help dissolve the gel. A QIAquick spin column was placed in a provided 2 ml collection tube. The DNA/buffer QG mixture was applied to the QIAquick column and centrifuged for 1 min to immobilize DNA. 0.5 ml of Buffer QG was added to QIAquick column and centrifuged for 1 min to remove any remaining traces of agarose which may interfere with subsequent steps. The immobilised DNA was washed by adding 0.75 ml of Buffer PE to QIAquick column and centrifuging for 1 min. The flow-through was discarded and the column was centrifuged for an additional 1 min to remove residual Buffer PE. The QIAquick column was place into a clean 1.5 ml microcentrifuge tube. The DNA was eluted by the addition of 50 µl of Buffer EB (10 mM Tris, pH 8.5) or deionized H_2O to the centre of the QIAquick membrane. The column was incubated at room temperature

for 1 min and then centrifuged for 1 min to collect the purified DNA solution. The eluted DNA was stored at -20°C.

2.4 *In vitro* amplification of DNA by the polymerase chain reaction (PCR)

The polymerase chain reaction (PCR) was used to amplify and modify specific genes from genomic DNA. The desired restriction sites were added to both termini of the target template by PCR using oligonucleotide primers to allow the ligation of amplified fragments from the initial cloning vector pGEM-T Easy into the expression vector.

Oligonucleotides used in this study were designed according to the published DNA sequence and synthesized by Invitrogen Custom Primers (Table 2.3).

Table 2. 3 – Oligonucleotides

Primer Name	Sequence (5′ to 3′)	Use (reference)
VceCF1	CAT ATG AAA AAT AGC GTT CAA ACG GTA GGT TTG TTG	PCR (chapter III)
VceCR	CTC GAG AGA TTC TGT TGT TTC AAA ACC GCC GCC TAG	PCR (chapter III)
VceCF2	GAG CTC AAA AAT AGC GTT CAA ACG GTA GGT TTG TTG	PCR (chapter III)
VceBR1	AAG CTT TTA ATG AAC AGC AGA GGT ATC CAC CGC TTG	PCR (chapter III)
VceC-E92CF	CAG CAC TCA CCA TCG TTA TGC ATG GCA ATG GCT CGG C	Mutagenesis (chapter III)
VceC-E92CR	GCC GAG CCA TTG CCA TGC ATA	Mutagenesis

	ACG ATG GTG AGT GCT G	(chapter III)
VceC-D115AF	CGC AAT TCG ATC ATT TGC GCT CGG GTT AGC GGC ATC C	Mutagenesis (chapter III)
VceC-D115AR	GGA TGC CGC TAA CCC GAG CGC AAA TGA TCG AAT TGC G	Mutagenesis (chapter III)
VceC-D153AF	CCT TGA ATT TTC AGT ATG CGT TTG ACT TTT GGG GAA AAA ATC GCG	Mutagenesis (chapter III)
VceC-D153AR	CGC GAT TTT TTC CCC AAA AGT CAA ACG CAT ACT GAA AAT TCA AGG	Mutagenesis (chapter III)
VceC-D303AF	CTG GGG CAT CGT GCT GCG ATC ACC GCT GCG CGT TG	Mutagenesis (chapter III)
VceC-D303AR	CAA CGC GCA GCG GTG ATC GCA GCA CGA TGC CCC AG	Mutagenesis (chapter III)
VceC-D328AF	CGC AAG CAC AAT TCT ATC CAG CGG TCA CTT TGT CAG CGT TTA TTG	Mutagenesis (chapter III)
VceC-D328AR	CAA TAA ACG CTG ACA AAG TGA CCG CTG GAT AGA ATT GTG CTT GCG	Mutagenesis (chapter III)
VceC-E382AF	CAG AAG CTC GTT ACC AAG CGG CGG TCG CTC AAT ACA ATG	Mutagenesis (chapter III)
VceC-E382AR	CAT TGT ATT GAG CGA CCG CCG CTT GGT AAC GAG CTT CTG	Mutagenesis (chapter III)
VceC-E397AF	GTG CAA GCC TTG CAT GCG ATC GCC GAT GTG GTG AC	Mutagenesis (chapter III)
VceC-E397AR	GTC ACC ACA TCG GCG ATC GCA TGC AAG GCT TGC AC	Mutagenesis (chapter III)

VceC-M93AF	GCA CTC ACC ATC GTT AGA GGC GGC AAT GGC TCG GCT TAA AG	Mutagenesis (chapter III)
VceC-M93AR	CTT TAA GCC GAG CCA TTG CCG CCT CTA ACG ATG GTG AGT GC	Mutagenesis (chapter III)
VceC-M93R97AF	GTT AGA GGC GGC AAT GGC TGC GCT TAA AGG GGC ACA AGG	Mutagenesis (chapter III)
VceC-M93R97AR	CCT TGT GCC CCT TTA AGC GCA GCC ATT GCC GCC TCT AAC	Mutagenesis (chapter III)
VceAF	GAG CTC AAA CAA CAG AAT CTT AAT AGG TTA TCT TCA	PCR (chapter V)
VceA12F	GAG CTC ATG AAT TCA AAT AAT AGC AAC ACC GAA TTT	PCR (chapter V)
VceA53F	GAG CTC CAC TTT ATC GGT TCA CGC TAC ATT TCC ACC	PCR (chapter V)
ECFP-F	GGA TCC GTG AGC AAG GGC GAG GAG CTG TTC ACC GGG	PCR (chapter V)
ECFP-R	GAG CTC CTT GTA CAG CTC GTC CAT GCC GAG AGT GAT	PCR (chapter V)
EYFP-F	GTC GAC GTG AGC AAG GGC GAG GAG CTG TTC ACC GGG	PCR (chapter V)
EYFP-R	AAG CTT TTA CTT GTA CAG CTC GTC CAT GCC GAG AGT GAT	PCR (chapter V)
VceAR1	GTC GAC GCC TTG CTC TGA AAC TTT GGC GGC GCG	PCR (chapter V)
VceApF	GAG CTC CAT CAT CAT CAT CAT CAT GAC GAT GAC GAT AAG CAC TTT ATC GGT TCA CGC TAC ATT TCC ACC	PCR (chapter V)
VceR-PF	AGT ATA ACT GTA CGG TAC GGT	DNA binding

	TTA GTT AT	(Chapter VI)
VceR-PR	ATA ACT AAA CCG TAC CGT ACA GTT ATA CT	DNA binding (Chapter VI)
VceBF	CAT ATG AGT CAT AAC GCT GAC AAT GAG ATG CAA CCT	PCR (chapter VI)
VceBR2	CTC GAG ATG AAC AGC AGA GGT ATC CAC CGC TTG CTT	PCR (chapter VI)
HisR	GAG CTC GTG GTG GTG GTG GTG GTG CTC GAG	PCR (chapter VI)
VceA-EKF	GAG CTC GAT GAC GAT GAC AAG AAA CAA CAG AAT CTT AAT AGG TTA TCT TCA	PCR (chapter VI)
VceAR2	AAG CTT TTA GCC TTG CTC TGA AAC TTT GGC GGC GCG	PCR (chapter VI)
VceRF	GGA TCC CGG GTA AAA AGT GAA GAA AAA AGG CAA GCA	PCR (chapter VI)
VceRR	CTC GAG ATC AGC TGC GTA CAG GCA TAA AAA GGC CTC	PCR (chapter VI)

In table 2.3, the oligonucleotide names that end with 'F' and 'R' represent pairs of forward and reverse primers used for the amplification or mutagenesis of specific genes; the underlined sequences correspond to the restriction enzyme sites that were incorporated into amplicons in order to facilitate downstream manipulation of the amplified fragments.

The oligonucleotides were supplied on a 50 nmol scale and in a desalted, lyophilised form. The oligonucleotides were resuspended in high purity deionized water to a final concentration of 50 pmol/μl and stored at -20°C.

The HotStarTaq™ DNA polymerase (Qiagen) was used to amplify DNA templates less then 2 kb, which is a modified form of the recombinant 94 kDa Taq DNA polymerase from *Thermus aquaticus* (Qiagen). HotStarTaq™ DNA Polymerase has no polymerase activity at ambient temperatures, which may prevent the formation of misprimed products and primer–dimers at low temperatures. It can be activated by incubation at 95°C for 15-minute before thermal cycling. These features of HotStarTaq™ DNA Polymerase usually result in high PCR specificity and increase the yield of the specific PCR product (Qiagen). The reaction compositions using HotstarTaq™ DNA Polymerase were listed in table 2.4 (Qiagen).

Table 2.4 PCR components for HotstarTaq™ DNA Polymerase

Component	Volume/reaction	Final concentration
Master mix		
10x PCR Buffer*	10 µl	1x
dNTP mix (10 mM of each)	2 µl	200 µM of each dNTP
Primer A	1 µl	0.5 µM
Primer B	1 µl	0.5 µM
HotStarTaq DNA Polymerase	0.5 µl	2.5 units/reaction
Distilled water	Variable	–
Template DNA	Variable	≤1 µg/reaction
Total volume	100 µl	

* *Contains 15 mM $MgCl_2$*

A typical PCR cycling program was outlined below (Qiagen).

Initial activation step:	15 min	95°C
3-step cycling		
Denaturation:	0.5–1 min	94°C
Annealing:	0.5–1 min	50–68°C
Extension:	1 min/kb	72°C
Number of cycles:	30–35	
Final extension:	10 min	72°C

The ProofStart™ DNA polymerase (Qiagen) was used to amplify DNA templates more then 2 kb to avoid base pair mismatch. ProofStart™ DNA polymerase is a recombinant 90 kDa DNA polymerase containing a 3' →5' exonuclease activity, which functions by recognizing and removing incorrectly incorporated deoxynucleotides. The reaction compositions using ProofStart™ DNA polymerase were listed in table 2.5 (Qiagen).

Table 2.5 PCR components for ProofStart™ DNA polymerase

Component	Volume/reaction	Final concentration
Master mix		
10x PCR Buffer*	5 µl	1x
dNTP mix (10 mM of each)	1.5 µl	300 µM of each dNTP
Primer A	1 µl	1 µM
Primer B	1 µl	1 µM
ProofStart DNA Polymerase	2 µl	5 units
Distilled water	Variable	–

Template DNA	Variable	100 ng – 1µg genomic DNA 1–50 ng plasmid DNA
Total volume	50µl	

* *Contains 15 mM $MgCl_2$*

A typical PCR cycling program was outlined below (Qiagen).

Initial activation step:	5 min	95°C
3-step cycling		
Denaturation:	0.5–1 min	94°C
Annealing:	0.5–1 min	50–68°C
Extension:	2 min/kb	72°C
Number of cycles:	35–45	
Final extension:	10 min	72°C

Thermal cycling parameters used in PCR were variable depending on the primer pairs in use and the nature of template DNA, etc. Optimal reaction conditions such as incubation time, temperature, and amount of template DNA were individually determined.

Thermal cycling was performed in an Eppendorf Gradient Thermocycler. This PCR machine is fitted with heated lid to prevent evaporation of the reaction mixture during the cycling process. The lid was pre-heated to 105 °C before thermal cycling begins.

After amplification, the PCR sample contains a complex mixture of specific PCR product and residual reaction components such as primers, unincorporated nucleotides, enzyme, salts, and probably nonspecific amplification products. It is necessary to remove these contaminants before the specific PCR product can be used in subsequent

experiments. PCR products were separated by agarose gel eletrophoresis. Extraction and purification of PCR product from agarose gels was performed using a QIAquick Gel Extraction Kit (Qiagen).

2.5 Directional cloning into plasmid vectors

2.5.1 Cloning PCR products into pGEM-T Easy vector

The PCR amplified DNA fragment was ligated into the pGEM®-T Easy Vector (Promega), allowing the amplified gene sequence to be maintained in a stable form in a heterologous host (*E.coli*).

Taq DNA polymerases such as the HotStarTaq™ DNA polymerase often add a single deoxyadenosine to the 3′-ends of the amplified fragments, which is independent of template. The pGEM®-T Easy Vector possesses complementary 3′ terminal thymidine to both ends. This vector system is convenient for the cloning of PCR products.

The ligation reaction was set up according to the protocol of the kit, which is described below.

2X Rapid Ligation Buffer, T4 DNA Ligase	5µl
pGEM®-T Easy Vector (50ng)	1µl
PCR product	1-3µl
T4 DNA Ligase (3 Weiss units/µl)	1µl
deionized water to a final volume of	10µl

The reaction mixture was incubate at 16°C for 1-2 hours and then at 4°C overnight.

Thermostable DNA polymerases with proofreading activity, such as ProofStart DNA polymerase generate blunt-ended fragments during PCR amplification. PCR fragments generated using these polymerases were modified using the A-tailing procedure and ligated into the pGEM®-T Easy Vector, which were outlined below (Promega).

Start with 1–7 μl of purified PCR fragment generated by ProofStart DNA polymerase

↓

Add 1 ul Taq DNA Polymerase 10X Reaction Buffer with $MgCl_2$.

↓

Add dATP to a final concentration of 0.2mM.

↓

Add 5 units of Taq DNA Polymerase.

↓

Add deionized water to a final reaction volume of 10 μl.

↓

Incubate at 70°C for 15–30 minutes.

↓

Use 1–2 μl in a ligation reaction with pGEM®-T Easy Vector.

2.5.2 Cohesive end DNA cloning

The desired restriction sites were usually added to both termini of the target gene sequence by PCR using oligonucleotide primers. The gene sequence on pGEM®-T Easy Vector was cut out by two appropriate restriction enzymes that cleave different sequence and generate different cohesive termini. Then the gene was separated by agarose gel eletrophoresis. Extraction and purification of the gene sequence from agarose gels was performed using a QIAquick Gel Extraction Kit. The purified gene sequence was then ligated into an expression vector digested by the same restriction enzymes to express heterologous protein in *E.coli*. The ligation of two fragments was performed using T4 DNA ligase (Promega). The ligation reaction was set up as described below.

10X Ligation Buffer, T4 DNA Ligase	1.5µl
Vector DNA	1-6µl
Insert DNA	1-6µl
T4 DNA Ligase (10 Weiss units/µl)	1.5 µl
deionized water to a final volume of	15µl

The reaction mixture was incubate at 16°C for 1-2 hours and then at 4°C overnight.

2.6 Preparation and transformation of competent *E.coli*

2.6.1 Preparation of competent *E. coli*

Chemically competent *E.coli* cells were used throughout this work to produce strains for cloning and expression by transformation of recombinant plasmids into the cells. The competent cells NovaBlue, TOP10 and XL1-Blue were obtained from a commercial source, whilst other competent cells were produced in the laboratory. The procedure below was used to prepare competent cells (Qiagen).

A trace of desired *E. coli* cells was removed from the vial with an inoculating loop, and streaked out on LB agar containing appropriate antibiotics. The LB plate was incubated at 37°C overnight. A single colony was picked and used to inoculate 10 ml of LB-antibiotics and grown at 37°C, 220 rpm, overnight. 1 ml overnight culture was used to inoculate 100 ml prewarmed LB medium containing appropriate antibiotics in a 250 ml baffled conical flask. The cell suspension was shaken at 37°C, 220 rpm until it reached an OD_{600} of 0.5, then the culture was immediately transferred to a sterile, round-bottom centrifuge tube and kept on ice for 5 min. The cells were pelleted in a refrigerated centrifuge at low speed (5 min, 4000 x *g*, 4°C). The cell pellets were resuspended gently in 50 ml of ice-cold TFB1 buffer (100 mM RbCl, 50 mM $MnCl_2$,

30 mM potassium acetate, 10 mM $CaCl_2$, 15% glycerol, pH 5.8) and the suspension was kept on ice for an additional 90 min. Then the cells were pelleted in a refrigerated centrifuge at low speed (5 min, 4000 x *g*, 4°C). The cell pellets were resuspended gently in 30 ml if ice-cold TFB2 buffer (10 mM MOPS, 10 mM RbCl, 75 mM $CaCl_2$, 15% glycerol, adjust to pH 6.8 with KOH) and the suspension was kept on ice for a further 30 min. Then the cells were pelleted in a refrigerated centrifuge at low speed (5 min, 4000 x *g*, 4°C). The cell pellets were resuspended carefully in 4 ml ice-cold TFB2 buffer. The cell suspension was split into 100–200 µl aliquots in sterile microcentrifuge tubes and frozen in liquid nitrogen. The competent cells were stored at –80°C.

2.6.2 Transformation of competent *E.coli* cells

Recombinant plasmids were transformed into commercial competent cells according to the manufacturer's procedure. Competent cells produced in the laboratory by the previously described procedure were transformed according to the protocol below (Qiagen).

An aliquot of the ligation mixture was transferred into a ice-cold sterile 1.5 ml microcentrifuge tube. An aliquot of frozen competent *E.coli* cells was removed from storage at –80°C and was allowed to thaw on ice. The cells were gently resuspended and 100-150 µl of the cell suspension was transferred into the microcentrifuge tube with the ligation mixture. The mixture was incubated on ice for 20-30 min. After the incubation period the cells were subjected to heat-shock in a 42°C water bath for 90 sec, and then the cells were returned to ice for 2 min. 500 µl Psi broth (LB medium, 4 mM MgSO4, 10 mM KCl) were added to the cells and the cells were incubated for 60–90 min at 37°C, 220 rpm. Subsequently, 50, 100, and 200 µl aliquots were spread onto LB-agar plates containing appropriate antibiotics. The plates were incubated in an inverted position at 37°C overnight.

2.6.3 Screening transformants for inserts

Transformants were initially screened using a LB-agar plate containing appropriate antibiotic(s) on the basis of the vector-encoded antibiotic resistance determinants. After this initial screening, the procedures below were used to further identify the presence of the recombinant insert, and then plasmid was prepared to confirm the construct by DNA sequencing.

2.6.3.1 Screening bacterial colonies using X-gal and IPTG: α-complementation

Successful cloning of an insert in the pGEM®-T Easy Vectors disrupts the coding sequence of galactosidase. Cells bearing self-ligated pGEM®-T Easy Vector are capable to hydrolyse the artificial lactose analogue X-Gal, producing a blue pigment; whereas cells bearing recombinant plasmids can not do so and thus remain white in colour. The recombinant clones can be identified by this white/blue colour screening on indicator plates with IPTG and X-Gal (Novagen).

2.6.3.2 PCR screening

The PCR reaction composition for amplification of the inserted gene using the HotStarTaq™ DNA polymerase was prepared. A colony was picked from an agar plate using a 200 µl pipettete tip and mixed with the PCR reaction composition by gently aspirating and dispensing. Then the thermal cycle was processed for 30 cycles. The result of amplification was analyzed by agarose gel electrophoresis.

2.6.3.3 Restriction enzyme digestion screening

Transformed colonies were selected using a LB-agar plate containing appropriate antibiotic(s), and then plasmids were isolated from a single colony to screen for the presence of the inserted sequence by restriction digestion.

2.7 *In vitro* site-directed mutagenesis

The QuikChange® Site-Directed Mutagenesis Kits and procedure (Stratagene) were used to make point mutations in target genes.

2.7.1 Mutagenic primer design

The mutagenic oligonucleotide primers for use in this work were designed individually according to the desired mutation. The following considerations were made for designing mutagenic primers.

Both of the mutagenic primers contained the desired mutation and annealed to the same sequence on opposite strands of the plasmid. Primers were between 25 and 45 bases in length, with a melting temperature (Tm) of $\geq$78°C. The desired mutation was in the middle of the primer with ~10–15 bases of correct sequence on both sides. The primers had a minimum GC content of 40% and terminated in one or more C or G bases.

2.7.2 Mutant strand synthesis reaction (Thermal Cycling)

The plasmid DNA template used for mutant strand synthesis reactions was isolated from dam+ *E. coli* strains, such as NovaBlue and XL1-Blue. The sample reaction was prepared as indicated below:

5 µl of 10× reaction buffer

X µl (5–50 ng) of dsDNA template

X µl (125 ng) of oligonucleotide primer #1

X µl (125 ng) of oligonucleotide primer #2

1 µl of dNTP mix

ddH_2O to a final volume of 50 µl

Then add 1 µl of *PfuTurbo* DNA polymerase (2.5 U/µl)

Cycling parameters for the QuikChange Site-Directed Mutagenesis Method were outlined in table below.

Segment	Cycles	Temperature	Time
1	1	95°C	30 seconds
2	16	95°C	30 seconds
		55°C	1 minute
		68°C	1 minute/kb of plasmid length

2.7. 3 *DpnI* digestion of the amplification products

One µl of the *DpnI* restriction enzyme (10 U/µl, Stratagene) was added directly to each amplification reaction and the reaction mixture was mixed gently by pipetteting the solution up and down several times. The reaction mixtures were spun down in a

microcentrifuge for 1 minute and then immediately incubated at 37°C for 1 hour to digest the parental (i.e., the nonmutated) supercoiled dsDNA.

2.7. 4 Transformation of XL1-Blue supercompetent cells

The mutated plasmids were transformed into XL1-Blue supercompetent cells according to the manufacturer's procedure (Stratagene), which is outlined below.

The XL1-Blue supercompetent cells were gently thawed on ice and resuspended gently. 50 µl of the supercompetent cells were transferred to an ice-cold 1.5 ml micrcentrifugate tube. One µl of the *DpnI*-treated DNA from the reaction mixture was added to the micrcentrifugate tube with supercompetent cells. The transformation reaction was mixed carefully and incubated on ice for 30 minutes. The transformation reaction was subjected to heat-shock for 45 seconds at 42°C and then placed on ice for 2 minutes. 0.5 ml of NZY^+ broth preheated to 42°C was added to the transformation reaction and the cells were incubated at 37°C, 220 rpm for 1 hour. The appropriate volume of each transformation reaction was spread onto agar plates containing the appropriate antibiotic for the plasmid vector. The transformation plates were incubated at 37°C for over 16 hours.

2.8 Expression of cloned genes in *E.coli*

Several expression plasmid systems, such as pET, pBAD and pQE series plasmids, have been used to express cloned genes in different *E.coli* strains (Table 2.2). Expression constructs were transformed into an appropriate expression host. The general methods to overexpress recombinant proteins were used as summarized below. Different proteins have subtle differences in their expression and purification, which are indicated in the relevant chapter.

A single colony from a freshly streaked plate was picked and used to inoculate 10 ml 2×YT broth containing the appropriate antibiotic(s) in a 25 ml universal container. The cells were grown in an orbital shaker at 37 °C, 220 r.p.m. overnight. A liter of prewarmed 2×YT broth (with antibiotic) in a 2 liter baffled flask was inoculated with 2-4 ml of the overnight cultures. The cells were grown in an orbital shaker at 37 °C, 220 rpm, and their growth was monitored by measuring the absorbance of the culture at 600 nm (A_{600}). The cells were induced at mid log phase (approx. A_{600}=0.5-0.6) with 0.1-0.5 mM IPTG (for pET and pQE series vector) or 0.01%-0.05% L-arabinose (for pBAD series vector). The temperature was dropped to 25°C-30 °C and the growth was continued with shaking at 200 rpm for several hours or overnight. The cells were harvested by centrifugation and stored as a frozen pellet at -80 °C or purification continued.

2.9 Purification of hexahistidine-tagged proteins by immobilized metal affinity chromatography (IMAC)

2.9.1 Cell lysis and sub-cellular fractionation by differential centrifugation

Sub-cellular fractionation was adopted to identify the cellular location of particular proteins. It was also performed as an initial crude purification procedure.

Bacterial cells were harvested by centrifugation at 4,350 ×g (Beckman JA-10 rotor) for 8 min at 4 °C, resuspended in lysis buffer typically containing 100-300 mM NaCl, 10-20% glycerol and a buffering salt (e.g. Tris-HCl) at 20mM, and then stored at −80 °C or purification continued. Cell pellets were typically resuspended with 5-10 ml of ice-cold buffer per gram (wet weight) of cells. DNase I was added to the cell suspension to a final concentration of 5 U/ml and EDTA-free protease inhibitor

mixture tablet (Roche Molecular Biochemicals) was added at 1 tablet /50ml before lysis of the cells.

The cells were lysed by three passages through a Constant System Cell Disrupter (model Z-plus 1.1 kW, Constant System) at a pressure of 15-20 KPsi and the cell lysate was kept on ice to avoid proteolysis. Cellular debris was sedimented by centrifugation at 24,000 ×g in a Beckman JA-20 rotor for 30 min at 4 °C. The supernatant liquid was further separated into two fractions, the soluble cytosolic components of the cell and the membrane vesicles, by high-speed ultracentrifugation at 122,400 ×g (Beckman Ti-50 rotor) for 90 min at 4°C. The soluble cytosolic components of the cell were used as the starting material for protein purification if the target is a soluble protein.

The membrane pellet was emulsified in a small volume of buffer by repeated passage through a syringe needle and then the emulsified membrane was diluted by solubilization buffer to a final concentration of total protein of about 5-10 mg/ml. The integral membrane proteins were extracted from membrane lipids using detergent, such as DDM. The cellular membrane/detergent mixture was gently stirred at 4 °C for 1-3 hours and was then ultracentrifuged at 122,400 ×g for 60 min to sediment insoluble membrane debris. The membrane solution was processed according to the protocol for protein purification.

2.9.2 Batch purification of His_6-tagged proteins from *E. coli* using IMAC

One ml of the 50% Ni-NTA slurry was equilibrated with buffer and then added to 50 ml cleared lysate (or membrane solution) and mixed gently by shaking at 4°C for 2-3 hours. The lysate (or membrane solution)–Ni-NTA mixture was loaded into a column with the bottom outlet capped. The bottom cap was removed and the column flow-through was collected. The column was washed twice with 20 ml wash buffer containing imidazole (the concentration of imidazole in wash buffer was optimized according to individual protein), the wash fractions were collected for SDS-PAGE analysis. The protein was eluted 4 times with 0.5 ml elution buffer. The eluate was collected in four tubes and analyzed by SDS-PAGE (Qiagen).

2.9.3 Purification of 6xHis-tagged proteins from *E. coli* using HiTrap Chelating column

The HiTrap Chelating column and procedure for this column (Amersham Biosciences) were adopted to purify His_6-tagged proteins, which are outlined below.

2.9.3.1 Column preparation

The peristaltic pump tubing was filled with distilled water. The column stopper was removed and the column was connected to the pump tubing (with the provided adaptor), "drop to drop" to avoid introducing air into the system. The snap-off end at the column outlet was removed. The column was washed with 5 ml distilled water. 1 ml of 0.1 M $NiSO_4$ in distilled water was loaded on column. Then the column was washed with 5 ml distilled water.

2.9.3.2 Purification

The column was equilibrated with 5-10 column volumes of buffer at a flow rate of 0.8 ml/min. The sample was applied to the column using a peristaltic pump at a flow rate of 0.8 ml/min. The HiTrap Chelating column was then connected onto an *AKTA Purifier*. The column was washed using a step gradient with buffer containing 30-100 mM imidazole until the A_{280} was stable. Then the column was eluted with buffer containing 300-400 mM imidazole with 2-5 column volumes.

2.10 Purification of recombinant proteins by gel filtration chromatography

Gel filtration chromatography was used to polish the proteins initially purified by IMAC. The HiLoad 16/60 Superdex 200 prep grade column (Amersham Biosciences) was used in this study. The column was run according to the specification provided (Amersham Biosciences), which is outlined below.

Before connecting the column to a chromatography system, the pump was started to remove all air from the system, particularly in tubing and valves. The column was mounted vertically, the domed nut was removed, and the inlet tubing was connected to the system "drop-to-drop". The transport syringe was removed and the column outlet tubing was connected to a monitor cell. The column was equilibrated with three column volume of dH_2O at 30 cm/h (1 ml/min), and three column volumes of buffer at 30 cm/h (1 ml/min). Initially the following conditions were tried:

Flow rate: 30 cm/h (1 ml/min).

Sample volume: 1% of the column volume (1.2 ml).

Buffer: The buffer type and pH were compatible with protein stability and activity, and buffer concentration was sufficient to maintain buffering capacity and

constant pH. Ionic strength in the buffer was equivalent to at least 0.15 M NaCl, to avoid nonspecific ionic interactions with the matrix.

If the results obtained were unsatisfactory, flow rate and (or) sample volume were decreased to improve resolution.

2.11 Purification of recombinant proteins by ion exchange chromatography

Ion exchange chromatography was also used to polish the proteins purified by IMAC. A RESOURCE Q 6 ml column (Amersham Biosciences) was used in this study. The column was run according to the specification provided (Amersham Biosciences), which is outlined below.

2.11.1 Equilibrating the column

The RESOURCE Q column is stored in 20% ethanol. The steps below were followed when equilibrating the column.

The column was flushed with two to five column volumes of dH_2O. This washes out most of the storage solution. Two column volumes of elution buffer were run through the column. Then the column was equilibrated with five column volumes of start buffer.

2.11.2 Applying and eluting the sample

The protein binding capacity for the RESOURCE Q 6 ml column is up to 150 mg. The practical protein loading range may vary depending on the proteins to be separated and on the running conditions and the detection wavelength used. There is

no limit for the sample volume loaded on the column when a gradient elution is used, however, it could affect the resolution of early eluting components.

2.11.3 Optimizing the gradient shape

The shape of the elution gradient was optimized to give the best separation. These conditions were tried first: starting with a linear gradient of 0–100% elution buffer in a volume of 120 ml. If the results obtained were unsatisfactory, the resolution was improved using lower flow rates and more shallow gradients. If the protein of interest was not eluted in the gradient, the elution buffer was changed to one with a higher salt concentration.

2.12 Analysis of recombinant expression

2.12.1 SDS-PAGE

NuPAGE® LDS Sample Buffer (4X)
106 mM Tris HCl
141 mM Tris base
2% LDS
10% Glycerol
0.51 mM EDTA
0.22 mM SERVA® Blue G250
0.175 mM Phenol Red
pH 8.5

NuPAGE® MOPS SDS
Running Buffer (20X)
50 mM MOPS
50 mM Tris base
0.1% SDS
1 mM EDTA
pH 7.7

The recombinant protein expression and the efficacy of protein purification procedures were routinely monitored by SDS-PAGE. All SDS-PAGE experiments were performed using the NuPAGE® Novex Pre-Cast Gels (4-12% gradient gels for

samples containing proteins of widely variant molecular weights, or 12% gels for less complex samples) and XCell *SureLock* Mini-Cell (Invitrogen). The NuPAGE® LDS Sample Buffer (4X) (Invitrogen) was used for preparing samples for denaturing gel electrophoresis with the NuPAGE® Gels. The procedure for electrophoresis of NuPAGE® Gels is outlined below (Invitrogen).

2.12.1.1 Electrophoresis of NuPAGE® gels

The NuPAGE® gel was vertically mounted. The comb was gently pulled out of the cassette and the sample wells were rinsed with 1X NuPAGE® SDS Running Buffer. 5-30 µl of protein samples were mixed with 10 µl NuPAGE® LDS Sample Buffer (4X) and an appropriate volume of deionized water to give a final volume of 40 µl. 10-40 µl of the mixture were loaded into each well of the gel. 20 µl of protein molecular weight markers (SeeBlue Markers, Invitrogen) were loaded in one well. The gel cassette was then clamped vertically in the Mini-Cell such that the notched well side of the cassette faces inwards toward the Buffer Core. The Upper and Lower Buffer Chamber were filled with 200 ml and 600 ml of running buffer, respectively. Then the gel was run at 200 V constantly for 50 minutes. After completion of the electrophoresis procedure the gel was removed gently from the cassette using a Gel Knife and rinsed for 10 minutes in de-ionised water to remove the running buffer.

2.12.1.2 Coomassie® R-250 microwave staining of SDS-PAGE

Staining solution:
0.1% Coomassie® R-250 in 40% ethanol,
10% acetic acid

Destain solution:
10% ethanol and 7.5% acetic acid.

The Coomassie® R-250 microwave staining procedure was used to stain SDS-PAGE gel (Invitrogen). One gel was incubated in a staining container containing 100 ml of staining solution and heated in a microwave oven at full power for 1 minute. Then the

gel was gently shaken for 15 minutes at room temperature on an orbital shaker in the fume cupboard. The stain solution was decanted and the gel was rinsed once with deionized water. Then the gel was soaked in 100 ml of destain solution and heated in a microwave oven at full power for 1 minute. Subsequently, the gel was gently shaken at room temperature on an orbital shaker in the fume cupboard until the desired background was achieved.

2.12.2 Western blotting

ELISA wash buffer:

Phosphate stock	30 ml
NaCl	4.38 g
Tween 20	250 ul
ddH_2O to	500 ml

Phosphate stock:

KH_2PO_4	2.586 g
K_2HPO_4	14.12 g
ddH_2O to	400 ml

Transfer Buffer:

Glycine	2.9 g
Tris Base	5.8 g
SDS	0.37 g
Methanol	200 ml
ddH_2O to	1000 ml

A section of PVDF (0.2 um, Bio-Rad) was cut to fit the size of the gel. The top right hand corner was marked with a biro. PVDF was placed in methanol for approximately 10 seconds and left to soak in the methanol-transfer buffer for 15 minutes or longer. Two pieces of Whatman Chromatography paper (3 MM Chr) were cut slightly larger then the size of the gel. The protein gel was run using SeeBlue Markers (Invitrogen). The gel was soaked in the methanol-transfer buffer for 15-30 minutes. The blotting apparatus was set up as follows in a tray with buffer (Figure 2.1). The cage was rapidly transferred to the blotting chamber which was about half full with buffer. The

blotting chamber was topped up with transfer buffer and run at 170 mA for 1.5 hours. Then the blotted gel was stained to check for successful transfer. The PVDF membrane was transferred to 25 ml Blocking Buffer (ELISA wash buffer plus 3% BSA) in a square Petri dish; the membrane was blocked for 90 minutes at room temperature with gentle agitation, and then was held at a 4 °C overnight. The membrane was washed three times with ELISA wash buffer at room temperature, five minutes each time. The membrane was incubated with the first antibody (1:3000 dilutions) in 25 ml fresh blocking buffer (ELISA wash buffer plus 0.5% BSA and 8 µl monoclonal antipolyhistidine antibody [Sigma]) at room temperature for 1 hour. The membrane was rinsed in ELISA wash buffer three times, for 10 minutes each time, and incubated with the second antibody (1:3000 dilutions) in 25 ml fresh blocking buffer (ELISA wash buffer plus 0.5% BSA and 8 µl goat anti-mouse IgG-AP conjugate, [Sigma]) at room temperature for 1 hour. The membrane was rinsed in ELISA wash buffer three times, for 10 minutes each time. The membrane was transferred to a sheet of clear film and 2 ml substrate was added. The membrane was wrapped with the film, and left on a bench for 5 minutes. The membrane was transferred to between two sheets of plastic and taped around the four sides in a developing case, and developed in the dark for 2 minutes.

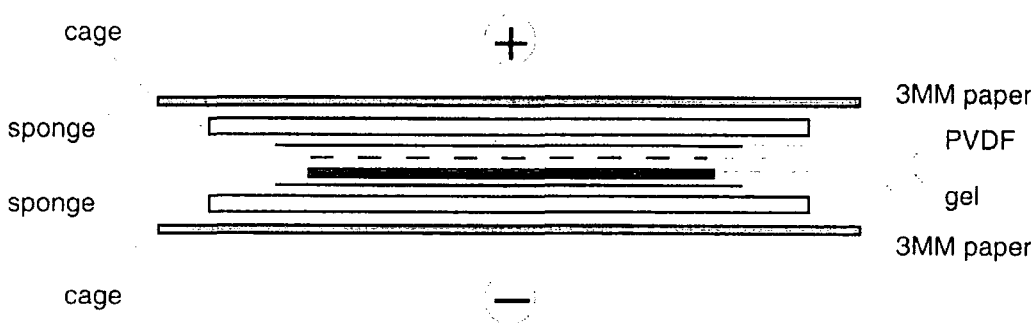


Figure 2.1 Schematic representation for setting up the blotting apparatus

2.13 Protein buffer change and concentration

HiTrap Desalting 5 ml column (Amersham Biosciences) and Slide-A-Lyzer® 10K MWCO Dialysis Cassettes (Pierce) were used to remove low molecular weight contaminants and perform buffer exchange.

2.13.1 Buffer changes using a HiTrap Desalting 5 ml column

The HiTrap Desalting column is packed with the size exclusion medium Sephadex G-25 Superfine, which has a fractionation range for globular proteins between Mr 1 000–5 000 and an exclusion limit of approximately Mr 5 000. This feature ensures group separations of proteins/peptides larger then Mr 5 000 from molecules with a molecular weight less than Mr 1 000. The column was run according to the specification provided (Amersham Biosciences), which is outlined below.

The desalting column was collected onto an *AKTA Purifier*. The column was washed with 25 ml of deionized water at 4 ml/min to completely remove the ethanol and then equilibrated with 25 ml buffer. The column was disconnected from the *AKTA Purifier* and then connected "drop to drop" to a syringe via the adaptor and 1.5 ml of the protein sample was applied to column; the eluted buffer from the column was discarded. Then 2 ml of buffer were applied to column and 2 ml eluates were collected.

2.13.2 Buffer changes using Slide-A-Lyzer® 10 kDa MWCO dialysis cassettes

The procedure for the Slide-A-Lyzer® 10 kDa MWCO dialysis cassettes (Pierce), used for buffer change, is summarized below.

A Slide-A-Lyzer® dialysis cassette was carefully loaded with protein sample using a syringe and then slipped into the groove of the buoy. This assembly was floated in the dialysis buffer of choice. The dialysis buffer was used at 200-500 times the volume of the sample. The sample was dialyzed for 2 hours at 4 °C with continuous stirring of the buffer to promote rapid equilibration; then the dialysis buffer was changed and the sample was dialyzed for another 2 hours at 4 °C; the dialysis buffer was changed again and the sample was dialyzed overnight at 4°C. After dialysis, the sample was carefully removed with a syringe from the dialysis cassette.

2.13.3 Protein concentration

The Vivaspin Concentrators are disposable ultrafiltration devices for the concentration of biological samples (Viva Science).

The Vivaspin concentrators with different molecular weight cut-off points were used to concentrate proteins. Firstly, several volumes of deionised water were passed through the concentrator by centrifugation using a fixed angle rotor to remove trace amounts of glycerine and sodium azide on the membrane. The concentrator was centrifuged at speed recommended for its membrane type and MWCO. Then the concentrator was filled with protein sample and inserted into the rotor with the printed window of the concentrator facing outwards. Once the desired concentration was achieved, the concentrator was removed and sample was recovered from the bottom of the concentrate pocket with a pipettete.

2.13.4 Protein concentration determination

The BCA™ Protein Assay Kit (Pierce) was used to determine protein concentration. This detergent-compatible method combines the reduction of Cu^{+2} to Cu^{+1} by protein in an alkaline medium (the biuret reaction) with colorimetric detection of the cuprous cation (Cu^{+1}) using a unique reagent containing bicinchoninic acid (BCA). The chelation of two molecules of BCA with one cuprous ion produces a water-soluble complex exhibiting a strong absorbance at 562 nm, which is nearly linear with increasing protein concentrations over a broad working range (20-2,000 μg/ml). The 'Test Tube Procedure' for this Kit (Pierce) was adopted, which is summarized as follows.

0.1 ml of each standard and unknown protein sample were transferred into a test tube and mixed with 2ml of the BCA working reagent. The tubes were then incubated at 37°C for 30 minutes to allow colour development. After the incubation period, reactions were cooled to room temperature and the absorbance of all the samples at 562 nm were measured against a dH_2O blank. The average 562 nm absorbance measurement of the blank standard replicates was subtracted from that of all other individual standard and unknown sample replicates. A standard curve was prepared by plotting the average blank-corrected 562 nm measurement for each BSA standard vs. its concentration in μg/ml. Then the standard curve was used to determine the protein concentration of each unknown sample.

2.14 Protein molecular weight and Stokes' radius determination by analytical gel filtration chromatography

Analytical gel filtration is an accurate method for determination of the molecular weight and Stokes' radius of native proteins under a wide variety of conditions of pH, ionic strength and temperature. The basic principle lies in that there is a sigmodial

relationship between the various elution volume parameters of protein molecules and the logarithm of their molecular weights. A standard calibration curve can be prepared by measuring the elution volumes of several standard protein molecules, calculating their corresponding Kav values, and plotting the Kav values versus the logarithm of their molecular weights. The molecular weight of an unknown protein molecule can be determined from the calibration curve once its Kav value is calculated from its measured elution volume. However, if the unknown protein molecule does not have the same molecular shape as the standards, the relationship between its molecular size (elution volume) and molecular weight may be different from that of the standards, and thus its molecular weight may not correlate well to the calibration curves established by the standards. For such a protein molecule, the molecular size parameters, such as Stokes' radius, may provide useful information. A calibration curve can be prepared by plotting $[-\log (kav)]^{1/2}$ versa Stokes' radius to determine the Stokes' radius of proteins (Amersham Biosciences).

The analytical gel filtration chromatography column Superdex 200 PC 3.2/30 column and Low Molecular Weight Gel Filtration Kit (Amersham Biosciences) were used to determine the molecular weight and oligomerization state of proteins in the native state. The protocol used is described below (Amersham Biosciences).

A fresh, filtered solution of Blue Dextran 2000 (1 mg/ml) was prepared in the running buffer. Thirty µl of Blue Dextran 2000 was applied to the column equilibrated with running buffer. The column was eluted at a flow rate of 20 µl/min to determine its void volume (Vo). The protein standards of Ribonuclease A (Mr 13.7 kDa, Å 16.4), Chymotrypsinogen A (Mr 25 kDa, Å 20.9), Ovalbumin (Mr 43 kDa, Å 30.5) and Albumin (Mr 67 kDa, Å 35.3) were dissolved in running buffer at concentrations of 10mg/ml, 3mg/ml, 7mg/ml and 7mg/ml, respectively; a few minutes were allowed for dissolution with gently stirring. The Ribonuclease A and Ovalbumin, Chymotrypsinogen A and Albumin were mixed separately at the ratio of 1:1. Thirty µl mixtures of Ribonuclease A and Ovalbumin were applied to the column and the

column was eluted at a flow rate of 20 µl/ min. The elution volumes (Ve) of Ribonuclease A and Ovalbumin were determined by measuring the volume of the eluent from the point of application to the centre of the elution peak; and then thirty µl mixtures of Chymotrypsinogen A and Albumin were applied to the column to determine their elution volumes. The Kav values for the protein standards were calculated as follows: Kav = (Ve-Vo)/ (Vt-Vo), where Ve = elution volume for the protein, Vo = column void volume =elution volume for Blue Dextran 2000, Vt = total bed volume. The Microsoft Excel program was used to prepare a calibration curve of Kav versus the logarithm of their molecular weights and calculate the linear regression equation of the curve. Thirty µl of protein sample were applied to the column to determine its elution volume. The corresponding Kav of protein was calculated and its molecular weight was determined from the calibration curve. The same data were used to prepare a standard curve of $[-\log (kav)]^{1/2}$ versa Stokes' radius and calculate the linear regression equation of the curve.

2.15 Crystal growth techniques

The protocols from Hampton Research were adopted for crystal growth, which are outline below:

The hanging drop and sitting drop vapour diffusion technique are two popular methods for the crystallization of macromolecules. In these techniques, a small droplet composed of protein and crystallizing solution (usually buffer, salt and precipitant) is placed in vapour equilibration with a liquid reservoir containing 0.5-1.0 ml of crystallizing solution. Typically the drop contains a lower crystallizing reagent concentration than the reservoir, which drives the system toward equilibrium by diffusion through the vapour phase. Both the protein and crystallizing reagent increase in concentration as water leaves the drop for the reservoir. The protein sample undergoes an increase in relative supersaturation and crystals start to form when the

drop and reservoir are at or close to equilibrium. Vapour diffusion drops can be prepared in two ways described below.

2.15.1 Hanging drop vapour diffusion crystallization

The Crystallization Tool (Nextal Biotechnologies) was adopted to set up crystallization trails using the hanging drop vapour diffusion method. The Crystallization Tool includes a 24 well plate with 24 Crystallization Supports. The greaseless sealing technology of the Crystallization Tool offers a fast and clean way to perform crystallization trials, as well as a safe environment for the crystal recovery.

One ml of crystallization solution was placed with a pipettete into reservoir A1 of the Crystallization Tool. One drop of 1-4 μl protein sample was placed into the centre of a Siliconized Crystallization Support. The droplet was made as spherical as possible. Then 1-4 μl of crystallization solution from reservoir A1 was place into the drop on the Crystallization Support containing the protein sample. The pipettete tip was left in the drop while gently aspirating and dispensing the drop with the pipettete. The Crystallization Support was inverted carefully yet without delay so the drop is hanging from the Crystallization Support, then the Crystallization Support was gently screwed into the well of Crystallization Tool and tightened to ensure a complete seal. This operation was repeated for reservoir 2 through 24.

2.15.2 Sitting drop vapour diffusion crystallization using Micro-Bridge

The pregreased XRL Plate and Micro-Bridge (Molecular Dimensions) were adopted to set up crystallization trails using the setting drop vapour diffusion method. The Micro-Bridge is a small bridge (inverted U) containing a smooth, concave depression in the centre of the top region of the bridge. The Micro-Bridge is inserted into the reservoirs of the XRL Plate to perform a sitting drop vapour diffusion experiment.

One ml of crystallization solution was placed with a pipettete into reservoir A1 of the XRL Plate. A clean Micro-Bridge was placed into the bottom of reservoir A1 such that the concave depression in the Micro-Bridge was facing up. One drop of 1-4 μl protein sample was placed into the Micro-Bridge in reservoir A1. Then 1-4 μl of crystallization solution from reservoir A1 was placed into the drop in the Micro-Bridge A1. The pipettete tip was left in the drop while gently aspirating and dispensing the drop with the pipettete. The cover slide was positioned onto the bead of grease on reservoir A1, gently pressed down onto the grease and twisted 45° to ensure a complete seal. This operation was repeated for the remaining 23 reservoirs.

The Cooled Crystallization Incubators (RUMED) were used for crystal growth and the dissecting stereomicroscopy (Carl Zeiss Ltd) was used to observe the crystallization drops. The crystallization drops were checked at 24h, the third day and the first week after setting up, and followed by weekly thereafter.

Chapter 3

The crystal structure of the outer membrane protein VceC from *Vibrio cholerae*

3.1 Introduction to the tripartite multidrug efflux pump VceCAB

Isolates of the Gram-negative pathogen *Vibrio cholerae* have been described that are resistant to antibiotics (e.g. ampicillin, penicillin, streptomycin, nitrofurantoin and erythromycin) and toxic metals (e.g. Pb^{2+} and Zn^{2+}) (Choudhury *et al.*, 1996). Accordingly, multidrug efflux pumps are likely to be important in drug resistance of this cholera pathogen (Baranwal *et al.*, 2002). Several multidrug efflux pumps from *Vibrio cholerea* have been characterized, such as the MF family transporter VceCAB (Woolley *et al.*, 2005), the RND family transporter AcrAB (Chatterjee *et al.*, 2004), the ABC family transporter VcaM (Huda *et al.*, 2003), the SMR family transporter QacEΔ1 (Kazama *et al.*, 1999) and the MATE family transporters VcmB, VcmD, VcmH and VcmN (Begum *et al.*, 2005).

The tripartite pump VceCAB of *Vibrio cholerae* is composed of the inner membrane antiporter VceB, the membrane fusion protein VceA and the outer membrane channel VceC (Woolley *et al.*, 2005). The *vceA* and *vceB* genes were initially identified by Colmer *et al.* (Colmer *et al.*, 1998), which share homology with the *emrAB* genes that encode a multidrug efflux pump in *E.coli* (Lomovskaya *et al.*, 1992). The DNA sequences of both chromosomes of *Vibrio cholerae* have been determined (Heidelberg *et al.*, 2000). Examination of the *Vibrio cholerae* DNA sequence revealed four adjacent open reading frames (ORF): The ORF VC1409 (accession no.: NP_231052), VC1410 (accession no.: NP_231053) and VC1411 (accession no.: NP_231054) correspond to *vceC*, *vceA* and *vceB*, respectively; and the ORF VC1408 (accession no.: NP_231051), termed *vceR*, was suggested to be a transcriptional regulator of the *vceRCAB* operon. The *vceRCAB* operon occupies a total of 4909 bp of the genome (*vceR*=603 bp, *vceC*=1455 bp, *vceA*=1221 bp and *vceB*=1536 bp). The *vceA* and *vceB*

genes are separated by 9 bp, while the 3' terminus of *vceC* and 5' terminus of *vceA* overlap by 20 bp. The *vceR* gene is transcribed divergently in relation to the *vceCAB* operon. A 28 bp inverted repeat in the *vceR-vceCAB* intergenic region was identified to be the promoter area of the *vceCAB* operon. Figure 3.1 shows a schematic representation of the *vceRCAB* operon.

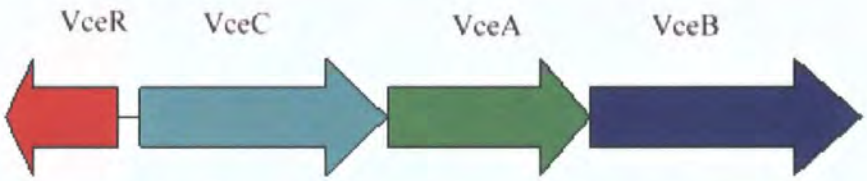


Figure 3.1 Schematic representation of the *vceRCAB* operon

Although the function of the VceCAB pump in *V. cholerae* has not been demonstrated, its components were found to collectively complement the multidrug resistance phenotype in *E. coli* null mutants of *tolC*, *acrB* and *emrB* (Colmer J et al., 1998). Among the compounds to which the VceCAB pump confers resistance are the uncoupler cyanide *m*-chlorphenylhydrazone (CCCP), the detergent sodium deoxycholate, phenylmercuric acetate (PMA) and several antibiotics such as chloramphenicol, nalixidic acid, erythromycin and rifampicin.

To investigate the role of the VceCAB pump in the multidrug resistance of *Vibrio cholerae*, I have characterized functionally and structurally the members of the VceCAB pump. In this chapter, I report the overexpression, purification and crystallization of the outer membrane component VceC.

3.2 PCR amplification of *vceA*, *vceB*, *vceC* and *vceR* genes from *V. cholerae* genomic DNA

The published sequences of *vceC*, *vceA*, *vceB* and *vceR* genes were used to design PCR primers for the specific amplification of the genes from genomic DNA. The *V.*

cholerae vaccine strain CVD101 (Hase *et al.,* 1994) was used as the source of genomic DNA for PCR. Figure 3.2 shows the successful amplification of the complete region of the operon and four individual genes, with the size and nucleotide sequences consistent with those published data for *vceCAB* and *vceR*.

The desired restriction sites were added to both termini of the target genes by PCR using oligonucleotide primers in order to subclone the amplified sequences from the initial cloning vector pGEM-T Easy into an expression vector. The amplified sequences were extracted from an agarose gel and ligated into the T-A cloning vector pGEM-T Easy. The ligation reaction was then transformed into competent *E.coli* NovaBlue cells to propagate the target genes.

3.3 Overexpression and purification of VceC

3.3.1 The pET-*VceC* expression construct

The pET21a vector was obtained from a commercial source (Novagen) and used to generate a recombinant expression construct for VceC. The *vceC* gene was amplified by PCR from genomic DNA of the *Vibrio cholerae* vaccine strain CVD101 using primers VceCF1/VceCR and ligated into pGEM-T Easy vector (Promega). After restriction enzyme digestion of the vector with *NdeI* and *XhoI*, the resulting *vceC* fragment was ligated into pET21a (Novagen), generating a 1479 bp open reading frame in which the native *vceC* gene (lacking the C-terminal termination codon) was fused to a short C-terminal sequence encoding a hexahistidine tag. The N-terminal ATG start code of *vceC* gene is located downstream of a bacteriophage T7 promoter, a *lac* operator sequence and an optimized *E.coli* ribosome binding site. Figure 3.3 shows the restriction digestion analysis of pET-*vceC*.

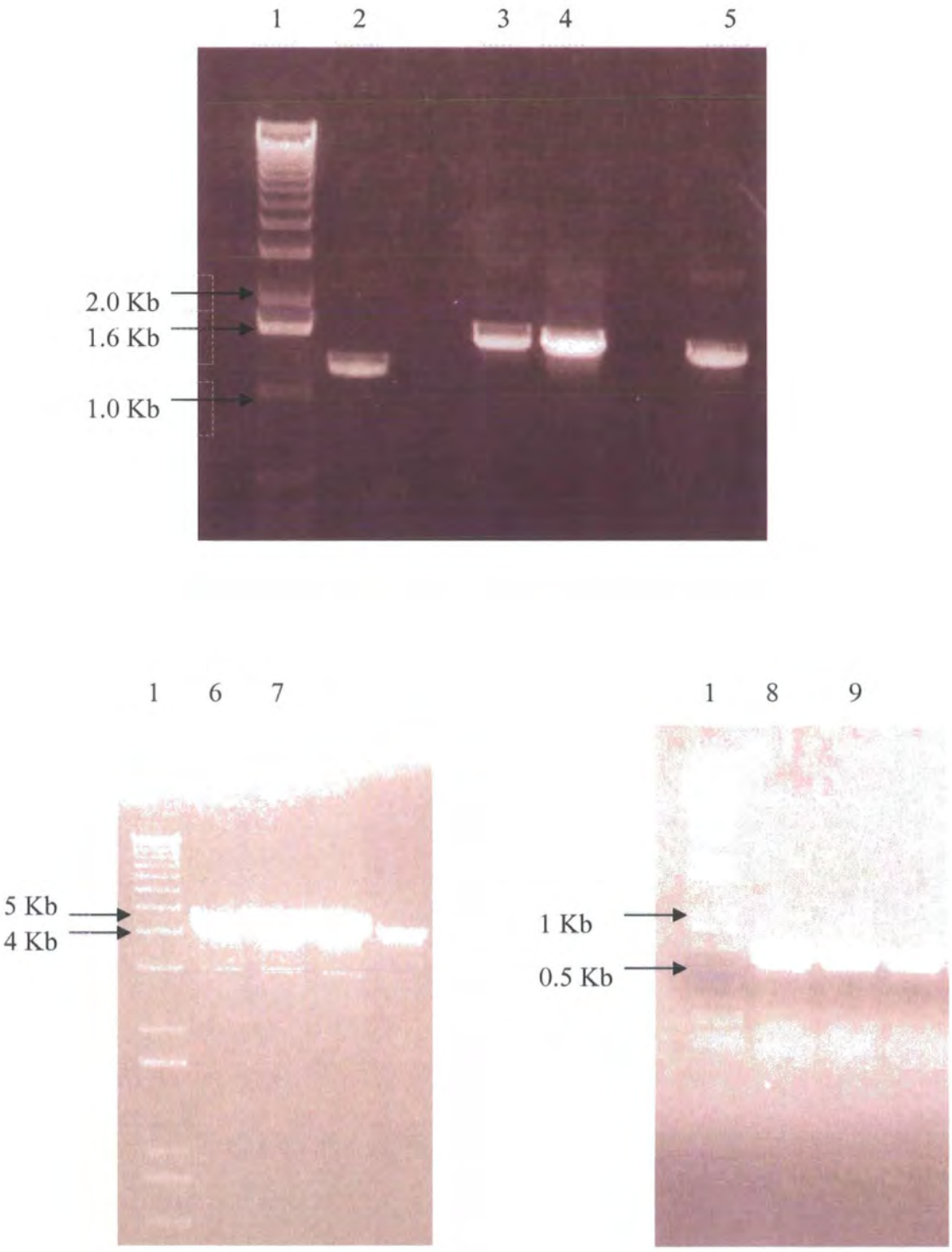


Figure 3.2. PCR amplification of *vceR, vceA, vceB* and *vceC* genes. Lane 1: 1 Kb Plus DNA Ladder (Invitrogen); Lane 2: *vceA* amplified using primers VceAF/VceAR2; Lane 3-4: *vceB* amplified using primers VceBF/VceBR2; Lane 5: *vceC* amplified using primers VceCF1/VceCR; Lane 6-7: *vceCAB* amplified using primers VceCF2/VceBR1; Lane 8-9: *vceR* amplified using primers VceRF/VceRR.

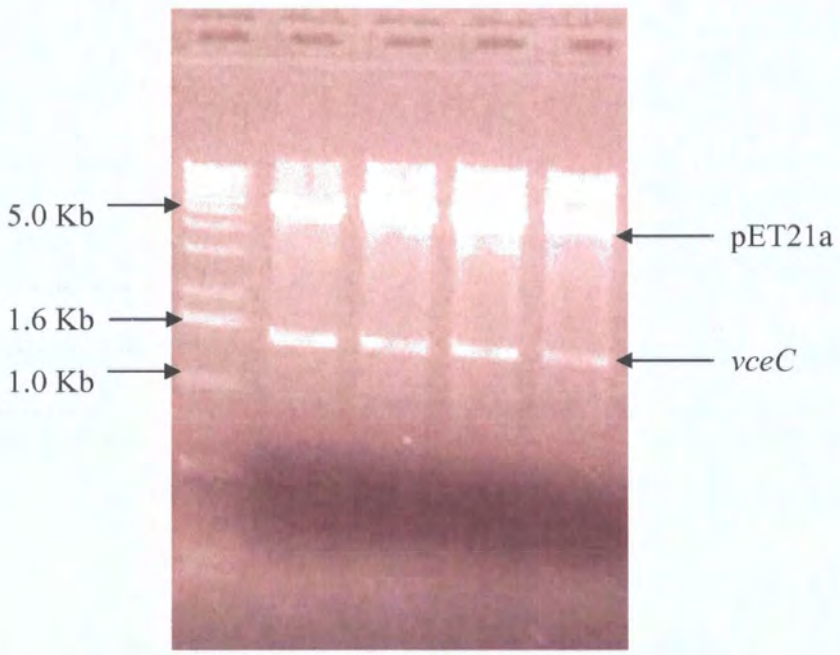


Figure 3.3. Restriction analysis of pET-*vceC* plasmids. The *NdeI/XhoI* bounded *vceC* gene was ligated into *NdeI/XhoI* digested pET21a vector and transformed into NovaBlue competent cells. The potential pET21a recombinants were analyzed by PCR first, then recombinant plasmids were prepared from PCR positive colonies and analyzed using *NdeI* and *XhoI* restriction digestion (37 ^{o}C for 3 hours). All recombinant plasmids exhibit *vceC* inserts.

3.3.2 Recombinant expression of VceC

The pET-*vceC* construct was transformed into *E.coli* strain BL21-AI (Invitrogen). Transformed colonies were selected on the basis of vector-encoded ampicillin resistance, and isolated plasmids were screened for the presence of the inserted *vceC* sequence by restriction digestion.

A single BL21-AI/pET-*vceC* colony from a freshly streaked plate was picked and used to inoculate 10 ml 2×YT broth containing carbenicillin (100μg/ml) in a 25 ml universal container. The cells were grown in an orbital shaker at 37°C, 220 rpm overnight. One liter of prewarmed 2×YT broth (with antibiotic) in 2 liter baffled flask was inoculated with 2 ml of the overnight culture. The cells were grown in an orbital shaker at 37 °C, 220 rpm, and their growth was monitored by measuring the absorbance of the culture at 600 nm (A_{600}).

The cells were induced at mid log phase (approx. A_{600}=0.5-0.6) with 0.2 mM isopropyl-β-D-thiogalactopyranoside (IPTG) and 0.02% L-arabinose. Then the temperature was dropped to 25 °C and the growth was continued with shaking at 200 rpm overnight. The cells were harvested by centrifugation at 4,350 ×g for 8 min at 4 °C, resuspended in buffer A (20 mM Na_2HPO_4 pH: 7.4, 300 mM NaCl, 20% Glycerol, 5 U/ml DNaseI, 1 tablet /50ml EDTA-free protease inhibitor mixture tablet) and stored at −80 °C or purification continued.

3.3.3 Purification of the His_6-tagged VceC using Ni^{2+} affinity chromatography

Purification of VceC-His_6 by Ni^{2+} affinity chromatography was performed using a pre-pack IMAC matrix, the HiTrap Chelating HP 1 ml column, purchased from Amersham Pharmacia Biotech.

The cells were lysed by three passages through a Constant System Cell Disrupter (15

Kpsi, model Z-plus 1.1 kW, Constant System). Cellular debris was removed by centrifugation at 24,000 ×g for 30 min at 4 °C. The supernatant liquid, containing soluble proteins and membrane vesicles, was carefully decanted so as not to disturb the pelleted material. The supernatant was further fractionated into membrane and soluble components by high-speed ultracentrifugation at 122,400 ×g for 90 min at 4°C. This procedure resulted in the generation of a supernatant containing cytosolic proteins and small molecules and a pellet of membranous material.

The membrane pellet was emulsified in a small volume of buffer B (20 mM Na_2HPO_4 pH: 7.4, 300 mM NaCl, 10% Glycerol, 0.5 mM THP) by repeated passage through a syringe needle, and then the emulsified membrane was diluted by buffer B to a final concentration of total protein of about 5-10 mg/ml. A powder of the detergent *n*-Dodecyl-ß-D-maltoside (DDM) was added to the emulsified membrane to a final concentration of 2% to solubilize the membrane. The mixture was gently stirred at 4 °C for 1.5 hours and then ultracentrifuged at 122,400 ×g for 1 hour at 4 °C. The membrane solution was carefully decanted so as not to disturb the pelleted material. NaCl was added to the membrane solution to a final concentration of 500 mM, and imidazole (6 M, pH: 7.4) to a concentration of 15 mM.

Extracted hexahistidine-tagged VceC was purified using the HiTrap chelating HP 1 ml column immobilized with Ni^{2+} equilibrated with buffer C (20 mM Na_2HPO_4, pH: 7.4, 500 mM NaCl, 10% Glycerol, 0.5 mM THP, 0.2% DDM). The membrane solution was loaded onto the HiTrap chelating HP 1 ml column using a peristaltic pump at a speed of 0.8 ml/min. The column was connected to the *AKTA Purifier* (Amersham Pharmacia Biotech) FPLC system. The column was washed with 100 mM imidazole in buffer C until the baseline of UV-absorbance at 280 nm (A_{280}) had no further change, and then VceC was eluted from the column with 500 mM imidazole, 0.2% DDM in buffer B. The eluates were collected in 0.5 ml fractions and analyzed by SDS-PAGE.

The full-length VceC is composed of 485 amino acids and the predicted molecular weight is about 52.2 kDa; the molecular weight of mature VceC should be lower with its signal peptide sequence being removed. Figure 3.4 illustrates an SDS-PAGE analysis of the eluates. The main protein band on SDS-PAGE is about 90% of total proteins. The molecular weight of this protein is approximately 50 kDa, which is within the size range of that expected for recombinant VceC-His_6.

3.3.4 Western blot identification of VceC-His_6

The Western blotting method was employed to confirm that the candidate 50 kDa protein was indeed VceC-His_6. The same sample shown in figure 3.4 was separated by SDS-PAGE, transferred to a polyvinglidene difluoride (PVDF) membrane by electroblotting, and then immunoblotted with a mouse monocloned primary antibody raised against the polyhistidine tag epitope. The blot was developed using an enzyme conjugated secondary antibody in conjunction with a chemiluminescent substrate (see chapter II). The result of Western blotting is shown in figure 3.5, which indicates that the candidate 50 kDa protein reacts strongly with the anti-polyhistidine tag antibody with high specificity, suggesting that this protein is VceC-His_6.

N-terminal sequencing of the protein further confirmed that the candidate protein is VceC-His_6, and revealed that the first amino acid corresponds to residue 42 of the cloned gene.

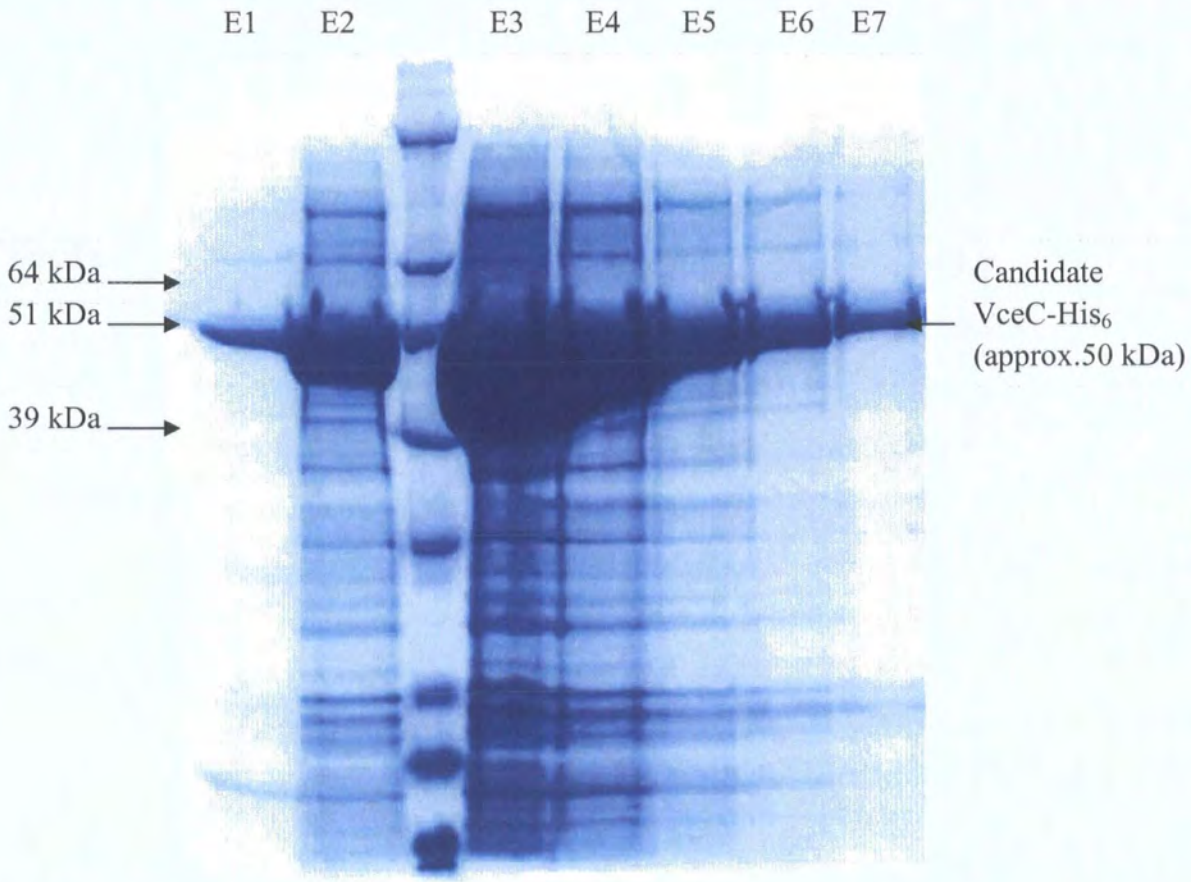


Figure 3.4. Coomassie blue stained SDS-PAGE gel showing partial IMAC purification of candidate VceC-His_6 protein from DDM solubilised membranes. Lane E1-E7 show elution fractions (0.5×7 ml) with buffer containing 500 mM imidazole. The left arrows mark the positions of protein standards; the right arrow marks the position of candidate VceC-His_6 protein.

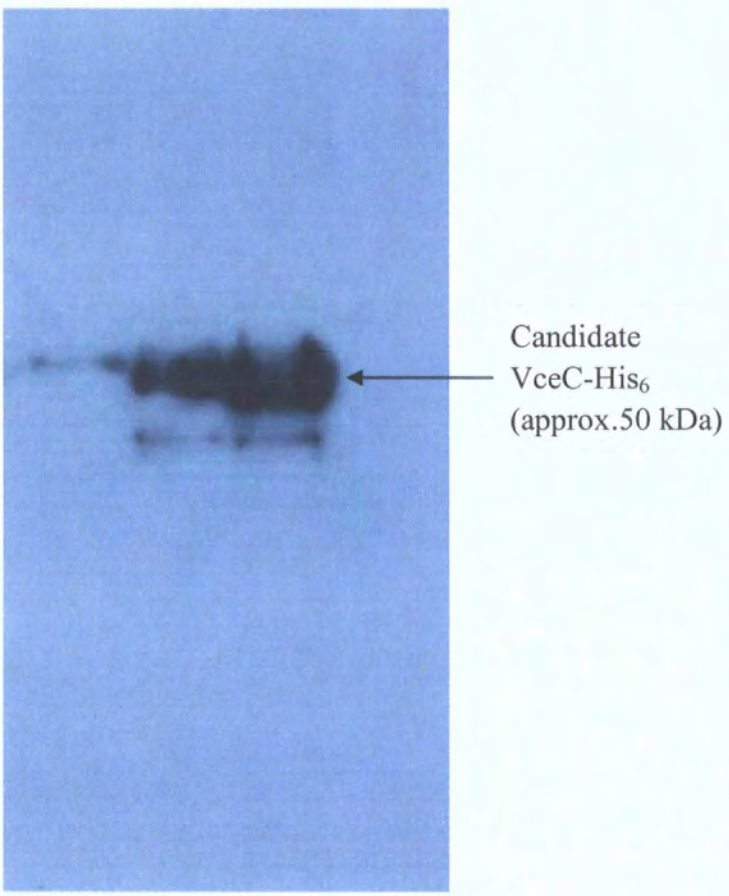


Figure 3.5. The Western blot above illustrates the strong reaction between an anti-His_6 antibody and an approximately 50 kDa protein with high specificity, suggesting that this protein is VceC-His_6.

3.3.5 Purification of VceC-His_6 by anion exchange chromatography

The purity of VceC protein purified by the Ni^{2+} affinity chromatography step is about 90% of total proteins, as seen on the SDS-PAGE analysis. It is obvious that a further purification step is inevitable to make the protein pure enough for crystallization. I adopted an anion exchange chromatography step using a RESOURCE Q 6 ml high resolution anion exchange column to purify VceC.

Buffer exchange of VceC into buffer D (20 mM Tris-HCl, pH: 7.8, 100 mM NaCl, 10% Glycerol, 0.5 mM THP, 0.2% DDM) was carried out on a HiTrap Desalting column. The RESOURCE Q 6 ml column was connected to *AKTA Purifier* system. The column was washed with 5 column volumes (CV) distilled water and 5 CV of buffer E (20 mM Tris-HCl, pH: 7.8, 500 mM NaCl, 10% Glycerol, 0.5 mM THP, 0.2% DDM), then equilibrated with 5 CV of buffer D.

The eluate from the HiTrap Desalting column was loaded onto the RESOURCE Q column equilibrated with buffer D. The column was washed with the same buffer, and then eluted at a flow rate of 1 ml/min with a gradient of 100-400 mM NaCl over 15 column volumes. Fractions containing purified VceC were pooled, diluted 2-fold with buffer D without NaCl, loaded onto the RESOURCE Q column again and eluted with 300 mM NaCl in the same buffer. Figure 3.6a shows the elution profile of anion exchange charomatography for VceC. Figure 3.6b shows the SDS-PAGE analysis of VceC polished by anion exchange charomatography. There are mainly two peaks on the profile, the sharp peak A correspond to the pure fraction of VceC, the purity of which is over 99% of total proteins; the broad peak B correspond to VceC and other contaminating protein.

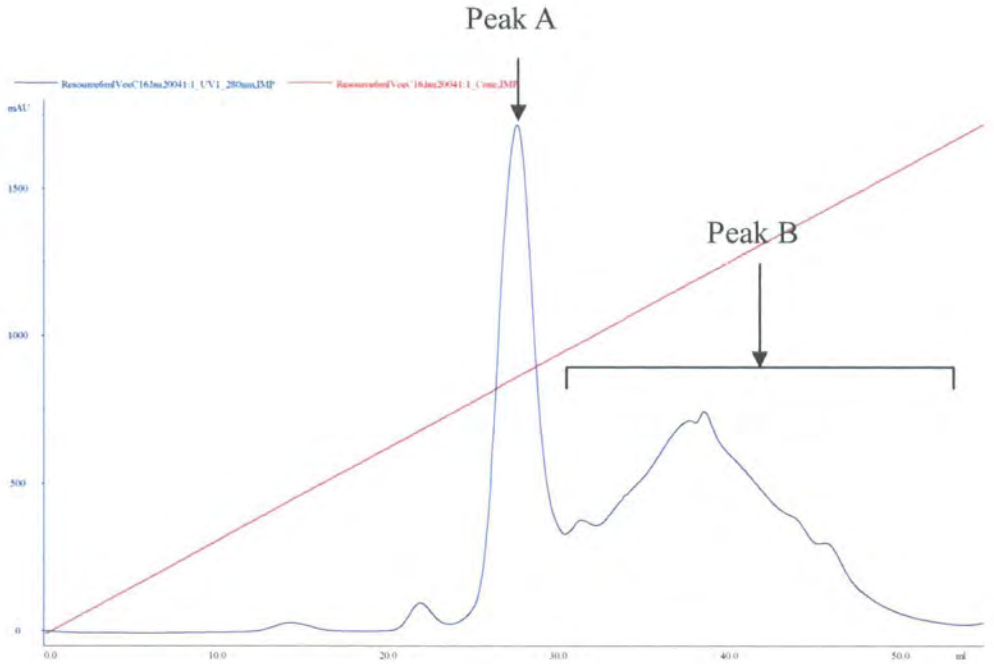


Figure 3.6a. The elution profile of anion exchange chromatography for VceC using a RESOURCE Q 6 ml column. Buffer D: 20 mM Tris-HCl, pH: 7.8, 100 mM NaCl, 10% Glycerol, 0.5 mM THP, 0.2% DDM; Buffer E: 20 mM Tris-HCl, pH: 7.8, 500 mM NaCl, 10% Glycerol, 0.5 mM THP, 0.2% DDM. Flow rate: 1 ml/min. Gradient: 0-70% buffer E in 15 CV.

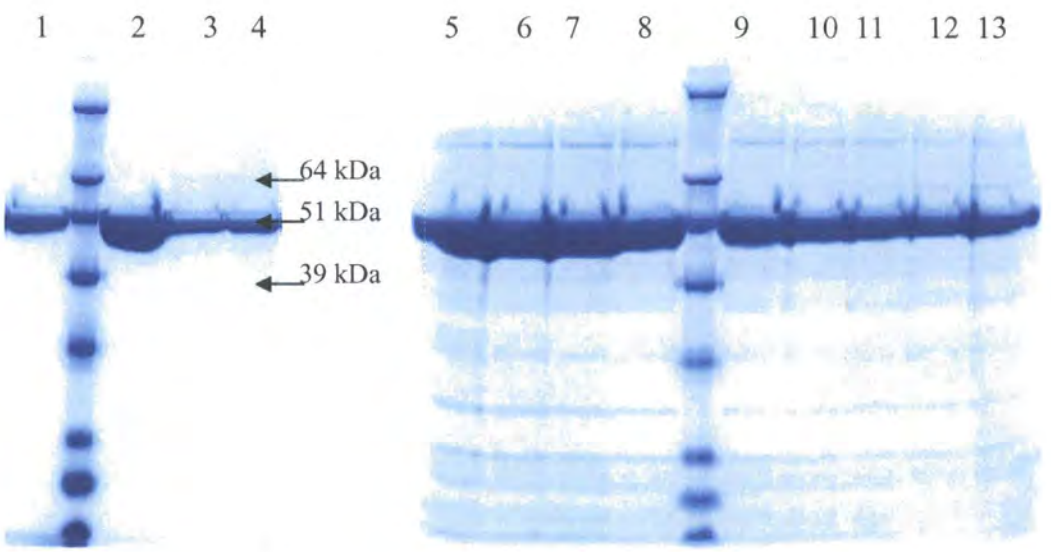


Figure 3.6b Coomassie blue stained SDS-PAGE gel showing VceC protein polished by anion exchange charomatography. Lane 1-4 correspond to fractions from peak A in Fig.3.6a; Lane 5-13 correspond to fractions from peak B in Fig.3.6a.

3.4 The production of a VceC derivative for phasing

The isomorphous replacement method is a commonly used technique for crystal phasing, which needs one or more heavy-atom crystal derivatives. In order to facilitate the binding of mercury compounds to VceC for phasing, a site-directed mutagenesis method was adopted to make a *vceC* cysteine mutant. The residue Glu92 is predicted to be inside the VceC channel according to the structure of its homolog protein TolC. The Glu92→Cys mutation was introduced into *vceC* using the QuikChange® Site-Directed Mutagenesis method (Stratagene) with PCR primers E92CF/E92CR. The construct pET-*vceC* was used as template for the site-directed mutagenesis experiments. The resulting construct pET-$vceC_{E92C}$ was transformed into *E.coli* strain BL21-AI. The VceC carrying the mutation Glu92Cys was produced and purified with the same protocol as native VceC.

3.5 Mutagenesis assays of VceC

Two rings of clustered negative charge of VceC locate approximately near the equatorial domain (residues Glu303 and Glu397) and at the boundaries between the α-barrel and β-barrel domains (residues Asp115, Asp153 and Asp328), respectively. The residues Met93, Arg97 and Glu382 of VceC are supposed to be involved in the interactions with nalidixic acid according to the structure of VceC in complex with nalidixic acid (Figure 3.16). The site-directed mutagenesis method was adopted to produce *vceC* mutants to identify the functional role of these residues.

The *vceCAB* genes were amplified by PCR using the primer pairs VceCF2/VceBR1 to introduce *SacI* and *HindIII* sites at the 5′ and 3′ ends of the amplified DNA, which was purified and ligated into pGEMT-Easy (Promega). The *vceCAB* genes were rescued from pGEMT-easy by restriction digest with *SacI* and *HindIII*; the fragment was ligated into *SacI/HindIII*-digested pQE100 (Qiagen) to create pQE-*vceCAB*. Figure 3.7 shows the restriction digestion analysis of pQE-*vceCAB*.

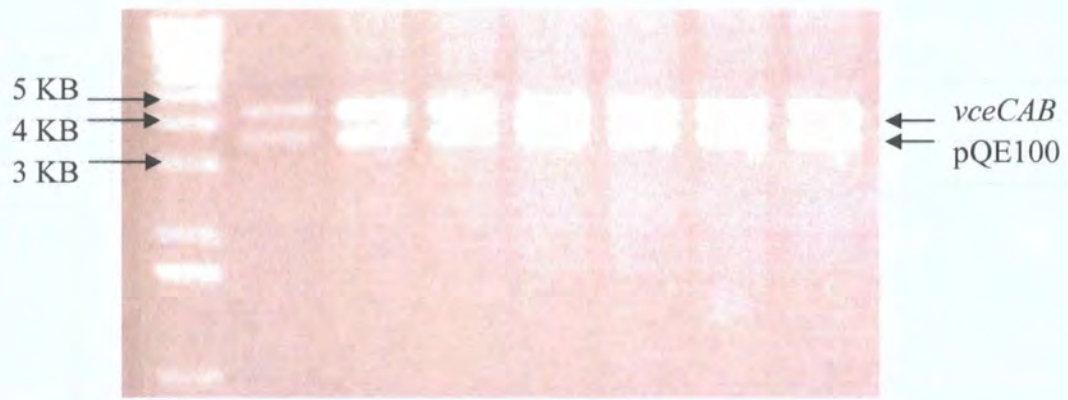


Figure 3.7. Restriction analysis of pET-VceCAB plasmids. The *SacI/HindIII* bounded *vceCAB* genes were ligated into *SacI/HindIII* digested pQE100 vector and transformed into NovaBlue competent cells. The potential pQE100 recombinants were analyzed by PCR first, then recombinant plasmids were prepared from PCR positive colonies and analyzed using *SacI* and *HindIII* restriction digestion (37 °C for 3 hours). All recombinant plasmids exhibit *vceCAB* inserts.

Table 3.1. The effect of *vceC* mutants on the susceptibility of the *acrA* mutantion strain N43 to various chemical compounds and antibiotics.

Compound	Minimum inhibitory concentration (μg ml^{-1})				
	pQE100	*vceCAB*	$vceC_{(M93A,R97A,E382A)}AB$	$vceC_{(D115A,D153A,D328A)}AB$	$vceC_{(E303A,E397A)}AB$
CCCP	16	64	64	64	32
Nalidixic acid	10	10	10	10	10
Rifampicin	4	4	4	4	4
Erythromycin	4	4	4	4	4
chloramphenicol	0.8	0.8	0.8	0.8	0.8

The Asp115Ala+ Asp153Ala+ Asp328Ala (primers: D115AF/ D115AR, D153AF/ D153AR and D328AF/ D328AR), Glu303Ala+ Glu397Ala (primers: E303AF/ E303AR and E397AF/ E397AR) and Met93Ala+ Arg97Ala+ Glu382Ala (Primers: M93AF/ M93AR, M93R97AF/ M93R97AR and E382AF/ E382AR) mutations were introduced into *vceC*, resulting in constructs pQE-$vceC_{(D115A,D153A,D328A)}AB$, pQE-$vceC_{(E303A,E397A)}AB$ and pQE-$vceC_{(M93A,R97A,E382A)}AB$, respectively.

These constructs were checked by automated DNA sequencing and were transformed into *E. coli* strain N43. MICs were measured according to the Microdilution Broth Method established by the National Committee for Clinical Laboratory Standards (NCCLS) (NCCLS Document M7-A2, Methods for dilution antimicrobial susceptibility tests for bacteria that grow aerobically vol. 10, No. 8.). Briefly, a single colony was picked from an LB/ampicillin-plate and used to inoculate 20 ml LB/ampicillin-medium and grown to an A_{600} of 0.1. Cells were transferred to a microtitre plate and mixed with serial two-fold dilutions of the drug to be tested; bacterial growth, in the presence of 0.1 mM IPTG, was monitored after 18 hours incubation at 37 °C.

The N43 cells transformed with pQE-*vceCAB*, pQE-$vceC_{(M93A,R97A,E382A)}AB$ and pQE-$vceC_{(D115A,D153A,D328A)}AB$ were resistant to CCCP, the well-known substrate of the VceCAB pump, with an increase in the MICs from 16 μg ml^{-1} for N43/ pQE100 to 64 μg ml^{-1} for N43/ pQE-*vceCAB*, pQE-$vceC_{(M93A,R97A,E382A)}AB$ and pQE-$vceC_{(D115A,D153A,D328A)}AB$; while the N43 cells transformed with pQE-$vceC_{(E303A,E397A)}AB$ was partially resistant to CCCP with an increase in the MIC to 32 μg ml^{-1}(Table 3.1).

3.6 Crystallization and data collection

Crystallization and data collection were carried out by Dr. Ben Luisi and his group at Cambridge University (Federici *et al.*, 2005).

The VceC protein purified by anion exchange chromatography was dialyzed against buffer F (20 mM Hepes pH 7.5, 300 mM NaCl, 0.05% DDM) and concentrated to 10 mg/ml using 100 kDa MWCO membrane.

The hanging drop vapour diffusion method and the screen kits Crystal Growth Matrices Wizard™ Screens I-II (Emerald BioSystems) were used to screen the initial crystallization conditions. Typically 1 ml of screen solution was pipetteted into the reservoir well. 1 µl of protein sample was mixed with 1 µl of a reservoir solution on a silinated cover slip and then the cover slip was sealed on a well. The crystals were grown in an incubator at 22 °C. The VceC protein yielded crystals in the condition of buffer G (35% 2-methyl-2,4-pentanediol [MPD] v/v, 100 mM Tris pH 7.0, 200 mM NaCl)(buffer no.24 in the Wizard™ Screens I kits). Crystals appeared after one month and reached a final size of 0.1x0.05x0.05 mm^3 (Figure 3.8a). Such crystals diffract up to 2.75 Å.

After the initial hits the condition that produced crystals were optimised to improve the diffraction quality of the crystals. Several additives, detergents and precipitants were tested in the crystallization trials for optimization. The crystals were also obtained in the conditions using PEG 4000 as precipitants (15-20% PEG 4000 at pH 7.5). VceC carrying the mutation Glu92Cys proved to be crystallized more easily, which also yielded crystals of bigger size (up to 0.3x0.2x0.2 mm^3).

The best crystals yielded in the following condition: a droplet made of 1 µl of protein solution, 1 µl of octyl-glucoside 245 mM and 1 µl of reservoir solution (45 % MPD v/v, 100 mM HEPES pH 7.0, 200 mM NaCl). Such crystals diffract up to 1.8 Å. They

appeared after two days and reached their final dimensions after a week (Figure 3.8b). Crystals were directly frozen after looping from the droplet.

The VceC-E92C crystals was soaked for 24 hours with 50 μM phenylmercuribenzoate or Hg acetate after treatment with reducing agent TCEP, but no Hg was found in Patterson map or in MR difference map. The reducing agent TCEP was chosen because it has no thiols, which can form a strong complex with Hg. The strategy was then changed to crystallize VceC-E92C in the presence of Hg salts and other heavy-atom compounds. After pre-treatment of the protein sample with TCEP, the VceC-E92C was successfully crystallized in the presence of 50 μM phenylhydroxyHgbenzoate.

Data collected using an in house X-ray diffractometer (Cambridge University) for the native and Hg labeled VceC-E92C crystals yield a resolution limit to 3.1 Å. Then data were collected at the ID14-4 beamline of ESRF synchrotron source (Grenoble). Complete data of VceC crystals were collected to a resolution of 1.8 Å. Figure 3.9 shows the diffraction map of the VceC crystal. Data were processed and scaled using DENZO and SCALEPACK, respectively. Unit cell dimensions together with statistics about the data processing are summarized in table 3.2.

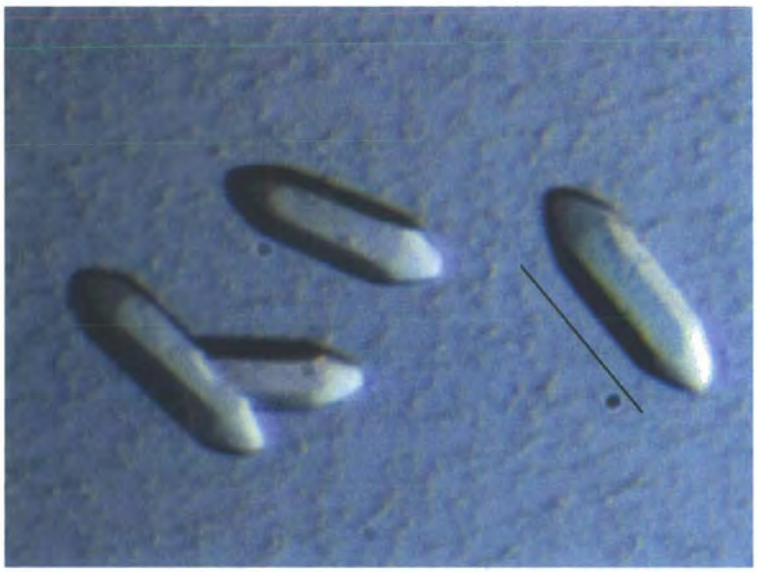


Figure 3.8a. A picture of the crystals from *Vibrio cholerae* VceC. The reservoir solution contained 35% MPD v/v, 100 mM Tris pH 7.0 and 200 mM NaCl. The bar represents about 0.1 mm.

Figure 3.8b. A picture of the crystal from *Vibrio cholerae* VceC. The reservoir solution contained 45 % MPD v/v, 100 mM Tris pH 7.0, 200 mM NaCl; the additives are 81.7 mM OG and 50 μM phenylhydroxyHgbenzoate. The bar represents about 0.15 mm.

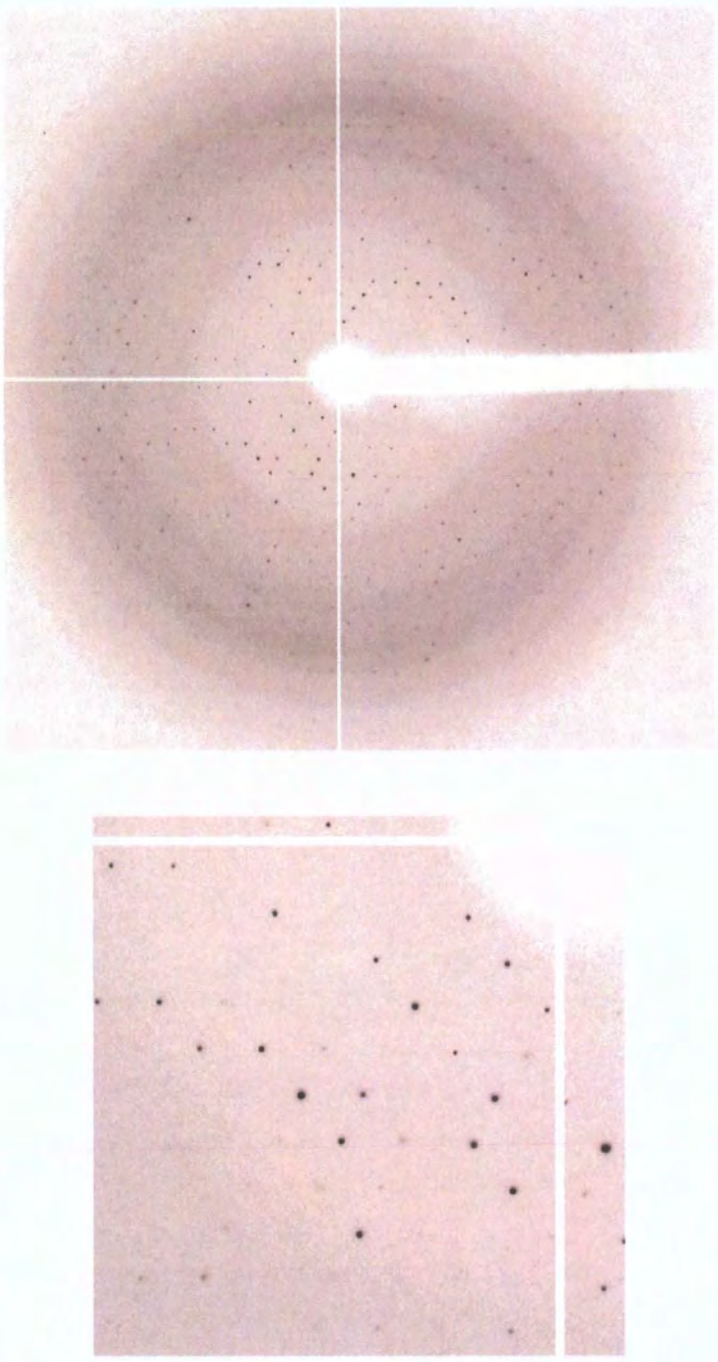


Figure 3.9 X-ray diffraction pattern of VceC crystal. The space group of the crystal is P321, with unit cell lengths a=b=71.93 Å, c=191.21 Å and angles $\alpha=\beta=90^{o}$, $\gamma=120^{o}$.

Table 3.2. Data collection and refinement statistics*

	Native	**Hg derivative**
	In house	ESRF at ID14.4
Data collections		
Wavelength (Å)	1.5418	0.9686
Resolution (Å)	26.17-3.30	63.63-1.80
(last shell)	3.42-3.30	1.90-1.80
R_{merge}	0.093(0.367)	0.076(0.419)
Unique reflections	9230	53173
Completeness	99.6 (99.7)	100 (100)
Multiplicity	4.2	8.0
$I/\sigma(I)$	14.6	5.5
Cell dimensions	a=b=71.929, c=191.213	a=b=71.458, c=190.702
Space group	P321	P321
Refinement		
R (working set)		0.1887
R_{free} (test set)		0.2212
R.m.s. deviations		
Bond lengths (Å)		0.019
Bond angles (Å)		1.751
Ramachandran statistics		
% of residues in allowed regions		96.2
% of residues in generously allowed		3.8
% of residues in not allowed		0
Model		

amino acids	410 of 441
water molecules	424
octyl-β-glucosides	1
mercury atoms	2

*Adapted from Federici *et al.*, 2005.

3.7 Structure solution and refinement

Structure solution and refinement were carried out by Dr. Ben Luisi and his group at Cambridge University (Federici *et al.*, 2005).

Initially, Dr. Ben Luisi's group tried to obtain phase by anomalous scattering from Hg derivative crystals of VceC. Two partially occupied mercury sites were identified by calculating the anomalous difference Patterson map. The phase extracted from anomalous scattering data was used to calculate the electron density maps. However, the resulting maps were not of sufficient quality to allow building of the protein model.

Then they changed strategy to solve the structure by molecular replacement using a polyalanine model of the *E. coli* TolC protomer (pdb code 1ek9) as phasing model (Molrep version 6.2.5 and CNS) (Vagin *et al.*, 2004; Brunger *et al.*, 1998). The best solutions were obtained using data collected in-house on a VceC-Glu92Cys crystal, which indicates that one VceC protomer occupies the asymmetric unit and the VceC trimer is generated by 3-fold crystallographic axis.

The VceC model obtained by molecular replacement was then used as a starting model for refinement against the ESRF 1.8 Å resolution dataset collected on the VceC-Glu92Cys crystal soaked with mercury. Two refinement and model building procedures were adopted. In the first procedure, the calculated phase from the

molecular replacement was used as initial phase to build the VceC model automatically using ARP/wARP (Perrakis *et al.*, 2001), which provided roughly 80% of the model. The rest of the model was then completed manually. In the second procedure, the original polyalanine model of VceC was rebuilt manually using program COOT (Emsley *et al.*, 2004). The ARP/wARP map was used as a guide for model building and the restrained refinement of the model was carried out using REFMAC5 (Murshudov *et al.*, 1999). These two procedures result in the same structure. The density of mercury atoms bound to Cys92 was found, which is consistent with the anomalous Fourier calculations. The density of an octyl-β-glucoside molecule between the β-barrel domains of two symmetry-related molecules was also identified. The final R_{work} is 18.9% and R_{free} is 22.1 %. The VceC model includes residues 63-480 of the cloned coding sequence and 424 water molecules. The first N-terminal 21 residues, the last 4 C-terminal residues and the residues 288-294 of the loop in the equatorial domain of the mature protein are missing due to the absence of clear density. The final quality of the model was checked using PROCHECK (Vaguine *et al.*, 1999), which indicates an excellent model. Statistics relative to the refinement and the quality of the model are shown in Table 3.2.

The programs JOY (Mizuguchi *et al.*, 1998), ODA (Fernandez-Recio *et al.*, 2005), COMPARER (Zhu *et al.*, 1992), GRASP (Nicholls *et al.*, 1991), PYMOL (DeLano Scientific, San Carlos, CA, USA), COOT and ICM (www.molsoft.com) were used to prepare figures.

3.8 Structure of VceC

3.8.1 Overall architecture of VceC

The mature form of VceC from *Vibrio cholerae* contains 442 amino acid residues. The

crystal structure of VceC was determined by molecular replacement at 1.8 Å resolution. The final model of VceC consisted of 410 amino acid residues comprising 93% of the mature protein. The first N-terminal 21 residues, the last 4 C-terminal residues and the residues 288-294 of the loop in the equatorial domain are missing.

The asymmetric unit in the unit cell of crystal is a VceC protomer. The functional unit of VceC is a trimer, which is generated by the crystallographic 3-fold axis. The VceC trimer has a cannon-like appearance with a total length of ~140 Å. VceC shares the same overall architecture as TolC. The trimeric VceC consisted mainly of three domains: the β-domain, the α-domain and the equatorial domain (Figure 3.10).

In the VceC trimer, the strands of the β-domain associate in an antiparallel orientation to form a 12-stranded β-barrel that is right-twisted. The α-helices also form a 12-stranded antiparallel barrel, but this α-helix barrel is left-twisted. The β-sheet and α-helix clusters are connected by four short loops consisting of three to seven amino acid residues.

The α-helix barrel is composed of two types of helix. The long helices (H3 and H7) extend for the entire length of the helical barrel, while short helices (H2-H4 and H6-H8) stack end-to-end so that they resemble pseudo-continuous helices that traverse the length of the helical barrel.

The N-terminal half of VceC (residues 63-286) and the C-terminal half (residue 295-480) share 19.8% sequence identity and the superimposition of the two halves yields an RMS deviation of 2.7 Å between related Cα atoms, suggesting that the protein arose as a result of an intragenic tandem duplication event.

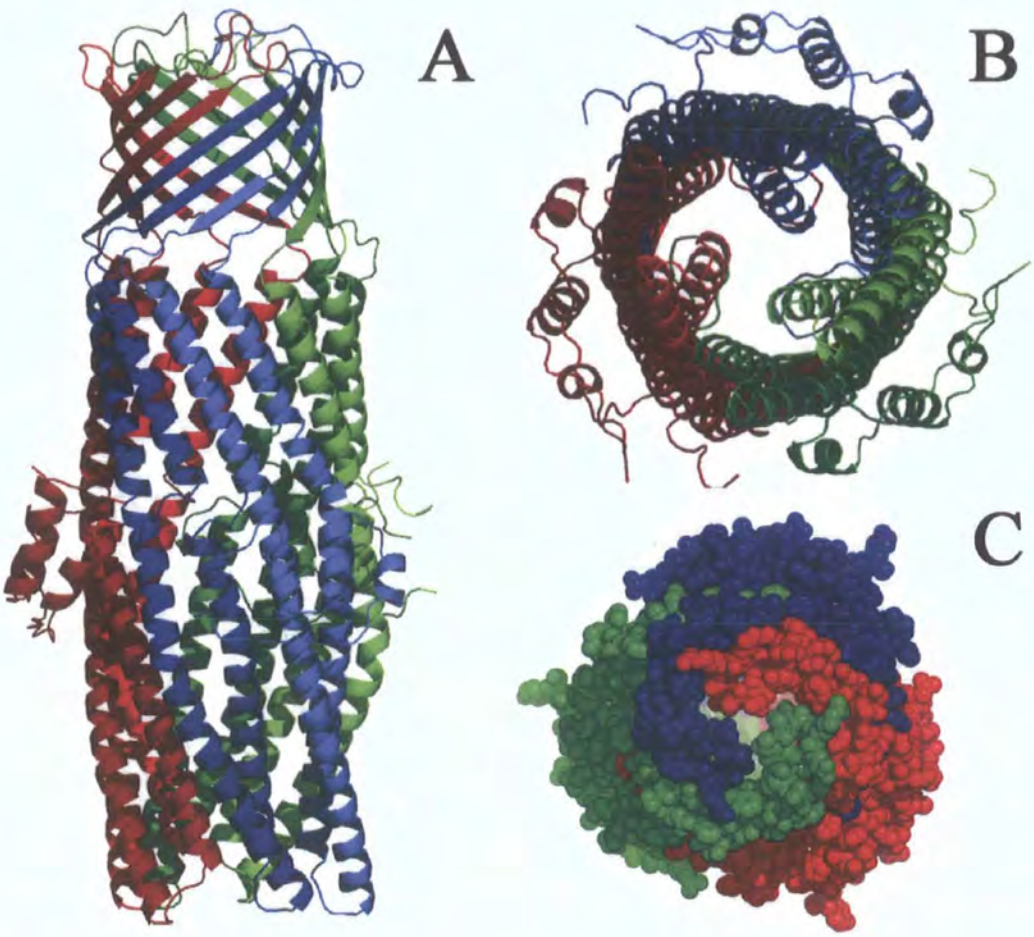


Figure 3.10. Structure of the VceC trimer. Three protomers are individually coloured (blue, green and red). (A) Side view of a ribbon representation, perpendicular to the molecular three-fold axis. (B) A ribbon representation of VceC trimer, viewed from the periplasmic side showing the closure of the α-barrel domain. (C) Space filling representation of the trimer, viewed from the extracellular space (Adapted from Federici *et al.,* 2005).

The porin-like β-barrel domains of VceC trimer pack in laminar sheets in the crystal (Figure 3.11A), commonly referred to as "type I crystal packing". Two octyl-β-glucoside molecules line end-to-end and span the height of the β-barrel domain, nestling between two adjacent β-barrel domains. (Figure 3.11A). The hydrophobic interactions between the aliphatic chains of detergent molecules with the protein, as well as the hydrogen bond interactions between the sugar moieties of detergent molecules with the protein (Figure 3.11B), might mimic protein-lipopolysaccharide interactions that could occur in the outer membrane of Gram-negative bacteria.

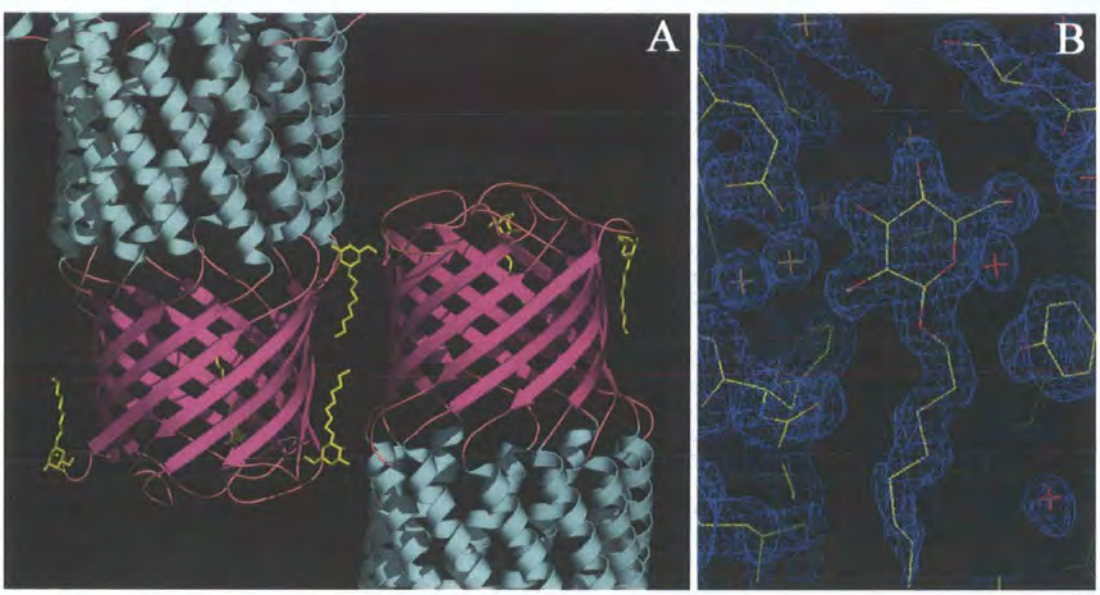


Figure 3.11. (A) Crystal packing of VceC in sheets that mimic the lipid bilayer. The octyl-β-glucoside molecules are coloured yellow. (B) A potion of the $2F_o$-F_c electron density map is shown for the octyl-β-glucoside molecule at the interface between symmetry related protomers. The map is contoured at 1.5 σ (Adapted from Federici *et al.*, 2005).

3.8.2 The β-barrel domain

The β-barrel domain of VceC is inserted in the outer membrane and it spans ~40 Å. The β-barrel is assembled from three protomers, with each protomer contributing four anti-parallel β-strands, which organize in a highly twisted and curved β-sheet that extend to two-thirds the distance along the β-barrel. The aromatic amino acid residues, such as Phe^{114}, Phe^{156}, Phe^{325} and Phe^{364}, cluster in a ring around the base of the β-barrel, delimiting the position of the inner edge of the lipid bilayer. A number of prolines and glycines participate in the precise transitions from right-twisted β-barrel into left-twisted α-helix barrel at the junctions of H2-S1, S2-H3, H6-S3 and S4-H7, respectively.

Strands 1-2 and 3-4 are connected by two long extracellular loops consisting of 13 and 10 residues, respectively. The loop between β-strands S1 and S2 rests over the concave face of the barrel. Three such loops in the trimeric VceC form a pore with diameter of ~6Å. The residues Ala^{134} and Thr^{135} of each protomer provide the boundaries of this pore. Dr. Ben Luisi's group tried to manually dock a variety of well-known substrates of the VceCAB pump, such as carbonyl cyanide *m*-chlorophenylhydrazone (CCCP), onto this pore constriction of VceC and found that they are too large to freely diffuse through the pore if this is considered rigid (data not shown). It is likely that the opening of the porin domain is necessary during transport of these drugs.

3.8.3 The α-helix domain

The α-helical barrel domain is contiguous with the β-barrel domain and protrudes into the periplasm. The α-barrel domain has a length of ~100 Å and constitutes the periplasm exit duct for substrates. The α-helices form a 12-stranded antiparallel barrel,

with each protomer contributing four α-helices (two continuous long helices H3 and H7, and two pairs of shorter helices H2-H4 and H6-H8).

From the β-barrel domain to the equatorial domain, the internal diameter is virtually uniform, which is about 30 Å. However, from the equatorial domain to the periplasmic periphery, the inward twist of helices H7 and H8 tapers the channel gradually, which is almost completely occluded at its periplasmic end. A small pore remains with a diameter of roughly 3.5 Å (Figure 3.10B). This pore is rich in hydrophobic residues from the loop between helices H7 and H8, such as Val^{445}, Ala^{438} and Gly^{435}. The periplasmic end is the narrowest point of the channel, which ends with a ring made by Leu^{441}. A charged ring made by Asp^{442} locates just before the periplasmic end, which might act as an anionic selectivity filter. The helix pairs H7 and H8 form a conventional antiparallel coiled-coil, curving towards the symmetry axis in their C-terminal and N-terminal parts, respectively. Another helix pair comprise a straight helix (H3) around which its partner (H4) is coiling.

This conformation represents the resting state of the protein that must encounter a conformational switch to allow substrates to diffuse into the channel, possibly due to interaction with its partner proteins VceA and VceB.

3.8.4 The equatorial domain

The segments of equatorial domain connect the helices forming the α-barrel domain. We distinguished a helix H1 and a short helix 3_{10} (G5) in this region; however the loop 288-294 was found to be flexible and could not be modelled due to lack of density. Interestingly this loop is located exactly between the two halves of the VceC sequence that constitute structural repeats.

3.9 The open state of VceC

The open state model of VceC was devised by Dr. Ben Luisi's group at Cambridge University. The sequence of the N-terminal halve of VceC resembles that of the C-terminal halve, suggesting that the protein arose as a result of an intragenic tandem duplication event, so that the two halves could make similar interhelical interactions. Dr. Ben Luisi's group aligned the N-terminal repeat (residues 63-286) with the C-terminal repeat (residues 295-480) using CLUSTAL-X (Thompson *et al.*, 1997). The homolog models for the C-terminal repeat were produced based on the crystal structure of the N-terminal repeat using Modeller version 6.2 (Fiser *et al.*, 2003). The resulting 15 models were ranked by energy and violations from ideal geometry, the best of which was further validated using Verify3D (Eisenberg *et al.*, 1997) and PROCHECK. The 'open-state' model of VceC was obtained by merging the experimental and modelled coordinates. The superimposition of the 'open-state' model to the crystal structure of VceC revealed that the β-barrel domain and the first part of the α-barrel domain were perfectly aligned whilst the terminal part of helices H7 and H8 adopted an open conformation.

In the 'open-state' model of VceC, the curvature of helices H7-H8 resembles that of helices H3-H4. The uncoiling movement of the helices H7 and H8 in this model produce a channel aperture of ~20 Å at the periplasmic side, allowing access of substrates into the channel. In Figure 3.12A, the superimposition of the 'open-state' model to the crystal structure of VceC illustrates how the channel may be opened.

The inward twisting of helices H7 and H8 seems to be maintained by two inter-protomer hydrogen bonds between residues Gln^{236}-Ser^{449}, Ser^{243}-Gln^{453} (Figure 3.12B), as well as two intra-protomer polar interactions at the interface of helices H7/H8 and helix H3: a hydrogen bond between residues Tyr^{433}-Glu^{229} and a salt bridge between Arg^{455}-Glu^{247}(Figure 3.12C). In the 'open-state' model of VceC, these

hydrogen bonds and salt bridge are disrupted so that the H4 lose its contact with H7/H8 of the adjacent protomer and the H3 lose its contact with H7/H8 of the same protomer. The hydrophobic residues at the interface of H4-H7, Ile^{428}, Ala^{425} and Val^{445}, become exposed to the solvent upon opening of the channel, which might play a role in binding the adaptor protein VceA when the pump assembles.

3.10 Electrostatic properties of the VceC

In figure 3.13, Dr. Ben Luisi's group has represented the electrostatic potential surface of VceC. In this figure, the front protomer is removed to show the channel interior. The channel interior is generally electronegative. Two rings of clustered negative charge locate approximately near the equatorial domain and at the boundaries between the α-barrel and β-barrel domains, respectively. The ring near the equatorial domain is made by residues Glu^{397} and Glu^{303}; the ring at the boundaries between the α-barrel and β-barrel domains is made by residues Asp^{153}, Asp^{115} and Asp^{328}. These charge distribution of VceC might facilitate substrate movement across the channel.

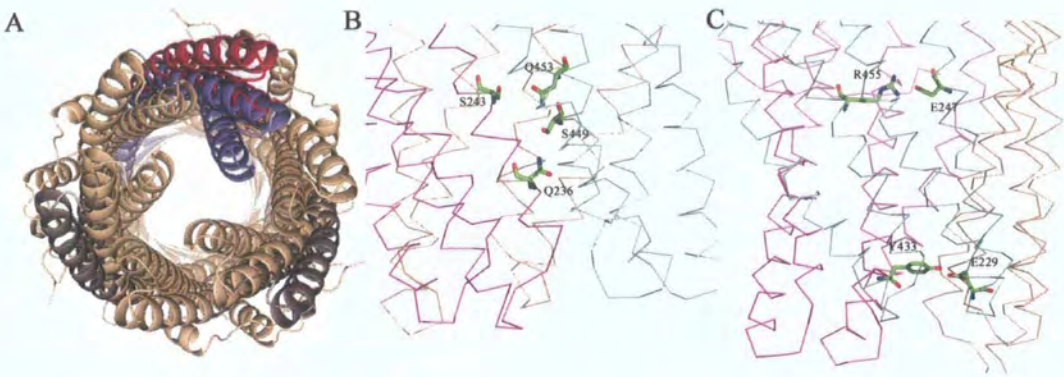


Figure 3.12. (A) The superimposition of the 'open state' model to the crystal structure of VceC illustrates how the channel may be open. The helices H7 and H8 of one protomer are coloured in purple for the closed state and magenta for the open state, respectively. (B) Polar interactions of inter- and intra-protomers that stabilize the closed state of VceC (Adapted from Federici *et al.*, 2005).

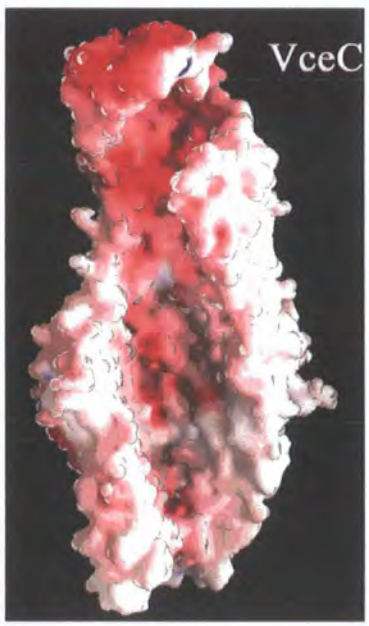


Figure 3.13. Surface representation of the VceC trimer, illustrating the charge distribution The front protomer was removed to show the channel interior. Charges are shown in red for electronegativity, blue for electropositivity, and white for neutrality (Adapted from Federici *et al.*, 2005).

3.11 Superimposition of VceC, TolC and OprM

The structure of VceC was superimposed on the structures of TolC and OprM (Koronakis *et al.,* 2000; Akama *et al.,* 2004). The superimposition of aligned Cα atoms of VceC and TolC yields an RMS deviation of 2.0 Å; the superimposition of VceC and OprM yields an RMS deviation of 1.7 Å; the RMS deviation between TolC and OprM was found to be 1.8 Å. In figure 3.14A, the protein sequence alignment of VceC, TolC and OprM is shown. Despite the very low RMS deviations found, with the overall fold essentially conserved in both the β-barrel and α-barrel domains (Figure 3.14B and 3.14C), the sequence identity between structurally aligned segments is only 8.3%. Moreover, most of the conserved residues of the three proteins are hydrophobic and are likely to play a structural rather than a functional role. For example, the conserved leucine residues play a role in the knobs-into-holes intermeshing between coiled-coiled helices; two conserved prolines participate in the transitions from right-twisted β-barrel into left-twisted α-helix barrel.

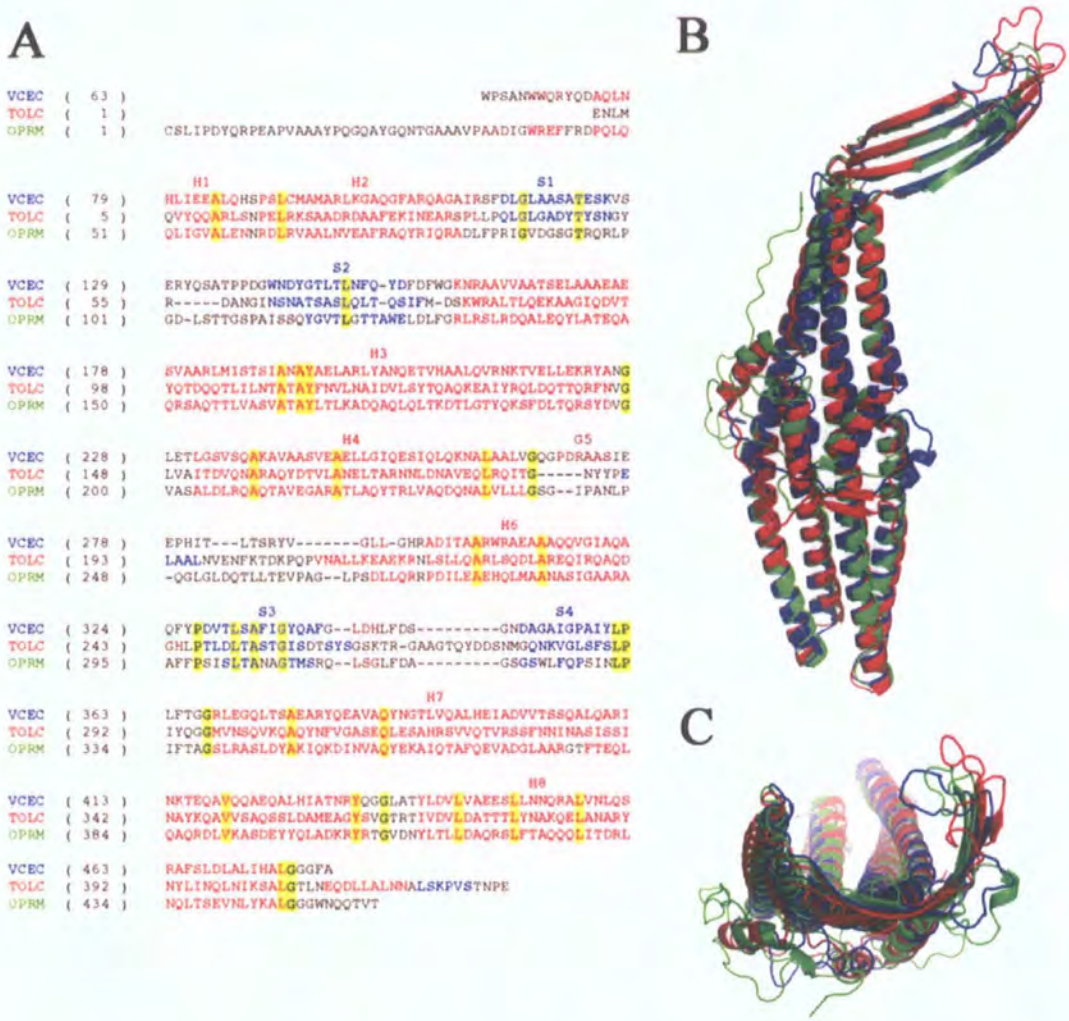


Figure 3.14. Structural comparisons of VceC, TolC and OprM monomers. (A) Sequence alignment of the three proteins. The helices, strands and helix 3_{10} are coloured red, blue and dark red, respectively, and the identical residues are coloured yellow. (B) and (C) Superimposition of monomer structures of VceC (blue), TolC (red) and OprM (green), viewed parallel and perpendicular to the membrane plane, shows the variation in the intra-strand loops of the β-barrel (Adapted from Federici *et al.,* 2005).

.

3.12 Proposed transport mechanism of VceC

This conformation of VceC represents the resting state of the protein that must encounter a conformational switch during transport to open the two closed ends of the channel.

The VceC channel is almost completely occluded at its periplasmic side by the inward twisting of helices H7 and H8. The helix pair H7 and H8 in the C-terminal half have a similar sequence to that of the helix pair H3 and H4 in the N-terminal half, suggesting that they could make similar interhelical interactions. The structure of the N-terminal repeat was used as a template for generating a model of the C-terminal repeat. In the resulting VceC model, the curvature of helices H7 and H8 resembles that of helices H3 and H4. The uncoiling movement of the helices H7 and H8 in this model produce a channel aperture of ~20 Å at the periplasmic side, and thus allow access of substrates into the channel.

The interaction between VceC and its protein partners in the VceCAB pump may trigger the uncoiling movement of the helices H7 and H8. The optimal desolvation energy for the VceC trimer was calculated using the procedure previously adopted for TolC (Fernandez-Recio *et al.*, 2004). The TolC α-helical barrel end is predicted to interact with a protein partner, such as the TolC docking domain of AcrB, the corresponding area in VceC is not (Figure 3.15). This result suggests that VceB and VceC are not in direct contact in the assembled tripartite pump. It is likely that the direct interaction between VceC and VceA is involved in the opening of the periplasmic end of the channel.

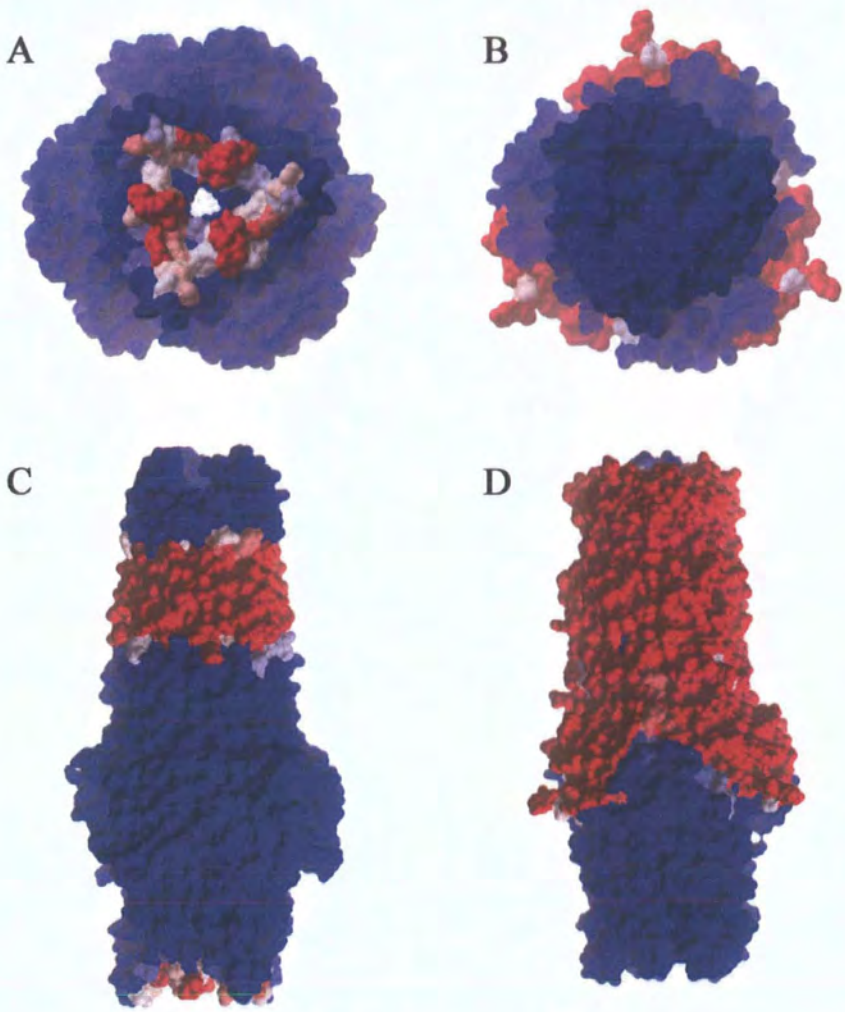


Figure 3.15. Two perpendicular views of the optimal desolvation areas (ODAs) calculated on the surface of TolC (A and C) and VceC (B and D). Residues are coloured red for ODA values <-15.0 kcal/mol, and blue otherwise (Adapted from Federici *et al.*, 2005).

The inward twisting of helices H7 and H8 seems to be maintained by inter-monomer hydrogen bonds between residues Gln^{236}-Ser^{449}, Ser^{243}-Gln^{453}, as well as two intra-protomer polar interactions at the interface of helices H7/H8 and helix H3: a hydrogen bond between residues Tyr^{433}-Glu^{229} and a salt bridge between Arg^{455} -Glu^{247}. In the model of the VceC open state, these polar interactions are disrupted so that H4 loses its contact with H7/H8 of the adjacent protomer and H3 loses its contact with H7/H8 of the same protomer. These could be replaced by contacts with VceA. The coiled-coil domain of VceA could re-pack against the unwound coiled-coils at the proximal end of VceC and thus stabilize its open state.

In the VceC structure, the loop between β-strands S1 and S2 rests over the concave face of the barrel, forming a pore with diameter of ~6Å. These loop regions have the thermal factors in the same range of the contiguous β-strands. Therefore it is hard to exclude that these loops play a role in controlling the diffusion of substrates. For instance, the hydrogen- bonding capability of Thr^{135} at the boundaries of the pore might be instrumental to bind selected substrates. The interaction might then promote a conformational change, allowing substrates to diffuse out into the solvent.

The substrate specificity in the tripartite pumps is supposed to be determined by the inner membrane component (Yu *et al.*, 2003). Once past the entrance of the outer membrane channel, substrates could readily diffuse in the extracellular space. The interior of the channel of VceC is generally electronegative. Two rings of clustered negative charge locate approximately near the equatorial domain and at the boundaries between the α-barrel and β-barrel domains, respectively. These negative rings should attract positively charged drugs into the outer membrane pore, facilitating their movement along the channel (Figure 3.13). The negative ring near the equatorial domain is made by residues Glu^{303} and Glu^{397}. The VceCAB pump partly loses it activity when these two residues were replaced with alanine, indicating their functional role during transport (Table 3.1).

The question remains as to how negatively charged substrates of the VceCAB pump, such as nalixidic acid, might diffuse along the channel. Dr. Ben Luisi's group obtained the crystal structure of VceC in complex with nalixidic acid. Nalidixic acid was identified inside the α-helical barrel domain of VceC (Figure 3.16). Mutagenesis experiments are in progress to test the importance of the surrounding residues. This structure suggests that the negatively charged substrates could be actively transported across the channel. The outer membrane protein VceC can not directly use energy from the proton motive force; however, the proton motive force could produce conformational change of VceB that may be transduced to VceC by a conformational coupling of VceA and thereby transport substrates across the channel.

The antimicrobial resistance feature afforded by two tripartite multidrug pumps, the MexA/MexB/OprM (Zhao *et al.*, 1998; Zhao *et al.*, 2002) and MtrC/MtrD/MtrE (Rouquette-Loughlin *et al.*, 2002) efflux pump, were shown to require TonB/ExbB/ExbD proteins, which has been identified in many Gram-negative bacteria and probably represents a conserved mechanism for the transfer of energy from the proton motive force to high affinity receptors within the outer membrane (Biswas *et al.*, 1997; Mey *et al.*, 2003). It is possible that the TonB/ExbB/ExbD complex facilitate the opening of the VceC channel and the substrates transport across the channel, in analogy to the function of OprM (Zhao *et al.*, 1998).

The VceC trimers pack in laminar sheets in the crystal. Within this membrane-like layer, the molecules of octyl-β-glucoside line tail-to-tail (end-to-end), bridging two VceC protomers. The hydrophobic interactions between the aliphatic chains of detergent molecules with the protein, as well as the hydrogen bond interactions between the sugar moieties of detergent molecules with the protein, might mimic protein-lipopolysaccharide interactions that could occur in the outer membrane of Gram-negative bacteria.

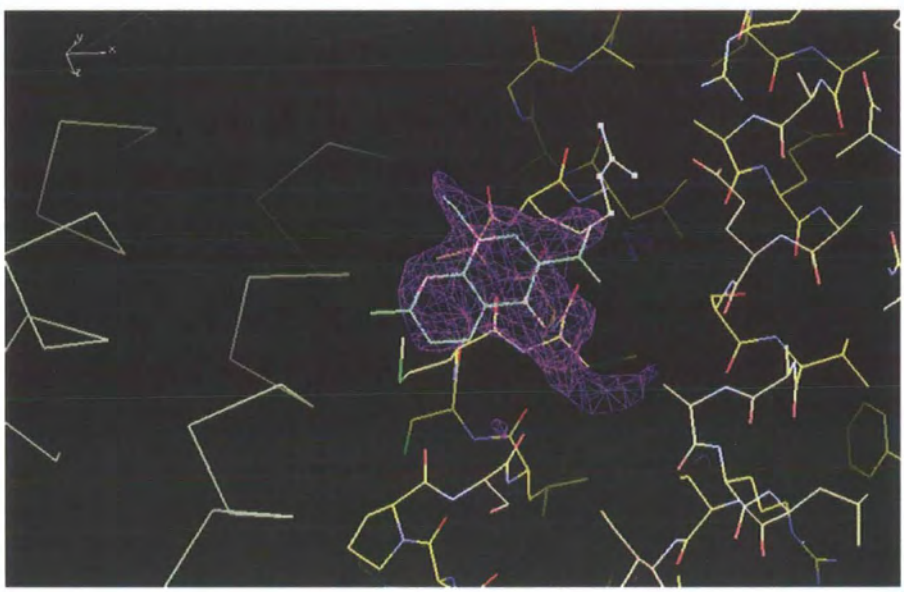


Figure 3.16. The structure of VceC in complex with nalidixic acid. The electron density map is shown for nalidixic acid, which was identified inside the α-helical barrel domain of VceC.

Chapter 4

Overexpression, purification and characterization of VceB and VceR

The crystal structures of the three components of tripartite multidrug efflux pump, AcrB, AcrA and TolC, have been determined. The question remains as to how the three components assemble to form a functional pump. Obviously, the crystal structures of the IMP-MFP-OMP complex or the complex of either of the two components will be critical to decipher the molecular mechanism of these pumps. Different models of the protomer stoichiometry for tripartite pumps have been proposed (Higgins *et al.*, 2004; Akama *et al.*, 2004; Fernandez-Recio *et al.*, 2004), however, these models are based on the structure of the RND family transporter protein AcrB, which in contrast to other types of transporter has a big periplasmic domain.

VceB is the inner membrane component of the VceCAB pump, which belongs to the MF family of proton antiporters. Like other MF family proteins, VceB is predicted to lack the extensive periplasmic domain that is typical of proteins belonging to the RND family (Colmer *et al.*, 1998). It seems likely that the MFS-based pumps have a different organization and use different types of protein-to-protein interactions in the tripartite pumps.

Membrane transport proteins are difficult for overexpresion, purification and crystallization, and the lack of structures of multidrug efflux pumps has retarded our understanding of these transporters. The cytosolic multidrug-binding transcriptional regulators of these pumps provide an alternative in studying the multidrug recognition mechanisms.

VceR is established as a repressor of the *vceCAB* operon (Woolley *et al.*, 2005; Borges-Walmsley *et al.*, 2005), which belongs to the TetR/CamR family of

transcriptional regulators. These regulators typically function as dimers that bind palindromic operators in which the termini of the inverted repeats are about 15 bp apart. However, some of these family regulators, such as QacR, function as a pair of dimers and bind to 28 bp palindromic operators (Grkovic *et al.*, 2001).

In this chapter, I report the successful overproduction of the complex of VceA (12-406) and VceB-VceA fusion protein, which could be a candidate to co-crystallize these two components of the VceCAB pump. I analyzed the oligomerization states of DNA-bound and drug-bound VceR.

4.1 Overexpression and purification of VceB-VceA fusion protein

4.1.1 Construction of the pET-B-A plasmid

The pET-B-A construct that was used to express the VceB-VceA fusion protein was generated by several steps. The *vceB* gene was amplified by PCR from genomic DNA of the *V. cholerae* vaccine strain CVD101 (primers: VceBF/VceBR2) and ligated into pGEM-T Easy vector. After restriction enzyme digestion of the vector with *NdeI* and *XhoI*, the resulting *vceB* fragment was then ligated into pET21a, generating the construct pET-B-I. The *vceB* gene in pET-B-I was amplified again using primer pairs VceBF/HisR, the resulting *vceB-6His* fragment bounded by *NdeI* and *SacI* was ligated into *NdeI* and *SacI* digested pET21a vector, generating the construct pET-B-II. The forward and reverse primers were designed so as to introduce an N-terminal enterkinase-tag to *vceA*, and the *SacI* and *HindIII* restriction sites were incorporated into these primers to allow fragment ligation into pET-B-II via *SacI* and *HindIII* sites (primers: VceA-EKF/VceAR2), generating the construct pET-B-A. This construct generated a fusion protein with VceA following the C-terminus of VceB and a His_6-enterkinase linker (LEHHHHHHELDDDDK) between them.

4.1.2 Recombinant expression of VceB-VceA fusion protein

The *E.coli* strain C43 (DE3) was used as the host strain to express the VceB-VceA fusion protein. The pET-B-A construct was transformed into C43 (DE3). Transformed colonies were selected on the basis of vector-encoded ampicillin resistance, and isolated plasmids were screened for the presence of the inserted *vceB-vceA* sequence by restriction digestion.

A single C43 (DE3)/pET-B-A colony from a freshly streaked plate was picked and used to inoculate 10 ml 2×YT broth, containing carbenicillin (100μg/ml), in a 25 ml universal container. The cells were grown in an orbital shaker at 37°C, 220 rpm. overnight. One liter of prewarmed 2×YT broth (carbenicillin 100μg/ml) in a 2 liter baffled flask was inoculated with 2 ml of the overnight culture. The cells were grown in an orbital shaker at 37 °C, 220 rpm, and their growth was monitored by measuring the absorbance of the culture at 600 nm (A_{600}). The cells were induced at mid log phase (approx. A_{600}=0.5-0.6) with 0.3 mM IPTG. The temperature was dropped to 25 °C and the growth was continued with shaking at 200 r.p.m. for 4-5 hours. The cells were harvested by centrifugation at 4,350 ×g for 8 min at 4 °C, resuspended in buffer A (Na_2HPO_4 20 mM, pH: 7.2, NaCl 300 mM, Glycerol 10%) at 5 ml per gram wet weight. DNase I (5U/ml) and EDTA-free protease inhibitor mixture tablet (1 tablet/100 ml) were added to the cell suspension and either stored at −80 °C or purification continued.

4.1.3 Purification of the His-tagged VceB-VceA fusion

Purification of VceB-VceA fusion protein by Co^{2+} affinity chromatography was performed using TALON Metal Affinity Resin purchased from Clontech.

The cells were lysed by three passages through a Constant System Cell Disrupter (15 Kpsi, model Z-plus 1.1 kW, Constant System). Cellular debris was removed by centrifugation at 24,000 ×g for 30 min at 4 °C. The supernatant, containing soluble proteins and membrane vesicles, was carefully decanted so as not to disturb the pelleted material. The supernatant was further fractionated into membrane and soluble components by high-speed ultracentrifugation at 122,400 ×g for 90 min at 4°C. This procedure resulted in the generation of a supernatant containing cytosolic proteins and small molecules and a pellet of membranous material.

The membrane pellet was emulsified in a small volume of buffer A by repeated passage through a syringe needle, and then the emulsified membrane was diluted by buffer A to a final concentration of total protein of about 5-10 mg/ml. A powder of the detergent DDM was added to the emulsified membrane to a final concentration of 1-2% to solubilize the membrane. The mixture was gently stirred at 4 °C for 1-2 hours and then ultracentrifuged at 122,400 ×g for 1 hour at 4 °C. The membrane solution was carefully decanted so as not to disturb the pelleted material. Imidazole (500 mM, pH 7.2) was added to the membrane solution to a final concentration of 10 mM.

Extracted histidine-tagged VceB-VceA fusion protein was purified using the TALON Metal Affinity Resin immobilized with Co^{2+}. Two ml of Talon affinity resin were washed twice with 20 ml of deionised water and twice with 10 ml buffer B (Na_2HPO_4 20 mM, pH: 7.2, NaCl 300 mM, Glycerol 10%, DDM 0.2%) by centrifugation at 4,000 ×g for 10 min. The Talon affinity resin was equilibrated in 10 ml buffer B for 1 hour on ice, and then sedimented by centrifugation at 4,000 ×g for 10 min.

Two ml of equilibrated Talon affinity resin were add to 50 ml of membrane solution and mixed gently by stirring at 4°C for 2 hours. Then the sedimented Talon affinity resin was washed, using a batch-wise procedure, three times with 10 ml of 30 mM imidazole in buffer B. The Talon affinity resin was then loaded onto a column with the bottom outlet capped. The bottom cap was removed and the column flow-through

was collected. The column was washed twice with 10 ml of 50 mM imidazole in buffer B, the wash fractions were collected for SDS-PAGE analysis. The protein was eluted 6 times with 0.5 ml of elution buffer (500 mM imidazole in buffer B). The eluates were collected for SDS-PAGE analysis.

4.2 Overexpression and purification of VceR

The *vceR* gene was amplified by PCR using the forward and reverse primers VceRF/VceRR, respectively. I incorporated *BamHI* and *XhoI* restriction sites into these primers to allow fragment ligation into pET21a, generating a construct pET-VceR to express the VceR with a C-terminus hexahistidine-tag.

The pET-VceR construct was used to transform *E. coli* strain C41 (DE3), providing expression of the His_6-tagged protein. A single colony of C41 (DE3)/pET-VceR was picked from a freshly streaked plate and used to inoculate 20 ml of LB broth, containing 100 μg/ml carbencillin. The culture was grown at 37°C overnight with vigorous shaking. One liter of prewarmed 2×YT broth (with antibiotic), in a 2 liter baffled flask, was inoculated with 3 ml of the noninduced overnight culture. The culture was grown at 37°C with vigorous shaking until an OD_{600} of 0.5-0.6 was reached, at which point IPTG was added to a final concentration of 0.5 mM, and the growth continued for 3-4 hours at the reduced temperature of 25 °C. The cells were harvested by centrifugation at 4,350 ×g at 10 °C for 8 min and resuspended in buffer C (50 mM Tris pH 7.0, 300 mM NaCl, 10% (v/v) glycerol, 1 mM THP, 1 tablet/100 ml protease inhibitor cocktail tablet, 5 U/ml DNase I).

The cell suspension was passed three times through a Constant Systems cell disrupter (15 Kpsi, model Z-plus 1.1kW, Constant Systems) at 4 °C to disrupt the cell, and then incubated on ice for 15 min. The lysate was centrifuged at 122,400 ×g for 60 min at 4 °C to pellet the cellular debris. VceR was isolated from the supernatant using a HiTrap

chelating HP 1 ml column immobilized with Ni^{2+} that had been equilibrated with buffer D (50 mM Tris pH 7.0, 300 mM NaCl, 10% Glycerol, 0.5 mM THP). Imidazole (5 M, pH 7.0) was added to the supernatant to a final concentration of 15 mM and the mixture was loaded onto the column using a peristaltic pump at a speed of 1 ml/min. The column was washed with 10-15 ml of 50 mM imidazole in buffer D. The VceR was then eluted from the column with 300 mM imidazole in buffer D. The eluates were collected as 1 ml fractions and analyzed by SDS-PAGE. Positive identification was achieved through Western blotting with Mouse-Anti-His (BioRad). Figure 4.1 shows the SDS-PAGE analysis of VceR.

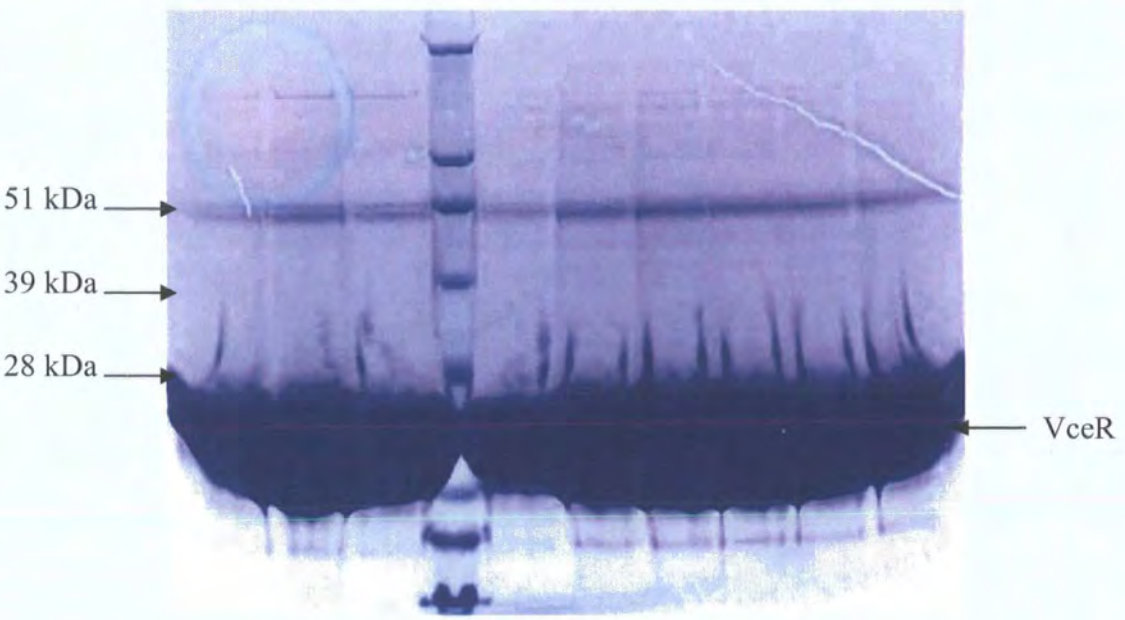


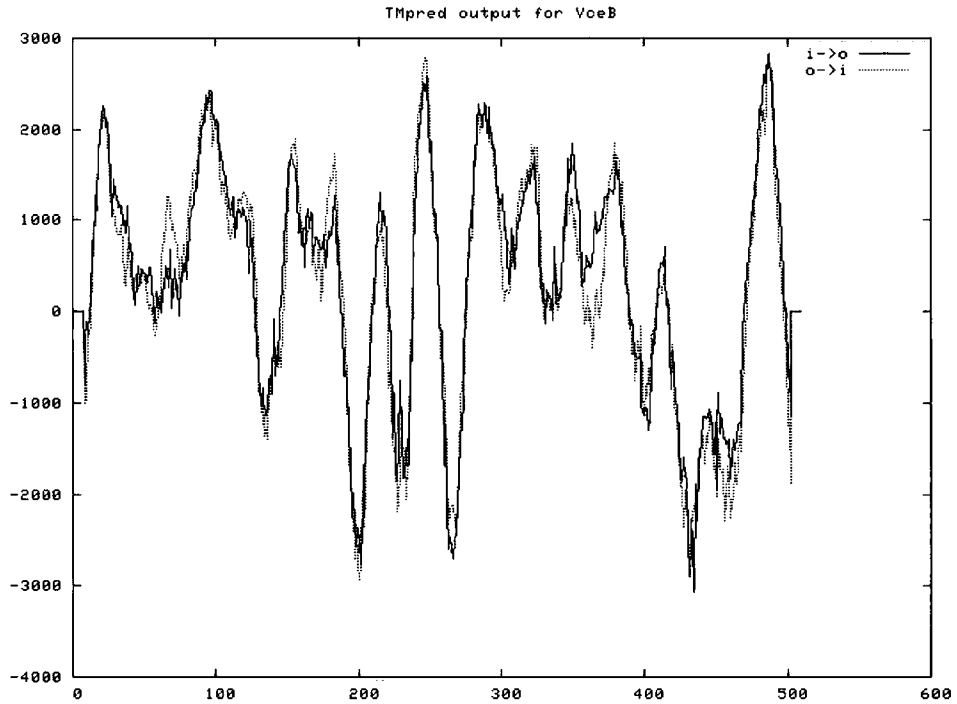
Figure 4.1. A Coomassie blue stained SDS-PAGE gel showing IMAC purification of VceR. The left side arrows mark the position of protein standards; the right side arrow marks the position of VceR protein.

4.3 Expression of VceB-VceA fusion protein

The *vceB* gene is 1536 bp long and encodes a protein of 511 amino acids. The VceB protein contains 48.92% hydrophobic amino acids (AILFWV), strongly suggesting that it associated with the cell membrane. Hydropathy analysis using the program TMPred indicated that VceB contains 14 transmembrane α-helices that may span the cytoplasmic membrane, with its N- and C-terminus in the cytoplasm (Figure 4.2). VceB shares 39.5% amino acid sequence identity, 60.3% sequence similarity to a characterized 14-TMS MF family transporter EmrB from *E.coli*, suggesting that VceB is a 14-TMS MF family transporter.

Figure 4.3 shows the SDS-PAGE analysis of the VceB-VceA fusion protein. On SDS-PAGE, the VceB-VceA fusion protein migrated at around 75 kDa, rather then at the predicted monomeric mass of 101.2 kDa. It is typical for membrane transporter proteins to migrate at 65-75% of their molecular weight on SDS-PAGE (Saidijam M et al., 2003; Mokhonov V et al., 2005), and thus the migration of the VceB-VceA band on SDS-PAGE would be expected at 65.8-75.9 kDa.

Matrix Assisted Laser Desorption Ionization Time-of-flight Mass Spectrometry (MALDI MS) analysis confirmed that the band 1 and band 2 were VceB-VceA and VceA (12-406), respectively (Figure 4.3). It is obvious that the construct pET-B-A expressed two proteins: the VceB-VceA fusion protein and VceA (12-406). VceA (12-406) was co-purified with VceB-VceA because it does not contain a His_6-tag.



STRONGLY prefered model: N-terminus inside, 14 strong transmembrane helices, total score: 26198

helices	from	to	length	score	orientation
1	14	33	(20)	2260	i-o
2	56	79	(24)	1278	o-i
3	87	105	(19)	2430	i-o
4	110	133	(24)	1299	o-i
5	143	166	(24)	1712	i-o
6	175	194	(20)	1737	o-i
7	208	224	(17)	1302	i-o
8	239	256	(18)	2806	o-i
9	277	300	(24)	2293	i-o
10	314	334	(21)	1834	o-i
11	340	361	(22)	1842	i-o
12	368	391	(24)	1853	o-i
13	404	426	(23)	707	i-o
14	480	499	(20)	2845	o-i

Figure 4.2. TMpred prediction of transmembrane regions and orientation of VceB. Prediction parameters: TM-helix length between 17 and 33. i-o: inside to outside helices; o-i: outside to inside helices.

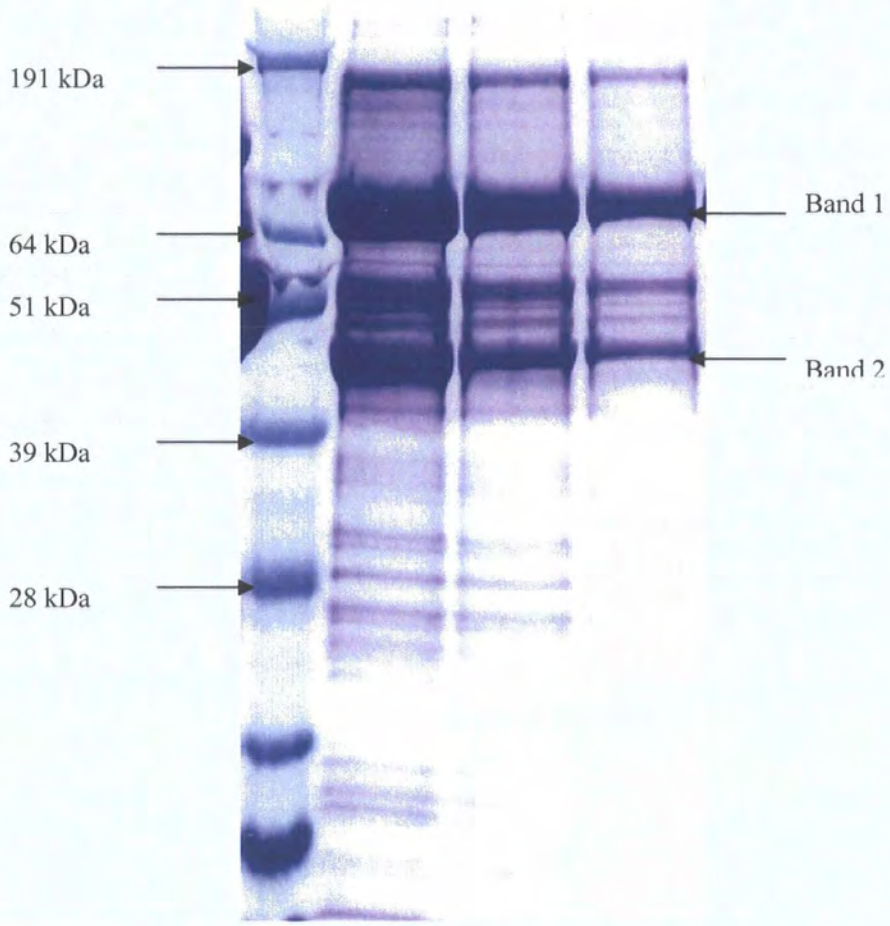


Figure 4.3. Coomassie blue stained SDS-PAGE gel showing partial IMAC purification of VceA(12-406) and VceB-VceA fusion protein from DDM solubilized membranes. The left side arrows mark the position of protein standards; the right side arrows mark the position of VceB-VceA fusion protein (band 1) and VceA (12-406) (band 2), respectively.

4.4 A pair of VceR dimers binds the *vce* promoter

The *vce* promoter was identified to be a 28 bp inverted repeat DNA sequence (Woolley *et al.*, 2005; Borges-Walmsley *et al.*, 2005). The complementary oligonucleotides corresponding to the 28 bp inverted repeat sequences of the *vce* promoter were annealed (oligonucleotides VceR-PF/VceR-PR) for DNA binding experiments. Analytical gel filtration chromatography was used to estimate the molecular weight (M_r) of VceR and of the VceR-DNA complex using a Superdex 200 PC 3.2/30 column, equilibrated with buffer E (50 mM Tris pH 7.0, 300 mM NaCl, 1 mM THP). The protein standards ribonuclease A, chmotrypsinogen A, ovalbumin and albumin were solubilized in buffer E and used to prepare a calibration curve. The molecular weights of VceR and VceR-DNA complex were determined using this calibration curve (see Materials and Methods).

Passage of the VceR protein through a gel filtration column produced a single protein peak, detected from the 280 nm absorbance, with a calculated M_r of 46.5 kDa that is consistent with a homodimer of VceR (with a predicted M_r of 49.8 kDa). However, passage of the VceR that was pre-incubated with molar ratios of between 0.1 and 0.5 of the 28 bp promoter DNA of *vceCAB*, through gel filtration column produced two peaks. One peak consistent with a homodimer of VceR, while the second peak had a calculated M_r of 118.9 kDa, which is in good agreement with the theoretical mass of 118.0 kDa for four VceR molecules bound to the *vceCAB* promoter DNA (Figure 4.4 a and b). The increase of molar ratio of promoter DNA to VceR shifted the profile so that more VceR proteins were located to the tetramer form, consistent with this interpretation. I failed to determine whether the substrates of the VceCAB pump, such as CCCP, could induce dissociation of the VceR(tetramer)–(dsDNA) complex, as CCCP precipitated upon application to the column. Interestingly, I found that VceR could bind ssDNA (e.g. either strand of the promoter dsDNA) to form a complex that had a calculated M_r of 28.4 kDa, which is in agreement with the theoretical mass of 34.1 kDa for one VceR molecule bound to the ssDNA, suggesting that the dimer

interface was destabilized by the presence of ssDNA (Figure 4.4b).

The oligomeric state of VceR was also examined in the absence of both DNA and drug, in the presence of dsDNA and in the presence of both dsDNA and drug, using an alternative technique of dynamic light-scattering (DLS). The DLS experiments were performed in a Zetasizer Nano instrument (Malvern Instruments, UK). The measurements were carried out in buffer E with 1 or 10% glycerol at 25 °C after filtration through a 0.22 μm filter. Test calibrations of the instrument were performed with lysozyme. The molecular weight (M_r) of VceR could be calculated from the particle size determined by DLS.

The DLS measurements of the VceR in buffer E containing 10% (v/v) glycerol yielded a molecular mass of 24.4 kDa (radius 2.330 nm) that is highly consistent with that predicted for the monomer (24.9 kDa), but there was no change in particle size upon addition of dsDNA (Figure 4.4c); while the DLS measurement of the VceR in buffer E containing 1% (v/v) glycerol yielded a molecular mass of 68.9 kDa (radius 3.634 nm). The addition of dsDNA caused a further increase to 141.1 kDa (radius 4.937 nm), and then the addition of CCCP reduced the molecular mass to 67.5 kDa (Figure 4.4c). In the absence and presence of dsDNA, the molecular weights of VceR are higher than expected for the VceR dimer (49.8 kDa) and tetramer–DNA complex (118 kDa). The VceR dimer and VceR(tetramer)–dsDNA complex may adopt an elongated conformation that leads to the M_r being over-estimated by this technique. Upon addition of CCCP to the VceR(tetramer)–dsDNA complex, the M_r was reduced to 67.5 kDa, indicating that two dimers of VceR are induced to dissociate from the dsDNA.

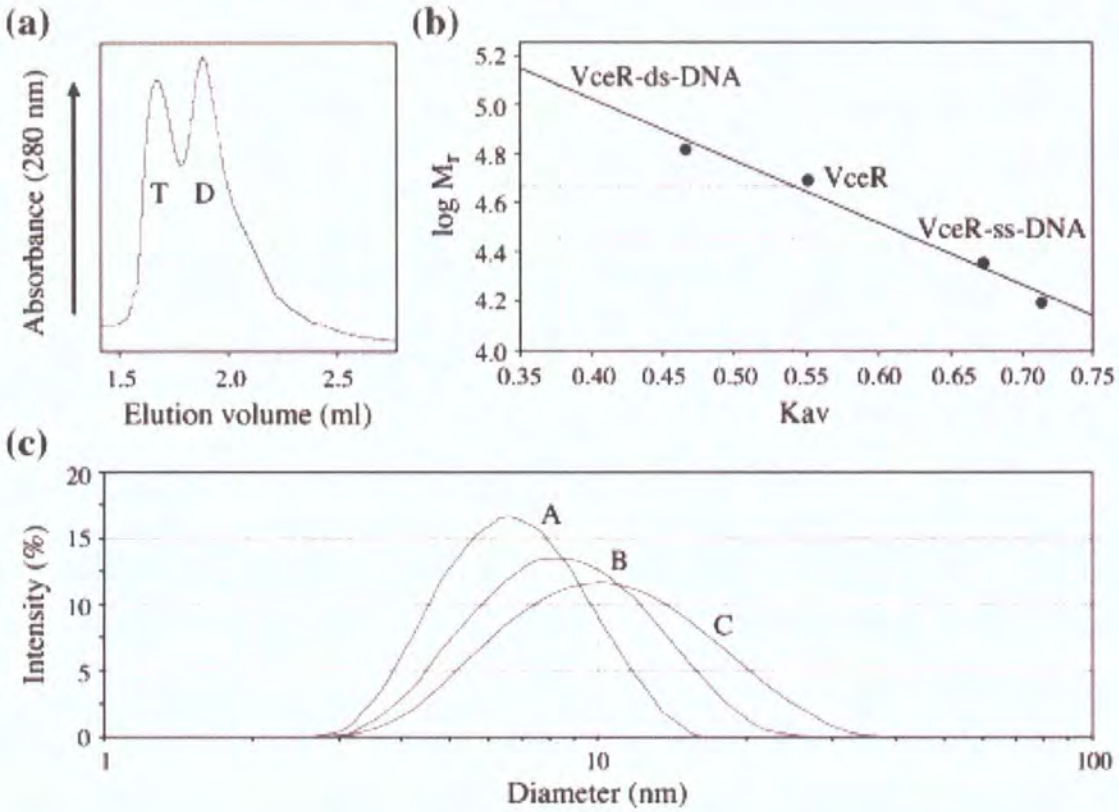


Figure 4.4. The oligomeric state of VceR. (a) The elution profile of VceR-dsDNA on a Superdex 200 PC 3.2/30 size-exclusion column. VceR and dsDNA were preincubated with a molar ratio of 2:1. Running buffer: 20 mM Tris pH 7.0, 300 mM NaCl, 1 mM THP. Flow rate: 20 µl/min. Peak D and peak T indicate VceR dimer and VceR (tetramer)–dsDNA complex, respectively. (b) A standard curve constructed from the elution volumes of the Low Molecular Weight Gel Filtration Kit. The positions of VceR, VceR–dsDNA and VceR–ssDNA complexes are shown on the curve. (c) Size distribution curves of VceR determined by dynamic light-scattering assays. Peak A, B and C show VceR in running buffer with 10% glycerol, 1% glycerol and 1% glycerol/dsDNA, respectively. Upon addition of CCCP to the VceR(tetramer)– dsDNA complex, a curve produced that was superimposed upon that for VceR alone in the presence of 1% glycerol (Adapted from Borges-Walmsley *et al.*, 2005).

4.5 Discussion

VceB

Overexpression of polytopic IMPs in *E.coli* is still a challenge. Several factors likely account for this difficulty. First, the biogenesis of most polytopic IMPs depends on the Sec-translocase. Heterologous expression of polytopic IMPs in *E.coli* will competitively exclude production of vital host membrane proteins, leading to toxicity (Van den Berg *et al.,* 2004; Roosild *et al.,* 2005). Second, the signal sequence of the IMP is critical for its targeting and insertion in the inner membrane. The signal sequence of helerologous IMPs may not be effectively recognized by *E.coli* (Dalbey *et al.,* 2004; Roosild *et al.,* 2005). Thirdly, the IMP needs to properly fold in the membrane and assemble with it partner subunits. Misfolding and misassembling of helerologous IMPs expressed in *E.coli* can be toxic to the cell. There is a quality control system that removes misfolded and misassembled helerologous IMPs. It has been discovered that FtsH, an ATP-dependent protein, is involved in such a quality control (Kihara *et al.,* 1999). Besides, the expression of toxic proteins in *E.coli* strains causes plasmid instability: bacteria which have lost the expression plasmids rapidly overgrow those retaining it (Dumon-Seignovert *et al.,* 2004).

The toxicity problem of expression was overcome for a number of membrane proteins using two host strains, called C41 (DE3) and C43 (DE3). By induced mutation of BL21 (DE3), these two host strains were selected that grew to high saturation density and continue to produce proteins at an elevated level without toxic effects (Miroux *et al.,* 1996; Dumon-Seignovert L *et al.,* 2004). Fusion partner proteins have been used to aid production of recombinant membrane proteins. The fusion partner protein Mistic bypasses the Sec-dependent pathway of IMP biogenesis. By using Mistic as a fusion partner protein, several IMPs have been overexpressed, which fully fold in the native conformation (Roosild *et al.,* 2005). The co-expression of IMPs with their partner subunits may improve the expression level of functional protein (Kihara *et al.,*

1995).

Overexpression of VceB in *E.coli* was highly toxic. We failed to overproduce this protein alone using several *E.coli* strains and expression vectors (data not shown). This problem was overcome by expression of the VceB-VceA fusion protein using C43 (DE3) as host strain. Like other tripartite multidrug efflux pumps, VceA and VceB could assemble together in the inactive state. The misassembling of VceB could result in toxicity to the cell when it is expressed alone.

The C-terminus end of VceB is supposed to be located at the cytoplasm side of the inner membrane, and VceA has an N-terminal in-outward transmembrane α-helix, it is reasonable for the VceB-C-terminus-fused VceA to express its periplasmic domain in the periplasm. However, the construct pET-B-A expressed two proteins: the VceB-VceA fusion protein and VceA (12-406). Colmer *et al.* (1998) identified two possible ribosome-binding sites 10-20 bp upstream of the ATG code of Met12. It is likely that the *vceA* gene starts at Met12 rather then postulated Met1 (Heidelberg *et al.*, 2000).

The characterized MFS transporters function as monomers and lack the large periplasmic domain present in RND-transporters (Abramson *et al.*, 2003; Huang *et al.*, 2003; Hirai *et al.*, 2004; Yin *et al.*, 2006). VceB could also function as a monomer. The periplasmic gap may vary between 100 and 250 Å according to the physiology of the Gram-negative bacteria (Dubochet *et al.*, 1983; Graham *et al.*, 1991). It is necessary for VceA to function as an oligomer to form a channel between VceB and VceC, so that the tripartite pump could span the periplasm. I have found that the periplasmic domain of VceA forms a trimer (this study). Therefore, it is not surprising that VceA (12-406) co-purified with VceB-VceA.

The key to understanding the mechanism of tripartite multidrug efflux pumps will lie in the structural solution of the IMP-MFP-OMP complex; however, this represents a

considerable challenge. Although it was shown that the three components of tripartite multidrug efflux pump form a stable complex in the inactive state (Tikhonova *et al.* 2004; Touze *et al.,* 2004; Gerken *et al.,* 2004), they are very easy to dissociate *in vitro*, resulting in the difficulty to co-crystallize the complex. The VceB-VceA fusion protein and VceA (12-406) form a stable complex during purification, which could provide an invaluable way for co-crystallization of these two components of the VceCAB pump.

VceR

The rate of the CCCP-induced conformational change of VceR determined by stopped-flow fluoresce spectroscopy revealed that the binding of CCCP to VceR in both the presence and absence of DNA is a highly co-operative process that occurs with a Hill coefficient of about four, suggesting that each functional unit of VceR interacts with four CCCP molecules (Borges-Walmsley *et al.,* 2005). The gel filtration and DLS experiments indicated that in the absence of promoter dsDNA, the protein exists predominantly as a dimer, while in the presence of promoter dsDNA, the protein exists as VceR (tetramer)-dsDNA complex. The DLS experiments also indicated that VceR (tetramer)-dsDNA complex dissociates into VceR dimers in the presence of CCCP. Taken together, these findings suggest that each VceR subunit has multiple drug-binding sites. Another TetR/CamR family transcriptional regulator, QacR, has also been shown to have multiple over-lapping sites that can simultaneously accommodate bound drugs (Schumacher *et al.,* 2001; Schumacher *et al.,* 2004). Taking into account the fact that the binding of CCCP to dimeric VceR is a co-operative process with a Hill coefficient of 4, it is reasonable to propose that the VceR monomer possesses a pair of binding-sites that can be simultaneously occupied by CCCP, and that these sites interact with one another and with those in the second subunit of the dimer. This is in contrast to QacR, in which the two subunits of the dimer do not appear to interact.

Chapter 5

Overexpression, purification and characterization of VceA

The membrane fusion protein (MFP) is a component of the tripartite multidrug efflux pump in Gram-negative bacteria. MFPs are anchored to the inner membrane by a single transmembrane helix, or via a fatty acid attached to its N-terminal cysteine residue (Borges-Walmsley *et al.*, 2003; Yoneyama *et al.*, 2000). MFPs contain a large periplasmic domain, which is predicted to have a α-helix hairpin and lipoyl structure (Johnson *et al.*, 1999).

Different models of the protomer stoichiometry for tripartite pump have been proposed. The MFP was predicted to be a trimer, hexamer or enneamer in an assembled tripartite pump (Fernandez-Recio *et al.*, 2004; Akama *et al.*, 2004; Higgins *et al.*, 2004). These models are based on the structure of RND family transporter AcrB, which has a big periplasmic domain. The problem remains as to how the MFPs of MFS-transporters or ABC-transporters interact with their IMP and OMP because they lack the large periplasmic domain present in RND-transporters. It is likely that different arrangement of MFPs can be expected in MFS and ABC-based tripartite pumps.

In this chapter, I report the initial structural characterization and oligomeric state of VceA, the MFP of an MF tripartite pump VceCAB.

5.1 Overexpression and purification of VceA

5.1.1 Construction of plasmids to overexpress VceA

VceA is predicted to anchor to the inner membrane by a single α-helical domain. Different *vceA* constructs were made to test this hypothesis. The full length *vceA* and two truncated fragments, *vceA* (12-406) and *vceA* (53-406), were generated by PCR using primers VceAF/VceAR2, VceA12F/VceAR2 and VceA53F/VceAR2, respectively. The plasmid pQE30 was used as expression vector to overexpress *vceA, vceA* (12-406) and *vceA* (53-406) in *E. coli* strain M15 (pREP4).

The pQE30-*vceA*, pQE30-*vceA* (12-406) and pQE30-*vceA* (53-406) were generated by subcloning the fragments of pGEMT/*vceA*, pGEMT/*vceA (12-406)* and pGEMT/*vceA (53-406)*, bounded by *SacI* and *HindIII*, into the MCS of expression vector pQE30. These resulted in the generation of constructs to express *vceA, vceA* (12-406) and *vceA* (53-406) with an N-terminal hexahistidine tag. Figure 5.1 shows the restriction enzyme analysis of these constructs.

5.1.2 Overexpression of the VceA

A single colony of M15 (pREP4)/ pQE30-*vceA* or M15 (pREP4)/ pQE30-*vceA* (12-406) or M15 (pREP4)/ pQE30-*vceA* (53-406) was picked from a freshly streaked plate and used to inoculate 20 ml of LB broth containing 100 μg/ml carbencillin and 25μg/ml kanamycin. The culture was grown at 37°C overnight with vigorous shaking. One liter of prewarmed 2×YT broth (with antibiotics) in a 2 liter baffled flask was inoculated with 3 ml of the non-induced overnight culture. The culture was grown at 37°C with vigorous shaking until an OD_{600} of 0.5-0.6 was reached. The cells were induced by adding 1 M IPTG to a final concentration of 0.2 mM, and then the culture was incubated for an additional 2-3 hours at 30 °C. The cells were harvested by centrifugation at 4,350 ×g at 4 °C for 8 min and the cell pellets were resuspended in

buffer A (20 mM Tris pH: 7.5, 300 mM NaCl, 20% Glycerol, 5 U/ml DNase I, 1 tablet/100 ml protease inhibitor).

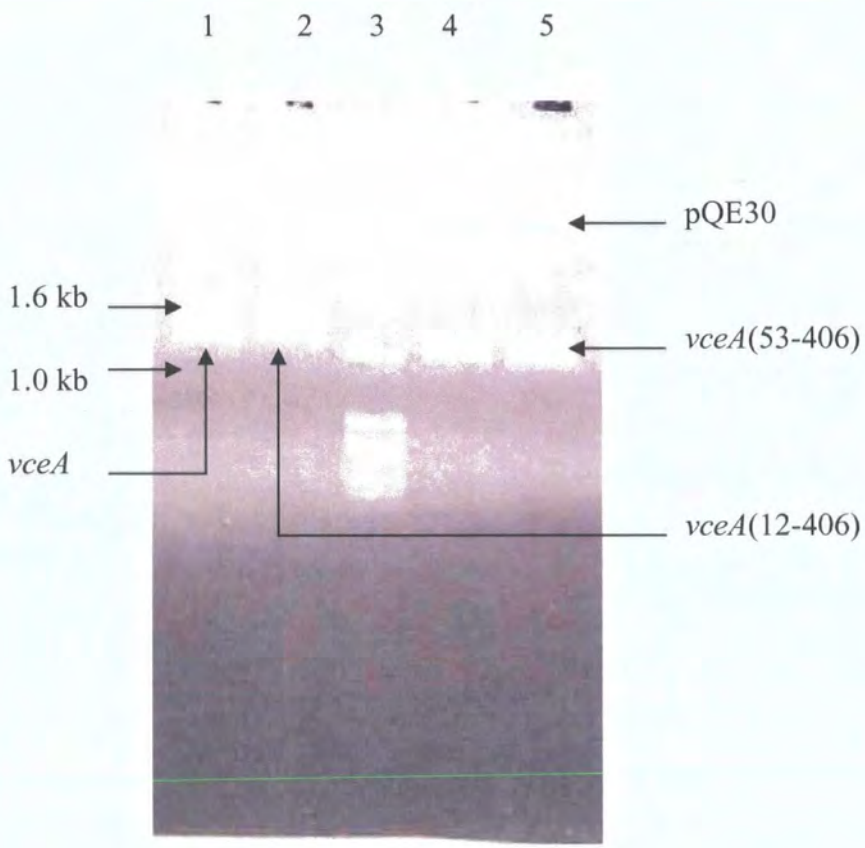


Figure 5.1. Restriction analysis of pQE30-*vceA*, pQE30-*vceA* (12-406) and pQE30-*vceA* (53-406) plasmids. The *SacI/HindIII* bounded *vceA*, *vceA* (12-406) and *vceA* (53-406) genes were ligated into *SacI/HindIII* digested pQE30 vector and transformed into NovaBlue competent cells. The potential pQE30 recombinants were screened by PCR first, then recombinant plasmids were prepared from PCR positive colonies and analyzed using *SacI* and *HindIII* restriction digestion (37 °C for 3 hours). All recombinant plasmids exhibit target gene inserts.

The cells were lysed by passing the cell suspension three times through a Constant Systems cell disrupter at a pressure of 15 Kpsi. The crude lysate was then fractionated into insoluble cellular debris, soluble cytosolic components of the cell and cellular membrane by differential centrifugation (see Materials and Methods). The VceA (53-406) was isolated from the soluble cytosolic components; the VceA and VceA (12-406) were isolated from the cellular membrane fraction.

5.1.3 Purification of His-tagged VceA using Ni^{2+} affinity chromatography

For purification of VceA (53-406), NaCl was added to the soluble cytosolic components to a final concentration of 500 mM, imidazole (500 mM, pH: 7.5) to 15 mM and THP to 0.5 mM, and then 2 ml of Ni-NTA (Qiagen) slurries were added to 100 ml of the mixture and mixed gently by stirring at 4°C for 2 hours. The mixture was loaded into a column with the bottom outlet capped. The bottom cap was removed and the column flow-through was collected. The column was washed with 40 ml of 40 mM imidazole in buffer B (20 mM Tris pH: 7.5, 300 mM NaCl, 20% Glycerol, 0.5 mM THP) and 20 ml of 60 mM imidazole in buffer B; wash fractions were collected for SDS-PAGE analysis. The protein was eluted 6 times with 0.5 ml elution buffer (500 mM imidazole in buffer B). The eluates were collected and analyzed by SDS-PAGE.

For purification of VceA and VceA (12-406), the membrane pellets were emulsified in a small volume of buffer B by repeated passage through a syringe needle, and then the volume of membrane mixture was adjusted using buffer B at 10 ml per gram of membrane (wet weight). DDM was added to the membrane suspension to a final concentration of 2% and THP to 0.5 mM. The mixture was rotated at 4°C for 2 hours to solubilize the membrane. The mixture was then ultracentrifuged at 122,400 ×g for 60 min to sediment insoluble membrane debris. NaCl was added to the membrane solution to a final concentration of 500 mM, and imidazole (500 mM, pH: 7.5) to 15

mM. Then 2 ml of Ni-NTA slurries were added to 100 ml of the mixture and mixed gently by stirring at 4°C for 2 hours. The mixture was loaded into a column with the bottom outlet capped. The bottom cap was removed and the column flow-through was collected. The column was washed with 40 ml of 40 mM imidazole, 0.2% DDM, in buffer B and 20 ml of 60 mM imidazole, 0.2% DDM, in buffer B; wash fractions were collected for SDS-PAGE analysis. The protein was eluted 6 times with 0.5 ml elution buffer (500 mM imidazole, 0.2% DDM in buffer B). The eluates were collected and analyzed by SDS-PAGE. Figure 5.2 shows an SDS-PAGE analysis of VceA, VceA (12-406) and VceA (53-406). VceA was also extracted from DDM solubilised cellular membranes without THP in the buffer during purification. There were mainly two protein bands (Figure 5.2 D1-3). After treating the protein sample (protein sample 20 μl +β-mecaptoethanol 1 μl+ protein loading buffer 4 μl, 99°C 30 min), one band disappeared (Figure 5.2 D4-6).

5.1.4 Western blotting identification of VceA

The Western blotting method was employed to confirm that the protein was indeed VceA. The VceA sample was separated by SDS-PAGE, transferred to a polyvinglidene difluoride membrane by electroblotting, and then immunoblotted with a mouse monoclonal primary antibody raised against the polyhistidine tag epitope. The result of Western blotting is shown in Figure 5.3. Western blotting analysis of untreated and treated (e.g. 20 μl protein, 1 μl β-mecaptoethanol, 4 μl protein loading buffer, 99°C 30 min) VceA samples established that the higher molecular mass band is a VceA dimer that dissociates when treated with β-mercaptoethanol and boiling.

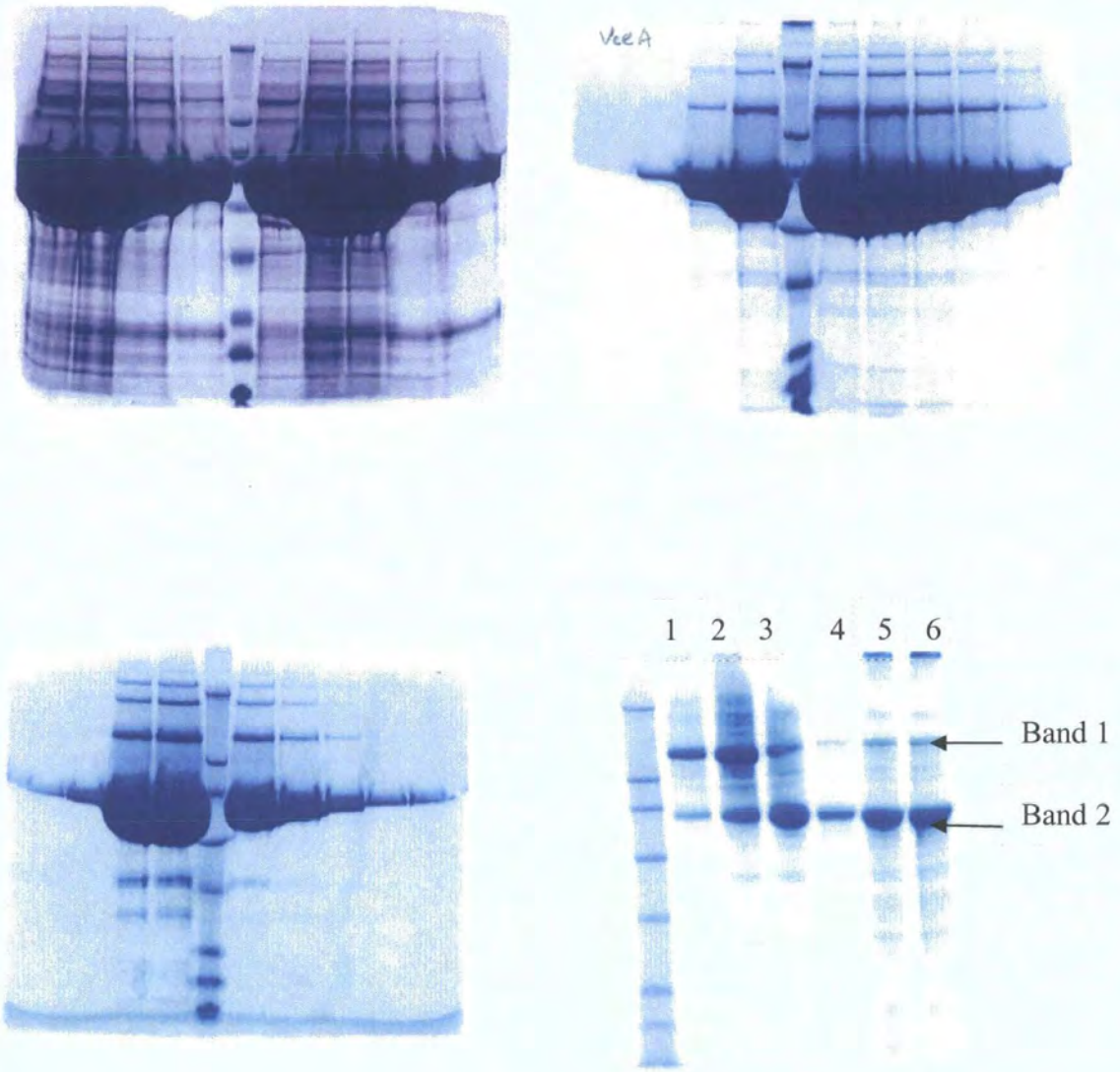


Figure 5.2 Coomassie blue stained SDS-PAGE gel showing partial purification of VceA, VceA (12-406) and VceA (53-406). (A) VceA (53-406) isolated from soluble cytosolic components of the cell. (B) and (C): VceA and VceA (12-406) extracted by DDM from cellular membrane, respectively. (D) VceA extracted by DDM from cellular membrane without THP in the buffer during purification. There are mainly two protein bands (lane 1-3). After treating the protein sample (20 µl protein, 1 µl β-mecaptoethanol, 4 µl protein loading buffer, 99°C 30 min), band 1 disappeared (lane 4-6).

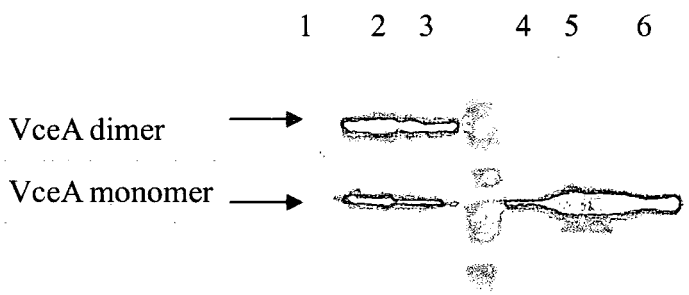


Figure 5.3. The Western blotting above illustrates the strong reaction between an anti-His_6 antibody and the approximately 50 kDa and 100 kDa proteins with high specificity. Lane 1-3: untreated VceA samples; Lane 4-6: treated VceA samples (20 µl protein, 1 µl β-mecaptoethanol, 4 µl protein loading buffer, 99°C 30 min). The higher molecular mass band disappears after treating the samples, indicating that it is VceA dimer.

5.2 Secreted expression and purification of the periplasmic domain of VceA

5.2.1 Construction of the pBAD/gIII-VceAp plasmid

VceA (53-406) was predicted to be the periplasmic domain of VceA (VceAp). The secreted expression of VceAp in the periplasm of *E.coli* strain TOP10 was conducted using expression vector pBAD/*gIII*. The pBAD/*gIII* is a pBR322-derived expression vector. The *araBAD* promoter (P_{BAD}) and the *araC* gene of the vector provide tight, dose-dependent regulation of heterologous gene expression, while the *geneIII* signal sequence allows secreted recombinant protein expression into the periplasmic space of *E.coli* (Invitrogen).

The forward and reverse primers were designed so as to introduce an N-terminus hexahistidine-enterokinase tag to VceAp (primers: VceApF/VceAR2). I incorporated *SacI* and *HindIII* restriction sites into these primers to allow fragment ligation into pBAD/*gIII*, generating a construct pBAD/*gIII*-VceAp to express the VceAp with an N-terminus secretion signal peptide G III and hexahistidine-enterokinase tag. Figure 5.4 shows the restriction enzyme analysis of pBAD/*gIII*-VceAp.

5.2.2 Overexpression and purification of VceAp-His_6 by Ni^{2+} affinity chromatography

A single colony of TOP10/pBAD/*gIII*-VceAp was picked from a freshly streaked plate and used to inoculate 20 ml LB broth containing 100 μg/ml carbencillin. The culture was grown at 37°C overnight with vigorous shaking. One liter of prewarmed 2×YT broth (with antibiotic) in a 2 liter baffled flask was inoculated with 3 ml of the noninduced overnight culture. The culture was grown at 37°C with vigorous shaking until an OD_{600} of 0.5-0.6 was reached. The cells were induced by adding 20% L-arabinose to a final concentration of 0.02%, and then the culture was incubated for

an additional 2-3 hours at 30 °C. The cells were harvested by centrifugation at 4,350 ×g at 10 °C for 8 min.

The cold osmotic shock procedure (Qiagen) was used to purify VceAp from the periplasm of TOP10. The cell pellets of TOP10 /pBAD/*gIII*-VceAp was resuspend in buffer C (30 mM Tris·Cl, 20% sucrose, pH 8.0) at 60 ml per gram wet weight. The cells were kept on ice and 500 mM EDTA was added dropwise to a final concentration of 1 mM. The cells were incubated on ice for 20 min with gentle agitation. The cell suspension was centrifuged at 4,350 ×g for 20 min at 4°C, all the supernatant was removed and the pellet was resuspended in the same volume of ice-cold 5 mM $MgSO_4$. The mixture was stirred for 20 min at 4°C and then centrifuged at 4,350 ×g for 20 min at 4°C. The supernatant is the osmotic shock fluid containing periplasmic proteins. Protease inhibitor was added to the supernatant (1 tablet per 500 ml). Tris·Cl (1 M, pH: 7.5) was added to the supernatant to a final concentration of 20 mM, NaCl to a concentration of 200 mM and imidazole (500 mM, pH: 7.5) to a concentration of 10 mM, and then 3 ml of Ni Sepharose™ 6 Fast Flow slurries were added to 500 ml of the mixture and mixed gently by stirring at 4°C overnight. The mixture was loaded onto a column with the bottom outlet capped. The bottom cap was removed and the column flow-through was collected. The column was washed with 80 ml of 30 mM imidazole in buffer D (20 mM Tris pH: 7.5, 300 mM NaCl, 20% Glycerol) and 20 ml of 45 mM imidazole in buffer D; wash fractions were collected for SDS-PAGE analysis. The protein was eluted 8 times with 1 ml elution buffer (500 mM imidazole in buffer D). The eluates were collected and analyzed by SDS-PAGE. Figure 5.5 illustrates an SDS-PAGE analysis of VceAp purified by this procedure. Detection with Western-blotting confirmed that the major band on the SDS-PAGE gel was VceAp-His_6.

Figure 5.4. Restriction analysis of pBAD/gIII-VceAp plasmids. The *SacI/HindIII* bounded *vceAp* gene was ligated into *SacI/HindIII* digested pBAD/gIII vector and transformed into TOP10 competent cells. The potential pBAD/gIII recombinants were analyzed by PCR first, then recombinant plasmids were prepared from PCR positive colonies and analyzed using *SacI* and *HindIII* restriction digestion (37 ^{o}C for 3 hours). All recombinant plasmids exhibit *vceAp* inserts.

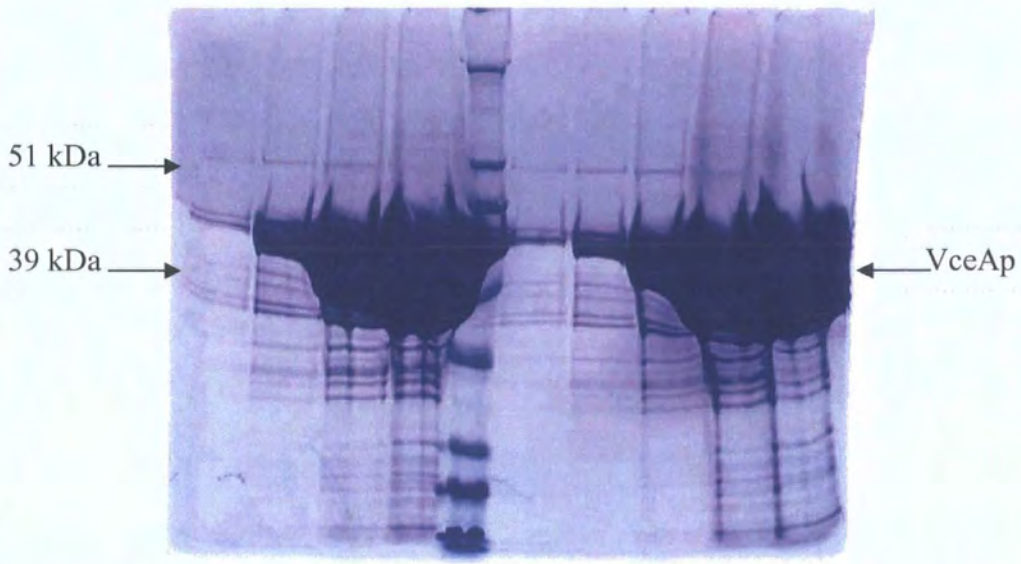


Figure 5.5. A Coomassie blue stained SDS-PAGE gel showing the purification of VceAp protein from the periplasm of *E.coli* strain TOP10 using the cold osmotic shock procedure. The protein was isolated from six liter cultures of TOP10/pBAD/gIII-VceAp. The left arrows mark the positions of protein standards; the right arrow marks the position of VceAp protein.

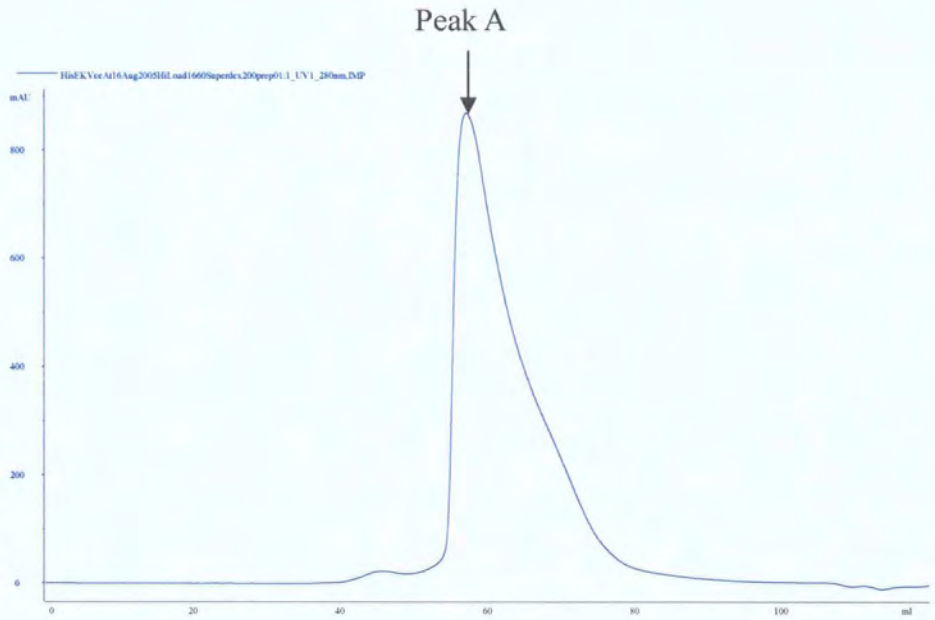


Figure 5.6A. An elution profile of VceAp on a HiLoad 16/60 superdex 200 prep grade size-exclusion column. Running buffer: 10 mM Tris, pH: 7.6, 300 mM KCl. Flow rate: 1 ml/min.

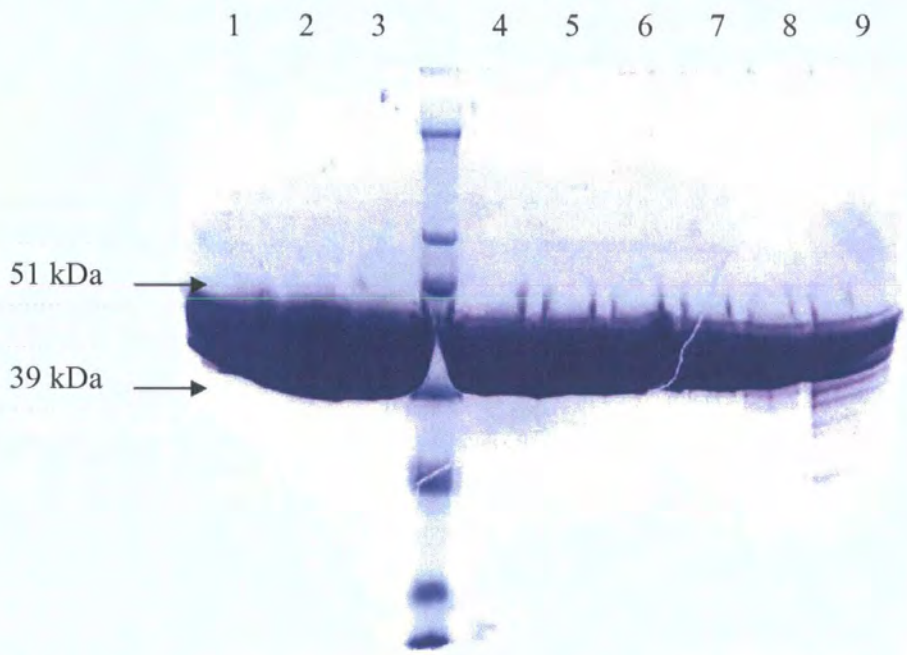


Figure 5.6B. Coomassie blue stained SDS-PAGE gel showing VceAp protein polished by size-exclusion charomatography. Lane 1-9 correspond to fractions from peak A in Figure 5.6A.

5.2.3 Purification of VceAp by gel filtration chromatography

The VceAp purified by immobilized Ni^{2+} absorption chromatography was subjected to gel filtration chromatography on a HiLoad 16/60 superdex 200 prep grade column run on an *AKTA purifier* (Amersham Biosciences) automated chromatography system. Three ml of proteins were loaded on the column equilibrated with buffer E (10 mM Tris, pH: 7.6, 300 mM KCl) and eluted at a flow rate of 1 ml/min. One ml fractions were collected at the peak area of elution profile shown in Figure 5.6A. Figure 5.6B shows the SDS-PAGE analysis of VceAp corresponding to fractions from peak A on the elution profile. The purity of VceAp polished by gel filtration chromatography reached up to 99%, which make it a good candidate for structural and functional research. The purified VceAp was concentrated to 25 mg/ml using Vivaspin concentrators with molecular weight cut-off point of 20 kDa MWCO, frozen in liquid nitrogen and stored at -80°C.

5.3 Overexpression and Purification of an ECFP-VceAp-EYFP fusion protein

The construct pQE-CApY and *E.coli* strain M15 (pREP4) were used to overproduce an ECFP-VceAp-EYFP (CApY) fusion protein. *ECFP* (primers: ECFP-F/ECFP-R) and *EYFP* (primers: EYFP-F/EYFP-R) were amplified by PCR from plasmids pECFP and pEYFP (Clontech). The resulting *ECFP* fragment bounded by *BamHI* and *SacI*, and the *EYFP* fragment bounded by *SalI* and *HindIII*, were inserted into pQE80L via *BamHI/SacI* and *SalI/HindIII* sites, generating construct pQE-CY. *vceAp* was amplified using primers VceAF/VceAR1 to introduce *SacI* and *SalI* restriction enzyme sites at its two ends and this fragment was inserted into pQE-CY via the *SacI* and *SalI* sites, generating a construct to express a fusion protein with ECFP preceeding the N-terminus of VceAp and EYFP following the C-terminus. The procedure for purification of VceA (53-406) from strain M15 (pREP4)/pQE30-VceA (53-406) was adopted to overexpress and purify CApY. Gel filtration chromatography was used to purify CApY. Figure 5.7A shows the elution profile of CApY and the

SDS-PAGE analysis of CApY fractions corresponding to the peak I of the elution profile are shown in Figure 5.7B.

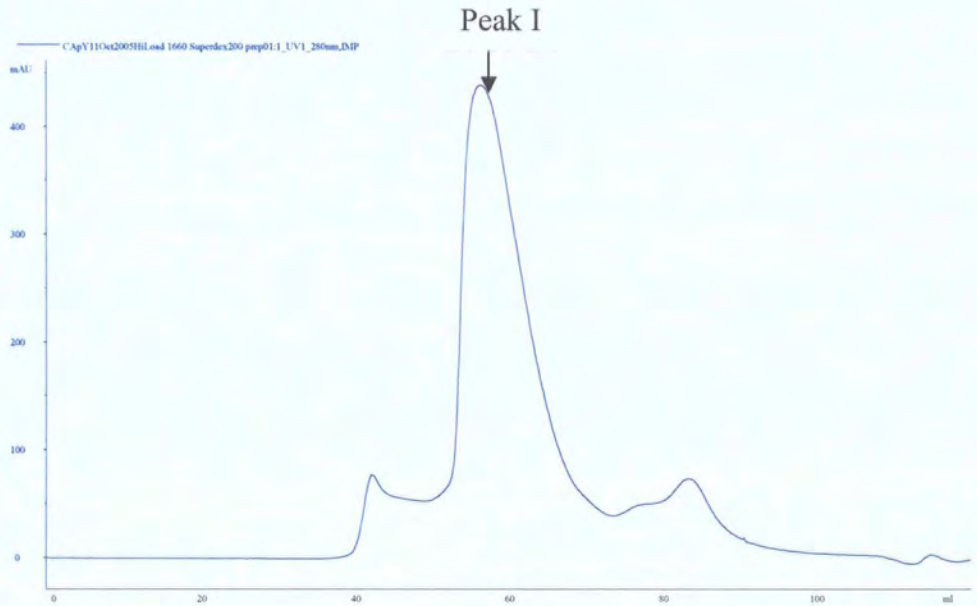


Figure 5.7A. Elution profile of CApY on a HiLoad 16/60 superdex 200 prep grade gel filtration column. Running buffer: 10 mM Tris, pH: 7.6, 300 mM KCl. Flow rate: 1 ml/min.

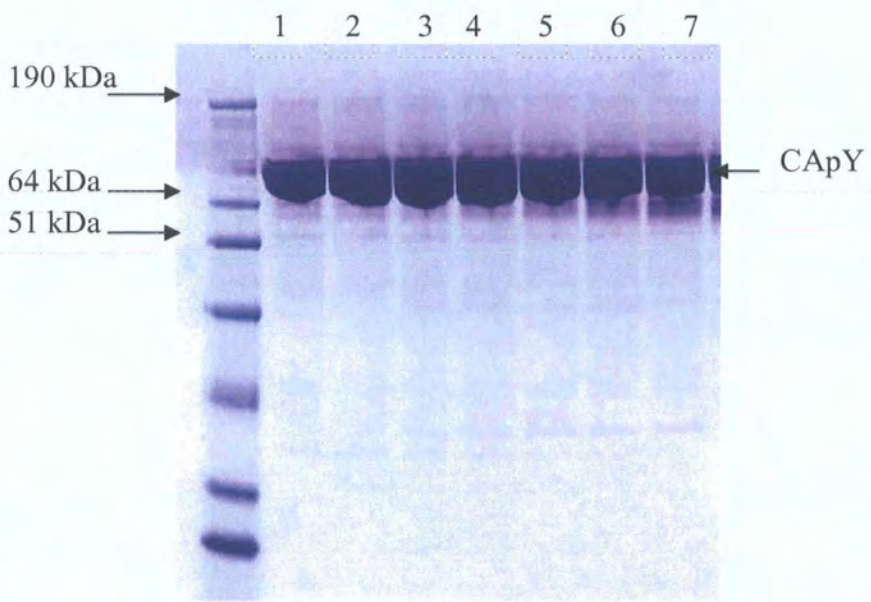


Figure 5.7B. Coomassie blue stained SDS-PAGE gel showing CApY protein polished by size-exclusion charomatography. Lane 1-7 correspond to fractions from peak I in Figure 5.7A.

5.4 Secondary structure prediction of VceA

The *vceA* gene is 1221 bp long encoding a protein of 406 amino acids. We used the secondary structure prediction program PSIPRED (McGuffin *et al.,* 2000) (http://bioinf.cs.ucl.ac.uk/psipred) to analyze the amino acid sequence of VceA. As shown in Figure 5.8, VceA consists of large α-helices sandwiched by β–sheet structures at the N- and C-termini.

We used the program TMPred (http://www.ch.embnet.org/ software/ TMPRED_ form.html) to perform hydropathy analysis of VceA, which indicated that residues 33-52 have a higher probability to form a transmembrane helix.

The program COILS was used to predict the probability of VceA to form a coiled-coil (Lupas *et al.,* 1991). The coiled-coil motif is composed of right-handed α-helices. The amino acid residues in the helix show characteristic heptad repeat $(abcdefg)_n$. The positions **a** and **d** tend to be hydrophobic residues, which form a hydrophobic core with the other coil; whilst the positions **e** and **g** are predominantly charged residues, which form ion pairs with oppositely charged residues on the other coil (Berger *et al.,* 1995). A schematic diagram of coiled-coil motif is shown in Figure 5.9A (Cohen *et al.,* 1990). Figure 5.9B shows the heptad repeat of a coiled-coil in 3D (David Gossard, Protein Folding Problem, MIT's OpenCourseWare, 2003 fall).

Figure 5.8. The secondary structure prediction of VceA using the program PSIPRED.

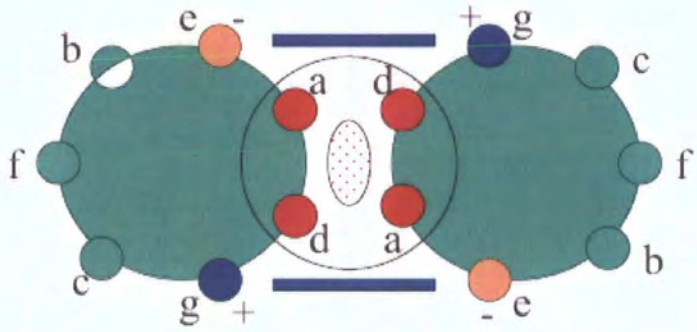
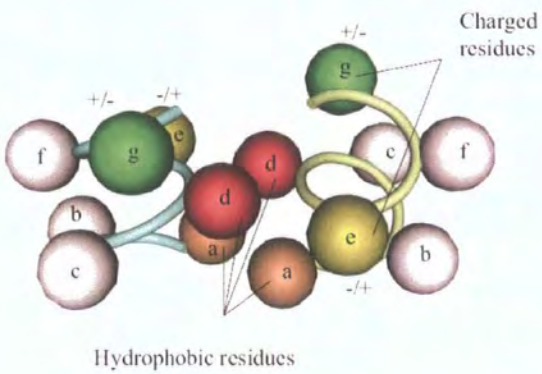


Figure 5.9A. Schematic diagram of coiled-coil motif, showing the hydrophobic residues at positions **a** and **d** and the charged residues at positions **e** and **g** (Adapted from Cohen *et al.*, 1990).



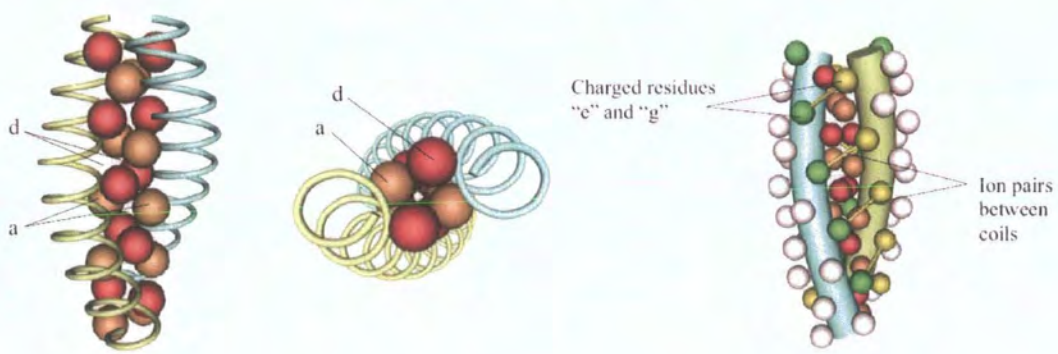


Figure 5.9B. Heptad repeat of coiled-coil in 3D (Figure adapted from David Gossard, Protein Folding Problem, MIT's OpenCourseWare, 2003 fall).

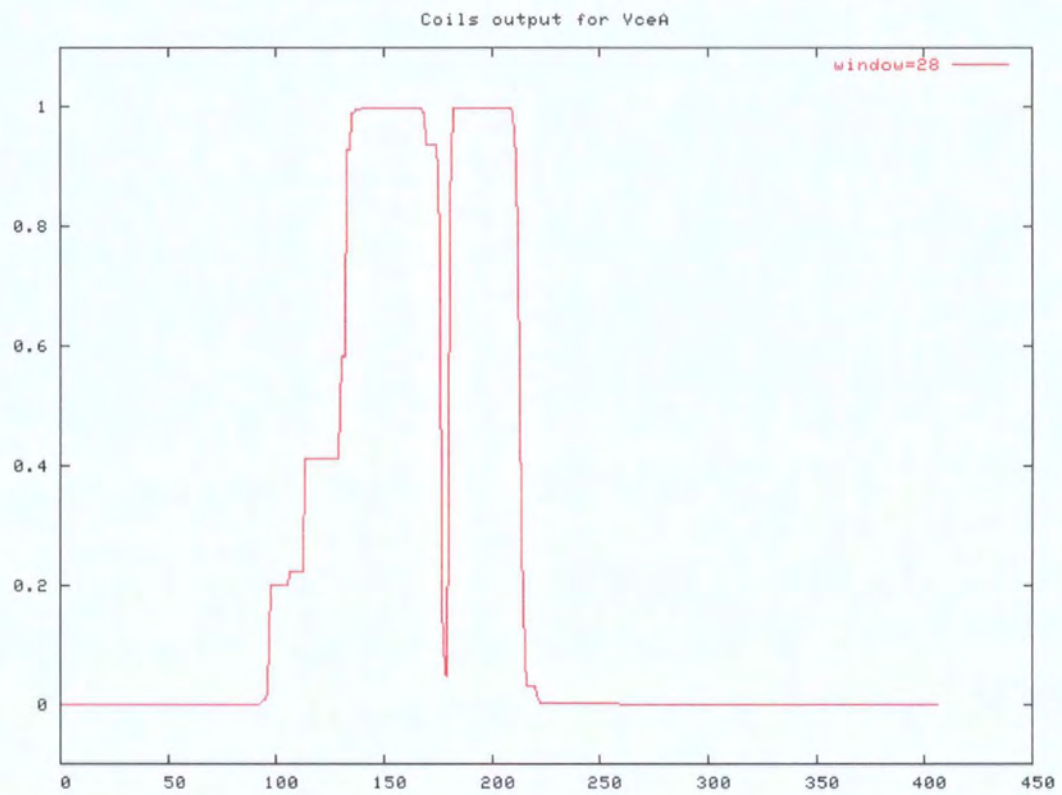


Figure 5.10. The probability of forming coiled-coils as a function of residue number of VceA. The graph was generated by the program COILS with a window of 28 residues.

The two long α-helices in the periplasmic domain of VceA are of similar length. As shown in Figure 5.10, these α-helices were predicted by COILS to have a high probability of forming a coiled-coil. The MULTICOIL program was used to analyze the oligomeric state of these coiled-coils, which indicated a high probability for the α-helices to form a dimeric coiled-coil (Wolf *et al.*, 1997). The prediction results indicated that the two long α-helices of VceA fold back to form a dimeric coiled-coil, with a gap of 10 amino acids between them. The dimeric coiled-coil includes a four heptad repeat.

5.5 VceA is anchored to the inner membrane by a single α-helical domain

The hydropathy analysis indicated that residuces 33-52 of VceA have a high probability of forming a transmembrane helix. To test whether VceA is anchored to the IM via this transmembrane helix, several constructs were made to overexpress the whole and truncated VceA.

Sub-cellular fractionation of cells expressing VceA, VceA (12-406) and VceA (53-406) indicated that VceA and VceA (12-406) were located in the cellular membrane, while VceA (53-406) was located in the soluble cytosol. There is no cysteine amongst the N-terminal 52 residues, ruling out the possibility of a lipid moiety. All these evidences suggested that VceA is anchored to the IM via the transmembrane helix.

The amino acid residues that flank a transmembrane domain of a membrane protein, especially the positively charged residues, are critical for membrane protein biogenesis (Dalbey *et al.*, 2004). Therefore, I did not try to construct a truncated VceA so as to locate the precise position of the transmembrane helix.

5.6 The periplasmic domain of VceA forms a trimer

To gain insight into the oligomeric state of the periplasmic domain of VceA, VceAp was expressed and purified from the periplasm of *E.coli.*

Analytical gel filtration chromatography was used to determinate the molecular weight and Stokes' radius of VceAp using Superdex 200 PC 3.2/30 column. The Blue Dextran 2000 and protein standards of Ribonuclease A, Chymotrypsinogen A, Ovalbumin and Albumin were dissolved in running buffer to prepare the calibration curve (see Materials and Methods). The VceAp sample was diluted to a concentration of 4 mg/ml and applied to the Superdex 200 PC 3.2/30 column equilibrated in running buffer F (20 mM Tris, pH: 7.0, 300 mM NaCl). The flow rate was kept at 20 µl/min. Figure 5.11A shows the elution profile for VceAp.

The Microsoft Excel program was used to prepare a standard curve of $[-\log (kav)]^{1/2}$ versa Stokes' radius and calculate the linear regression equation of the curve. The same data were used to calculate the molecular mass of VceAp using the plot of Kav versus the logarithm of their molecular weights. Figure 5.11 B and C show the standard curves.

The corresponding Kav of VceAp was calculated to be 0.41. Its Stokes' radius was determined from the standard curve to be 40.2 Å, and its molecular weight was determined to be 96.1 kDa. Considering that the monomeric VceAp has a molecular mass deduce from its amino acid sequence of 37.98 kDa, our findings indicate that VceAp forms a trimer.

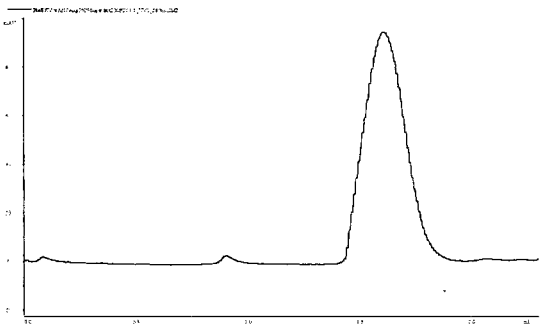


Figure 5.11A. The elution profile of VceAp on a Superdex 200 PC 3.2/30 size-exclusion column. Running buffer: 20 mM Tris, pH: 7.0, 300 mM NaCl. Flow rate: 20 µl/min.

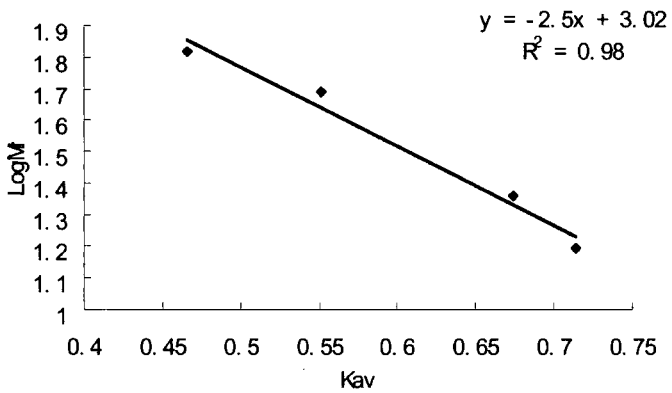


Figure 5.11B. A standard Kav:logMr curve constructed from the elution volumes of the Low Molecular Weight Gel Filtration Kit.

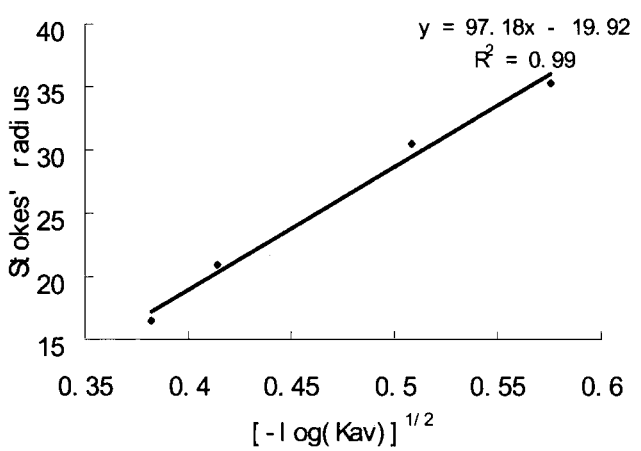


Figure 5.11C. A standard $[-\log (Kav)]^{1/2}$:stokes' radius curve constructed from the elution volumes of the Low Molecular Weight Gel Filtration Kit.

5.7 FRET evidence for the conformational state of VceAp

To gain insight into the conformational state of VceAp, I employed a GFP-based fluorescence resonance energy transfer (FRET) method. FRET is a radiationless transfer of excitation energy from a donor fluorescent molecule to an acceptor by way of dipole-dipole interactions during excitation of the donor molecule. It is a distance-dependent interaction between the donor and acceptor chromophores. The efficiency of energy transfer (E) is proportional to the inverse sixth power of the intermolecular separation, as described by the Förster equation: $E=Ro^6/ (Ro^6+ R^6)$, where R is the donor-acceptor radius and Ro is the radius at which FRET efficiency is 50% (Förster radius). FRET typically occurs in the 10-100 Å region and small changes in R (1–2Å) and small orientation changes between the donor and acceptor fluorophores can dramatically affect the efficiency of FRET. This means that FRET is a good indicator of close proximity and changes in intermolecular distances (Green *et al.*, 2005).

The GFP variants ECFP and EYFP were used as the fluorescence protein pairs for FRET. The ECFP variant has fluorescence excitation maxima at 433 nm (major peak) and 453 nm (minor peak), and the emission maxima at 475 nm with a small shoulder at 501 nm; the EYFP variant has fluorescence excitation maxima at 513 nm, it also can be efficiently excited at 488 nm, and its emission maxima is at 527 nm. These protein pairs have excitation and emission properties favorable for FRET (Clontech). Fig5.12A and B show the emission spectra of ECFP and EYFP.

A spectrofluorometer based method was adopted to detect FRET (Zheng *et al.*, 2004). The experiments were performed in buffer G (20 mM Tris, pH 7.5, 150 mM NaCl) at room temperature. The fluorescence spectra were recorded on a JASCO FP-750 spectrofluorometer using a 10 mm path-length quartz cuvette. The excitation and emission band widths were set up to 5 nm. The fluorescence emissions for ECFP and CApY were monitored between 450 and 550 nm with excitation at 433 nm; whilst the

fluorescence emission for EYFP was monitored between 510 and 560 nm with excitation at 500 nm. By exciting at 433 nm, the emission spectrum of CApY showed two peaks at 475 nm and 527 nm, respectively. Figure 5.12C shows the emission spectrum of CApY protein excited at 433 nm; Figure 5.12D shows the superimposition of the emission spectra of ECFP and CApY proteins excited at 433 nm. It is obvious that the peak at 475 nm corresponds to ECFP and the peak at 527 nm arises mainly from EYFP. The result indicated that strong FRET between ECFP and EYFP existed without substrate induction, suggesting that the N-terminus of VceAp is in close proximity to its C-terminus.

I tried to determine the drug-induced FRET change of CApY. The fusion protein CApY assumed strong FRET, but substrates induced no change in FRET.

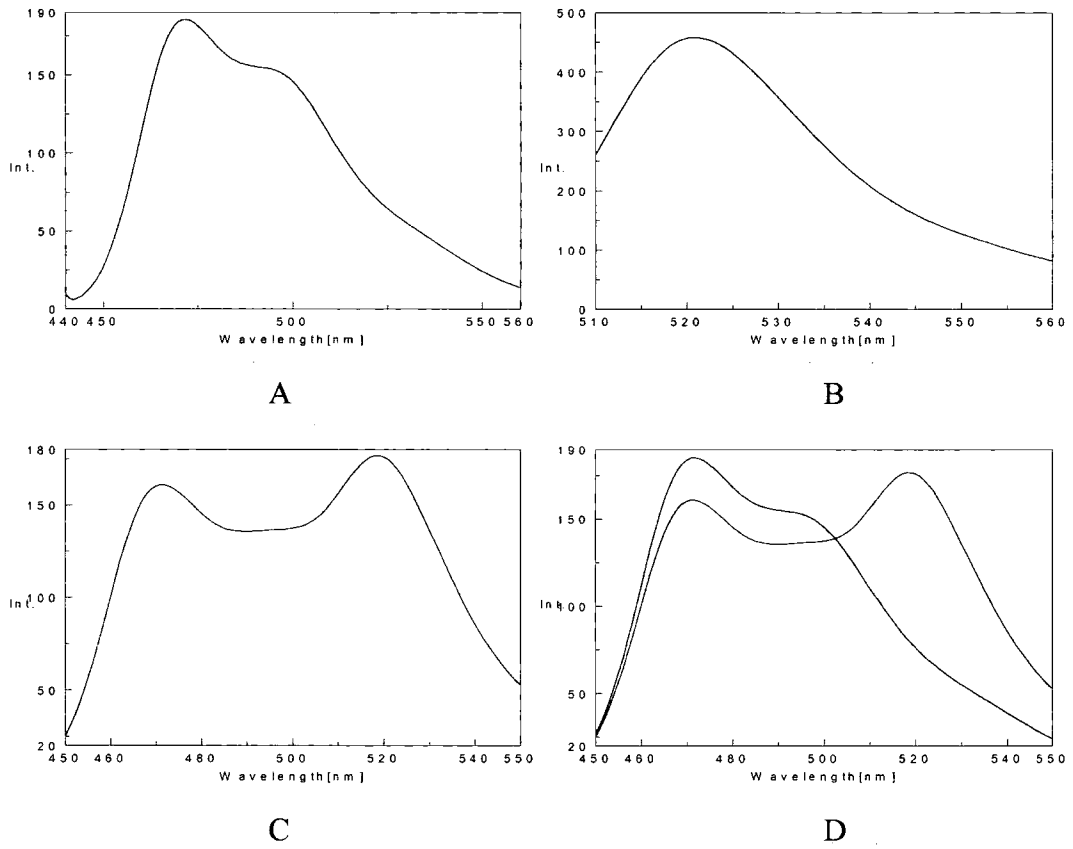


Figure 5.12. (A) Emission spectrum of ECFP (excitation at 433 nm). (B) Emission spectrum of EYFP (excitation at 500 nm). (C) Emission spectrum of CApY (excitation at 433 nm). (D) A superimposition of emission spectra of ECFP and CApY(excitation at 433 nm).

5.8 Discussion

Structural information on membrane fusion proteins is essential to understand their functional role in the tripartite transporters. In this study we provide an initial structural characterisation of VceA, the MFP of the tripartite multidrug extrusion pump VceCAB.

Sub-cellular localization of the whole and truncated VceA indicated that VceA is anchored to the inner membrane by residues 1-52, consistent with the prediction that residues 33-52 form a transmembrane α-helix. As the amino acid residues flank a transmembrane helix, especially the positively charged residues, are critical for membrane protein biogenesis (Dalbey *et al.*, 2004), I did not try to construct a truncated VceA in order to locate the precise position of the transmembrane helix.

Two models of MFP function have been proposed (Figure 5.13). In the first model, the MFP is presumed to fold upon itself using its coiled-coil domain to allow a transient interaction between the IMP and OMP for substrate transfer, with the N-terminus anchored in the inner membrane and C-terminus interacting with the OMP or outer membrane. In the second model, the MFP is supposed to form a channel between the IMP and OMP, allowing direct transfer of the drug (Johnson *et al.*, 1999; Elkins *et al.*, 2003). I used the FRET method to test which of these hypothetical models is reasonable. I fused the ECFP and EYFP to the two termini of the periplasmic domain VceAp. Strong FRET between ECFP and EYFP existed without substrate induction, indicating that the N-terminus of VceAp is in close proximity to its C-terminus. FRET evidence does not support the first model of MFP function which requires the MFP to span the periplasm in the inactive complex so that its N- and C-terminus are not in close proximity. The crystal structures of AcrA and MexA have been solved recently, which provide structural evidence that MFP form a channel between the IMP and OMP.

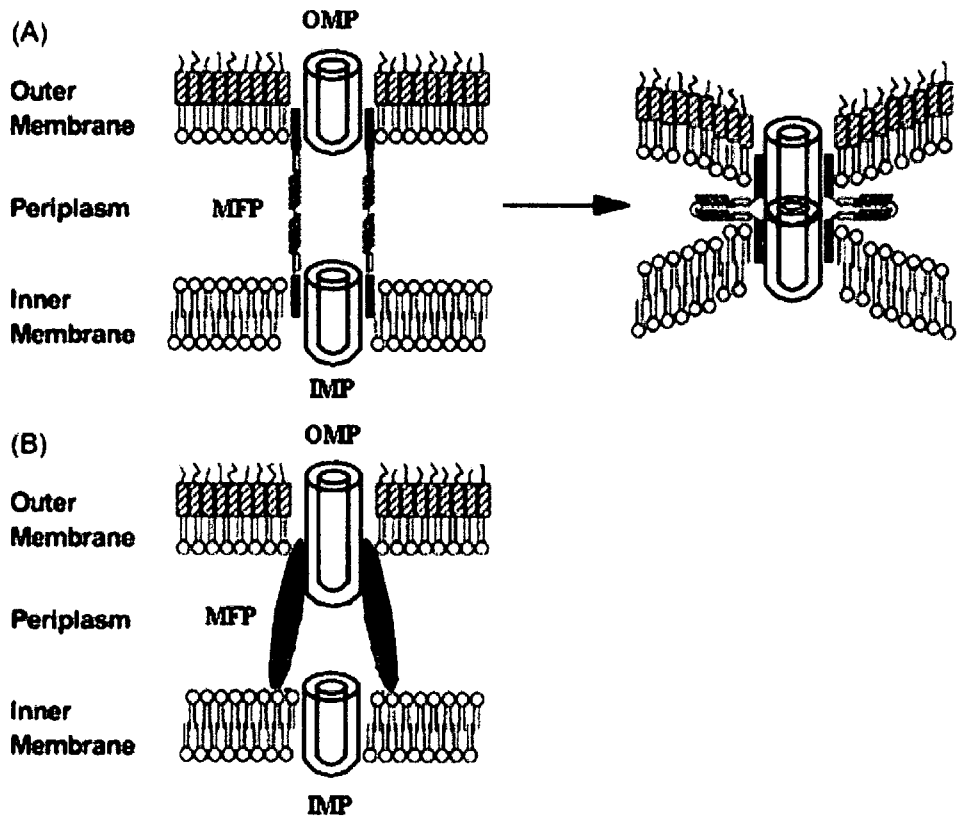


Figure 5.13. Schematic representation of two models of the MFP function. (A) The MFP is postulated to fold upon itself using its coiled-coil domain, allowing a transient interaction between IMP and OMP for substrate transfer. (B) The MFP is supposed to form a channel between IMP and OMP, allowing direct transfer of the drug (Adapted from Elkins *et al.*, 2003).

EmrA has been reported to bind drug molecules directly (Borges-Walmsley *et al.*, 2003). FRET can detect small changes in distance (1–2Å) and orientation between the donor and acceptor fluorophores. I adopted this method to determine if VceA also binds drug molecules using ECFP and EYFP as donor and acceptor fluorophores. The fusion protein CApY assumed strong FRET, but substrates induced no change in FRET. The CApY monomer has a similar molecular weight with the VceAp trimer, and ran on the gel filtration column with similar elution volumes (Figure 5.6A, Figure 5.7), indicating that CApY is a monomer. It is likely that the oligomeric interfaces of VceAp were destabilized by the fusion of the GFP variants, which may account for the inability of VceAp to interact with drugs.

The periplasmic domain of MFP expressed in the cytoplasm can exist as a monomer, dimer and trimer (Borges-Walmsley *et al.*, 2003). I expressed VceAp in the periplasm to test its oligomeric state. By using analytical gel filtration chromatography, the molecular mass of VceAp was calculated to be 96.1 kDa. Considering that the monomeric VceAp has a molecular mass deduce from its amino acid sequence of 37.98 kDa, my findings indicate that VceAp forms a trimer. Cross-linking of *in vivo* complexes using the short-arm chemical cross-linker DSG has identified MFP trimers in both drug efflux and protein export tripartite system (Thanabalu *et al.*, 1998; Zgurskaya *et al.*, 2000). By using a blue native PAGE method, several MFP were also identified to be homo-trimers (Stenberg *et al.*, 2005). These findings suggest that trimeric MFP is its oligomeric state in the tripartite pump, which support the structural model for tripartite multidrug efflux pumps proposed by Fernandez-Recio *et al* (2004).

Chapter 6

Crystallographic analysis of the RND family multidrug transporter MtrD from *Neisseria gonorrhoeae*

The *mtr* (multiple transferable resistance) operon of *Neisseria gonorrhoeae* encodes an energy-dependent tripartite efflux pump, which is composed of the membrane fusion protein MtrC, the inner membrane protein MtrD and the outer membrane protein MtrE. Expression of the *mtrCDE* operon is negatively regulated by the product of the *mtrR* gene, which is divergently transcribed (Hagman *et al.*, 1995). The inner membrane component, MtrD, belongs to the RND family transporter (Hagman *et al.*, 1997). The MtrC/MtrD/MtrE efflux pump confers broad-spectrum resistance to hydrophobic antibiotics, dyes and detergents similar to that of the MexA/MexB/OprM efflux pump of *P. aeruginosa* and AcrA/AcrB/TolC efflux pump of *E.coli*. Besides, the MtrC/MtrD/MtrE pump also exports cationic antimicrobial peptides (CAMPs), such as PG-1 and LL-37 (Shafer *et al.*, 1998).

Multiple alignments of amino acid sequences from multidrug transporters of the RND family have identified four strongly conserved motifs: motifs A, B, C and D (Figure 6.1. Putman *et al.*, 2000). The conservation of these motifs suggests that they play an important structural or functional role in the transporters. MtrD is an exception, which contains motif A only (Paulsen *et al.*, 1996). Phylogenetic analysis revealed that the majority of RND family multidrug efflux proteins fall within a single closely related cluster, with only MtrD being somehow divergent (Paulsen *et al.*, 1996). It is likely that there are structural and/or functional differences between MtrD and other RND family transporter, such as AcrB, MexB and AcrD.

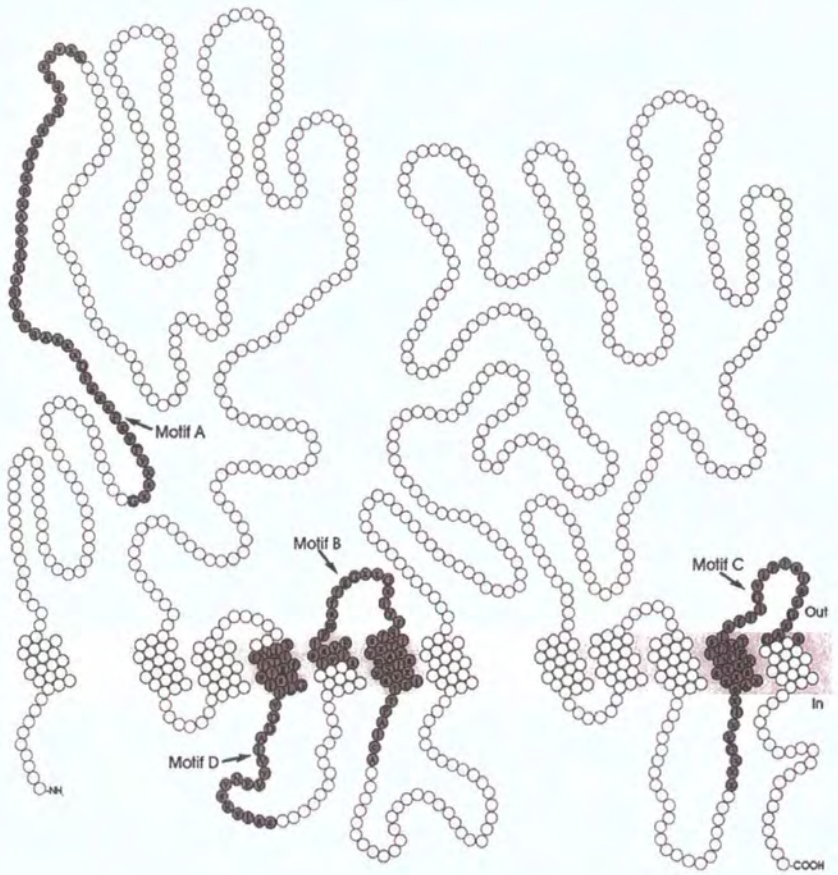


Figure 6.1. Schematic two-dimensional representation of a typical member of the RND family transporters in the cytoplasmic membrane (gray). The locations of the conserved motifs A, B, C, and D are highlighted in black (Putman *et al.*, 2000).

The determination of the structure of AcrB may represent the most significant advances on the understanding of drug efflux pumps in recent years. The knowledge of the difference between MtrD and AcrB will be critical for understanding the mechanism of these family transporters. In this chapter, I report the successful overexpression, purification and crystallization of MtrD, and the optimisation of resolution of MtrD crystals to 7-10 Å.

6.1 Overexpression and purification of MtrD

6.1.1 Recombinant expression of MtrD

The pET21a vector was used to make a recombinant expression construct pET-MtrD. The *mtrD* gene (bounded by *NdeI* and *XhoI*) was inserted into pET21a, resulting in a construct which expresses the *mtrD* gene with a C-terminal hexahistidine tag. The pET-MtrD construct was transformed into the *E.coli* strain C43 (DE3).

A single C43 (DE3)/pET-MtrD colony from a freshly streaked plate was picked and inoculated into a 25ml universal container containing 10 ml of 2×YT broth and Carbenicillin (100μg/ml). The cells were grown in an orbital shaker at 37°C, 220 rpm. overnight. 1 liter of prewarmed 2×YT broth (with antibiotic) in a 2 liter baffled flask was inoculated with 2 ml of the overnight culture. The cells were grown in an orbital shaker at 37°C, 220 rpm, and their growth was monitored by measuring the absorbance of the culture at 600 nm (A_{600}).

The cells were induced at mid log phase (approx. A_{600}=0.5-0.6) with 0.5mM IPTG. The temperature was dropped to 25°C and growth was continued with shaking at 200 r.p.m. for 4-5 hours. The cells were harvested by centrifugation at 4,350 ×g for 8 min at 4 °C, resuspended in buffer A (20mM Tris pH: 7.0, 20% Glycerol, 5 U/ml DNase I, 1 tablet /100ml EDTA-free protease inhibitor mixture tablet) and stored at −80°C or purification continued.

6.1.2 Purification of the His-tagged MtrD using Ni^{2+} affinity chromatography

Purification of MtrD by Ni^{2+} affinity chromatography was performed using a HiTrap Chelating HP 1 ml column.

The cells were lysed by three passages through a Constant Cell Disruptor System and the cell debris was removed by centrifugation at 24,000 ×g for 30 min at 4 °C. The supernatant, containing soluble proteins and membrane vesicles, was carefully decanted so as not to disturb the pelleted material. The supernatant was further fractionated into membrane and soluble components by high-speed ultracentrifugation at 122,400 ×g for 90 min at 4°C.

The membrane pellet was emulsified in a small volume of buffer B (20mM Tris pH: 7.0, 20% Glycerol, 0.5mM THP) by repeated passage through a syringe needle. The emulsified membrane was then diluted by buffer B to a final concentration of about 10 mg/ml of total protein. Imidazole (5 M, pH: 7.0) was added to the emulsified membrane to a final concentration of 15mM, NaCl was added to 100mM, and the pH of the mixture was adjusted to 7.0 using 50 mM NaOH. A detergent powder of the *n*-Dodecyl-ß-D-maltoside (DDM) was added to the emulsified membrane to a final concentration of 2% to solubilise the membrane. The mixture was gently stirred at 4 °C for 45 min and then ultracentrifuged at 122,400 ×g for 1 hour at 4 °C. The membrane solution was carefully decanted so as not to disturb the pelleted material.

Extracted histidine-tagged MtrD was purified using a HiTrap chelating HP 1 ml column immobilized with Ni^{2+} equilibrated with buffer C (20 mM Tris pH: 7.0, 20% Glycerol, 0.2% DDM). The membrane solution was loaded onto the column using a peristaltic pump at a speed of 0.8 ml/min. The column was washed with 10 ml of 50 mM imidazole in buffer C, and then with 20 ml of 100mM imidazole in buffer D (20 mM Tris pH: 7.5, 300 mM NaCl, 20% Glycerol, 0.2% DDM). The MtrD was eluted from the column with 400 mM imidazole in buffer D. The eluates were collected at

0.5 ml fractions and analyzed by SDS-PAGE. MtrD is composed of 1068 amino acids and the predicted molecular weight is about 114 kDa. Figure 6.2 illustrates an SDS-PAGE analysis of the MtrD.

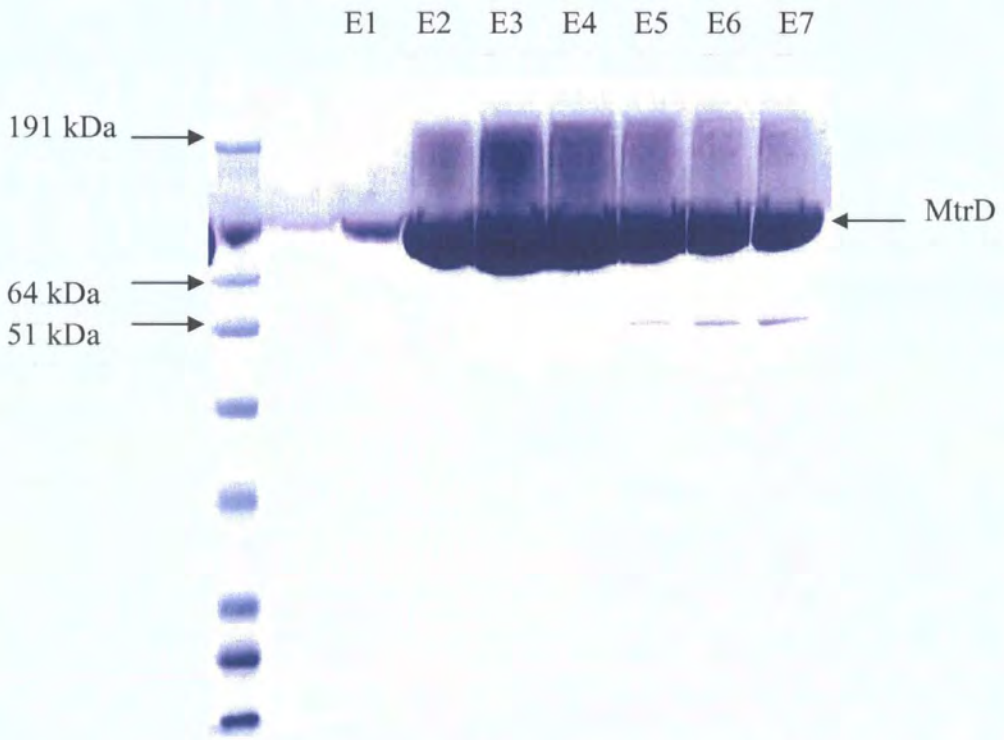


Figure 6.2. Coomassie blue stained SDS-PAGE gel showing IMAC purification of MtrD-His_6 protein from DDM solubilised membranes. Lane E1-E7 show elution fractions (0.5×7 ml) with buffer containing 400 mM imidazole. The left arrows mark the positions of protein standards; the right arrow marks the position of MtrD-His_6 protein.

6.1.3 Protein buffer exchange and concentration

The purity of the MtrD protein purified by Ni^{2+} affinity chromatography is about 98% of total proteins, as indicated by the SDS-PAGE analysis in Figure 6.2. Buffer exchange of MtrD into buffer E (10 mM Tris, pH: 7.5, 50 mM NaCl, 10% Glycerol, 0.1% DDM) was carried out on a HiTrap Desalting column. The protein was then concentrated using a RESOURCE Q 6 ml anion exchange chromatography column. The column was washed with 5 column volumes (CV) of distilled water and 5 CV of buffer F (10 mM Tris-HCl, pH: 7.5, 500 mM NaCl, 10% Glycerol, 0.1% DDM), then equilibrated with 5 CV of buffer E.

The eluate from the HiTrap Desalting column was loaded onto the RESOURCE Q column equilibrated with buffer E at a flow rate of 2 ml/min. The column was washed with 2 CV of the same buffer, and then eluted in 0.5 ml fractions with 400 mM NaCl in buffer E at a flow rate of 1 ml/min. During this step the MtrD protein was concentrated to 2-3 mg/ml. Figure 6.3 shows the SDS-PAGE analysis of MtrD concentrated by ion exchange chromatography.

The buffer of MtrD was exchanged again into buffer G (10 mM Tris-HCl, pH: 7.5, 100 mM NaCl, 10% Glycerol, 0.1% DDM) using the HiTrap Desalting column, and the eluate was dialyzed against buffer G using a Slide-A-Lyzer 10 kDa MWCO Dialysis Cassettes at 4°C for 3-4 hours. The MtrD protein was then concentrated to 25-30 mg/ml using VIVASPIN Concentrators with 100 kDa MWCO. The protein was used to set up crystallization trials or frozen in liquid nitrogen and stored at −80°C.

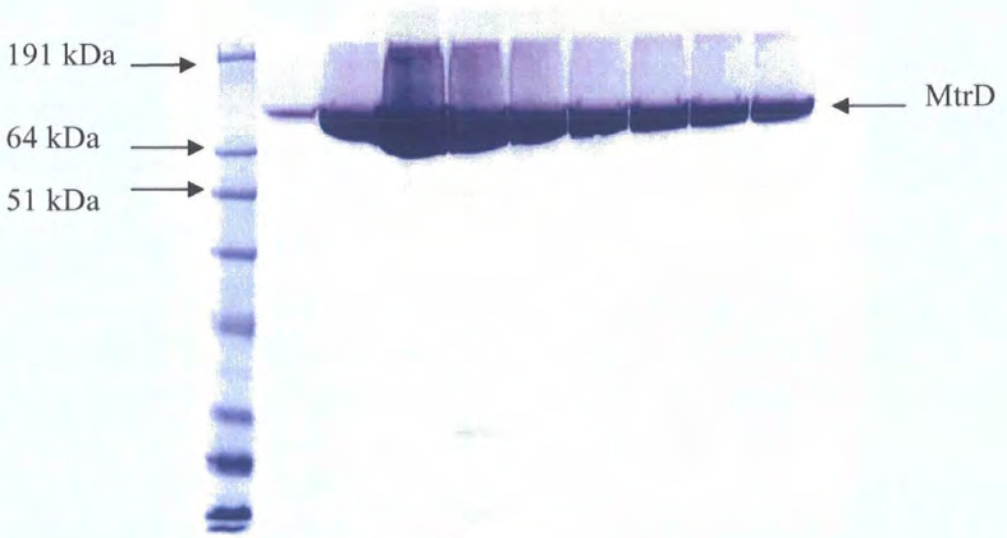


Figure 6.3. The SDS-PAGE analysis of MtrD concentrated by anion exchange chromatography.

6.2 Crystallization and data collection

MtrD purified in DDM was subjected to crystallization screens using the membrane protein crystal growth screens kits MbClass and MbClass II Suites (Nextal Biotechnologies). Crystallization trials were set up using the hanging drop vapour diffusion and sitting drop vapour diffusion methods.

Initially, I directly concentrated MtrD to 10-100× times using the Vivaspin Concentrators with 100 kDa MWCO. The DDM micellar aggregations appeared in some of the crystallization drops, indicative of a high DDM concentration (Figure 6.4). The DDM may have been concentrated because of its large micelle size, which may result in the inability to crystallize the protein. Then I changed strategy to concentrate the protein with two steps: the MtrD protein was concentrated to 2-3 mg/ml using an anion exchange chromatography column, then it was farther concentrated to 10× times using the Vivaspin Concentrators with 100 kDa MWCO, and we obtained MtrD crystals.

Crystals were obtained from the hanging drop vapour diffusion method by mixing 1 µl of protein sample on a silinated Crystallization Support (Nextal Biotechnologies) with 1 µl of a reservoir solution buffer H (30% PEG 400 v/v, 100 mM MES pH 6.5). The Crystallization Support was sealed on a well of the Crystallization Tool (Nextal Biotechnologies) containing 1.0 ml of the same reservoir solution and equilibrated at 18°C. Crystals appeared after two weeks with initial MtrD crystals growing as rectangular rods, finally reaching a size of 0.1x0.05x0.05 mm^3 after one month (Figure 6.5a). The crystals diffracted X-rays to 15-20 Å. However, this crystallization condition was less reproducible. The additives were screened using the Additive Screen™ Kits (Hampton Research). Improvements in the reproducibility were made with several additives. The addition of 1, 6-Hexanediol (3%), Jeffamine M-600 (5%, pH: 7.0), Heptane-1,2,3-triol (1.5%) and Polypropylene Glycol P400 (4%) in the drops improved the reproducibility of crystallization condition, but the crystals grew as needles, thin sheets or microcrystals (Figure 6.5b).

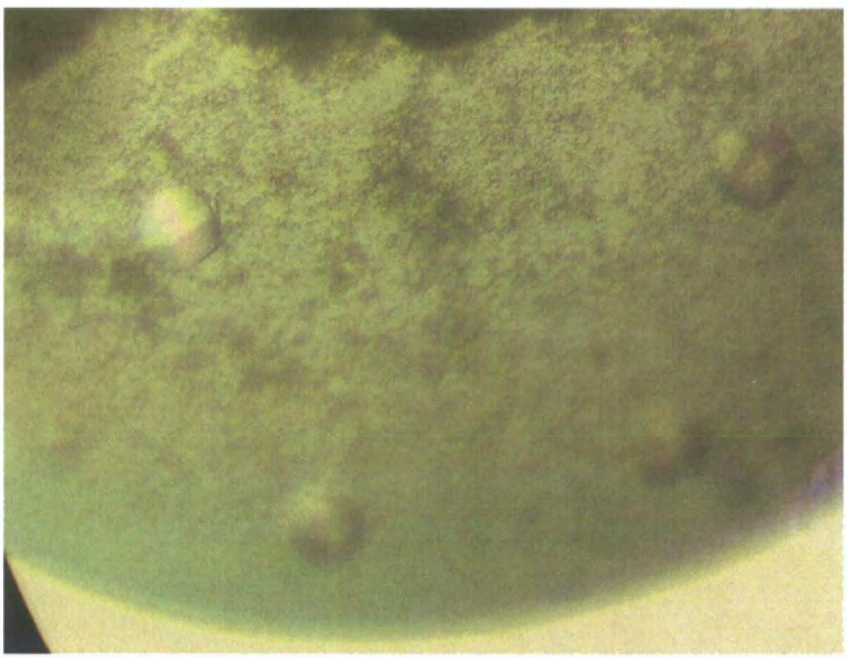


Figure 6.4. The DDM micellar aggregations with birefringence resembling crystalline birefringence.

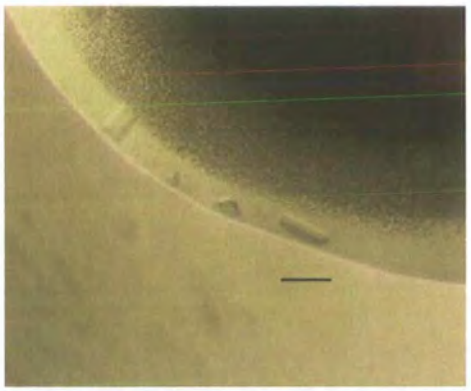


Figure 6.5a. A picture of the MtrD crystals. The reservoir solution contained 30% PEG 400 v/v, 100 mM MES pH 6.5. The bar represents about 0.1 mm.

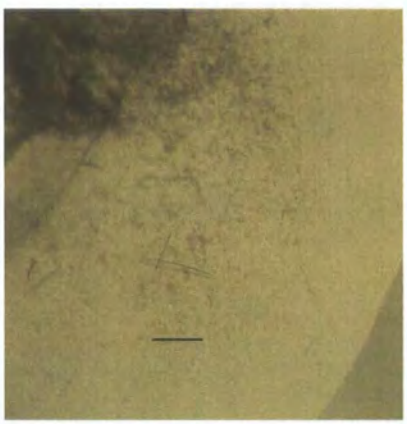


Figure 6.5b. A picture of the MtrD crystals. The reservoir solution contained 30% PEG 550 MME v/v, 100 mM HEPES pH 7.5. The bar represents about 0.1 mm.

Figure 6.5c. A picture of the MtrD crystals. The reservoir solution contained 30% PEG 400 v/v, 100 mM MES pH 6.5. The ratio of MtrD/reservoir solution/CYMAL-6 in the drop is 1:1:1. The bar represents about 0.4 mm.

The crystallization conditions were optimized with a series of screens that consisted of PEG or its MME derivative as the precipitant: PEG 400, PEG 550MME, PEG 1500, PEG 2000MME or PEG 4000 and pH from 6.5 to 8.0 at 0.5 unit increments. The precipitant concentrations in the PEG/pH screening kits ranged from 10% to 30% for PEG 400, PEG 550MME, PEG 1500 and PEG 2000MME in steps of 5%, and 4% to 12% for PEG 4000 in steps of 2%. The crystals were obtained using PEG 550 MME (30% PEG 550 MME, 100mM HEPES pH 7.5) and PEG 4000 (6% PEG 4000, 100mM HEPES pH 7.5) as precipitants. The crystals also grew as needles and thin sheets (Figure 6.5b).

Several detergents, such as cyclohexyl-*n*-hexyl-β-D-maltoside (CYMAL-6), *n*-decanoylsucrose, *n*-nonyl-β-D-thiomaltoside or *n*-octyl-β-D-thiomaltoside (at 1.9mM , 8.3mM, 10.7mM and 30mM final concentration, respectively), were also tested in the crystallization trials to improve the diffraction quality of the crystals. The best results were obtained by the hanging drop vapour diffusion method under the following condition: a droplet made of 1 µl of protein solution, 1 µl reservoir solution and 1 µl of CYMAL-6 (5.6mM) was equilibrated against a reservoir solution buffer H. Crystals appeared after 7-10 days and reached their final size of 0.4x0.15x0.15 mm^3 after three weeks (Figure 6.5c). The crystals diffracted X-rays up to 7-10 Å.

The MtrD crystals were soaked in cryo-protectant buffer (30% PEG 400 v/v, 100mM MES pH 6.5, 100mM NaCl, 10% Glycerol v/v, 0.1% DDM) for about one minute and frozen in liquid nitrogen. The diffraction data of MtrD crystals were collected by Dr. Rick Lewis at the ID14-4 beamline of ESRF synchrotron source (Grenoble).

6.3 Discussion

Membrane protein crystallography is still a challenge. To date, over 30,000 soluble protein structures have been determined; however, about 100 membrane protein

structures have been solved.

The integral membrane proteins are extracted from lipid bilayers using detergents. Above the critical micelles concentration (CMC), detergent monomers aggregate to form micelles, into which membrane proteins can insert. Lipids typically co-purify with membrane proteins, with complete delipidation of membrane proteins often causing denaturation and aggregation. Therefore, membrane proteins are solubilised in aqueous solution in the form of a protein/detergent/lipid complex.

It is critical for a membrane protein to crystallize that it is in a homogeneous state. Lipids can stabilize the membrane protein during purification; however, it results in the heterogeneity of the protein-lipid complex (Wang *et al.*, 2003; Lemieux *et al.*, 2003). The degrees of delipidation of membrane proteins are different after each chromatography purification step. Membrane proteins are usually purified by less than two chromatography steps. Purification with a third chromatography step may result in protein aggregation due to complete delipidation, especially for polytopic membrane proteins (Boulter *et al.*, 2001; Lemieux *et al.*, 2002).

Initially, we developed a procedure to purify MtrD by two chromatography steps: a Ni^{2+} affinity chromatography step followed by an anion exchange chromatography step. The protein was stable and the purity was over 99% of the total protein, but it could not be crystallized. We therefore developed a procedure of one step purification of the protein by Ni^{2+} affinity chromatography. The protein purity was less then that purified by the two chromatography steps, but it was crystallized easily. It is likely that the number of phospholipids molecules co-purified with MtrD is critical for its crystallization.

The presence of unbound detergent in solution is another source of the heterogeneity of membrane protein, when in excess, it may prevent proper packing of protein-detergent micelles in the crystal lattice (Scarborough, 1994; Wang *et al.*, 2003;

Lemieux *et al.,* 2003). We adopted the Vivaspin Concentrators with 100 kDa MWCO to one step concentrate MtrD. Because of the large micelles size, DDM may have been concentrated simultaneously. The DDM micellar aggregation appeared in some of the crystallization drops, resulting in the inability to obtain MtrD crystals. We therefore adopted two steps to concentrate MtrD: first, MtrD was concentrated to 2-3 mg/ml by anion exchange chromatography, and then further concentrated to 25-30 mg/ml using the Vivaspin Concentrators. No DDM micellar aggregation appeared under the same conditions and we obtained MtrD crystals (Scarborough, 1994). The reproducibility of the crystallization trials was greatly improved by the addition of Heptane-1, 2, 3-triol in the drops, which may be because Heptane-1, 2, 3-triol reduces the size of detergent micelles so that improves the packing of protein-detergent micelles into the crystal lattice.

The mixture of DDM and CYMAL-6 in the crystallization drops improved the resolution of the MtrD crystals to 7-10 Å. The mixture of detergent improved the crystal packing, which presumably was due to the more favourable packing or newly exposed protein surface areas for crystal contacts (Koronakis *et al.,* 2000; Huang *et al.,* 2003; Wang *et al.,* 2003).

Chapter 7

Final Discussion

In bacteria, multidrug efflux systems have been identified as significant determinants of resistance. These resistance pumps are widely distributed in bacterial species and many pathogenic bacteria possess them, which play an important role in their intrinsic and acquired multidrug resistances (Poole, 2004). The RND and MATE family transporters have also been shown to be involved in the pathogenicity of bacteria (Piddock *et al.*, 2006). Knowledge of the structure and mechanism of these transporter proteins would be exceedingly useful in the design of inhibitors. In recent years, the structural determination of several multidrug efflux pumps and their regulators, such as AcrB, EmrD, Sav1866, TolC, MexA, AcrA and QacR, etc., probably represent the most significant progress in understanding the structural basis of mutidrug resistance in bacteria (See Chapter I).

To investigate the role of the VceCAB pump in the multidrug resistance of *Vibrio cholerae*, I have characterized functionally and structurally the three components of the VceCAB pump and the regulator protein VceR.

The crystal structure of VceC, the outer membrane component of the VceCAB pump, was determined at 1.8 Å resolution. VceC shares the same overall architecture as TolC and OprM (Koronakis *et al.*, 2000; Akama *et al.*, 2004), consisting of three domains: the β-domain, the α-domain and the equatorial domain, despite the very low degree of sequence identity between them (8.3%). The conserved residues of the three proteins are mainly hydrophobic and are likely to play a structural rather than a functional role.

The residues that maintain the closed state of the channel of TolC and VceC are completely different. In TolC, the intra-monomer hydrogen bonds between residues Asp^{153}-Try^{382} and Gln^{136}-Glu^{359}, as well as the inter-monomer salt bridge interaction

between Asp^{153} of a monomer and Arg^{367} of the adjacent one, are shown to stabilize the closed state (Andersen *et al.,* 2002); while VceC is likely stabilized in its closed conformation by the inter-monomer hydrogen bonds between residues Gln^{236}-$Ser^{449}$and Ser^{243}-Gln^{453}, as well as an intra-monomer hydrogen bond between residues Tyr^{433}-Glu^{229} and an intra-monomer salt bridge interaction between Arg^{455}-Glu^{247}. It seems that different members of these family proteins reach their close conformation through different subsets of interactions.

The channel interiors of VceC, TolC and OprM are generally electronegative (Koronakis *et al.,* 2000; Akama *et al.,* 2004). In addition, VceC contains two rings of clusters of negative charge. The ring made by residues Glu^{397} and Glu^{303} is conserved in OprM, but is not present in TolC. The VceCAB pump partly loses it activity when these two residues of VceC were replaced with alanine, indicating their functional role during transport. However, it will be necessary to rule out the possibility that the mutations of these residues lower the expression level of VceC so affecting the function of the VceCAB pump.

In the resting state of the proteins, the α-barrel of VceC, TolC and OprM at the periplasmic ends are all closed; the β-barrel region of VceC and OprM are also closed, whereas the β-barrel region of TolC is widely open to the extracellular medium. Obviously, both ends of VceC and OprM need to be opened during transport. The interaction of the OMP with its partner IMP and MFP could induce opening of the α-barrel closures. The opening mechanism of the β-barrel closure is not clear yet. It is noticed that the intrinsic multidrug resistance of *P. aeruginosa* is compromised in a *tonB* deletion strain and that the operation of MexAB-OprM is facilitated by the TonB protein (Zhao *et al.,* 1998). It is likely that drug efflux *via* these pumps require a TonB or a TonB-like protein for the operation of the outer membrane components, which may facilitate the opening of the β-barrel closures of VceC and OprM.

The optimal desolvation area (ODA) on the surface of VceC was calculated, which is different from that of TolC (Federici *et al.*, 2005), indicating distinct architectures of VceC-based and TolC-based tripartite pumps. TolC is the outer membrane component of different tripartite pumps in *E.coli*, such as EmrAB-TolC, AcrAB-TolC, MacAB-TolC and HlyDB-TolC systems. The genes encoding TolC and other components are not present in the same operons. The inner membrane component EmrB is a MF family transporter which functions as a monomer (Lomovskaya *et al.*, 1992); AcrB belongs to the RND family transporter that functions as trimer (Murakami *et al.*, 2002); while MacB and HlyB belong to ABC family transporters which function as dimers (Thanabalu *et al.*, 1998; Kobayashi *et al.*, 2003). AcrB and MacB have an extensive periplasmic domain, whereas EmrB and HlyB lack it. These pumps must have a different organization and use different types of protein-to-protein interactions in the tripartite pumps, and the surface of TolC must match these distinct architectures. The genes encoding VceR, VceC, VceA and VceB are present in the same *vce* operon, indicating that VceC is the outer membrane component of the VceCAB system only. This might explain the difference in ODAs between VceC and TolC.

MFPs are anchored to the inner membrane by a single transmembrane helix, or via fatty acids attached to its N-terminal cysteine residue (Borges-Walmsley *et al.*, 2003; Yoneyama *et al.*, 2000). The function of the MFP is determined mainly by the periplasmic domain (Yoneyama *et al.*, 2000). Sub-cellular fractionation of cells expressing full length and truncated VceA suggested that VceA is anchored to the IM *via* a transmembrane helix. I expressed the periplasmic domain of VceA (VceAp) in the periplasm of *E.coli*. Analytical gel filtration chromatography experiments indicated that VceAp is a trimer, which could represent its oligomeric state in the VceCAB pump, and thus the VceC:VceA:VceB ratio in the assembled tripartite pump could be 3:3:1.

The crystal structures of the IMP-MFP-OMP complex or the complex of either of the

two components will be important to understand the molecular mechanism of these pumps. The fact that the three components of tripartite pumps are easy to dissociate *in vitro*, make it difficult to co-crystallize them. I tried to overproduce and crystallize VceB-VceA and VceC-VceA fusion proteins. The VceB-VceA fusion protein was successfully overexpressed. It is still necessary to determine if this fusion protein functions as wild type VceB and VceA. I tried to express a fusion of the periplasmic domain of VceA to the C-terminus of VceC, but I failed to produce this fusion protein (data not shown). Some amino acid residues at the C-terminus of bacterial outer membrane proteins, such as phenylalanine, have been shown to have important roles in protein folding, oligomerization and outer membrane incorporation (de Cock *et al.*, 1997). This may also be the case for VceC. The fusion of VceA to the C-terminus of VceC could impair the efficient and correct assembly of VceC into the outer membrane, resulting in failure to overproduce it.

Analytical gel filtration and DLS experiments indicated that the basic functional unit of VceR is a dimer. Stopped-flow fluoresce spectroscopy experiment revealed that the binding of CCCP to VceR occurs with a Hill coefficient of about four (Borges-Walmsley *et al.*, 2005). These evidences suggested that four CCCP molecules bind to dimeric VceR, and thus each VceR subunit binds at least two CCCP molecules. The basic functional unit of QacR is also a dimer. In QacR, each subunit can bind two different drug molecules, but it binds only one molecule of the same drug. Moreover, the binding of drug to QacR induces a coil-to-helix transition, which shift the position of the C-terminus of that subunit such that it blocks the entrance of the second drug-binding pocket within the same dimer (Schumacher *et al.*, 2001; Grkovic *et al.*, 2002; Schumacher *et al.*, 2004). Therefore, only one QacR subunit in the dimeric QacR binds drug. Another TetR/CamR family transcriptional regulator, TetR, was shown to bind two tetracycline molecules per dimer (Hinrichs *et al.*, 1994). The stoichiometry of drug/VceR-subunit is different from that of other transcriptional regulators in the TetR/CamR family.

Multiple alignments and phylogenetic analysis of the amino acid sequences of RND family transporters reveal that there are differences between MtrD and other RND family multidrug efflux pumps. Knowledge of these differences will be important for understanding the mechanism of this family of transporters. In this study, we have successfully overexpressed, purified and crystallized MtrD. The resolution of MtrD crystals was optimised to 7-10 Å at present. Different conditions, such as salts, temperatures, additives, precipitates, and detergents, still require optimisation in order to improve the crystal packing.

References

Abramson J, Iwata S, Kaback HR: **Lactose permease as a paradigm for membrane transport proteins (Review)**. *Mol Membr Biol* 2004, **21**:227-236.

Abramson J, Kaback HR, Iwata S: **Structural comparison of lactose permease and the glycerol-3-phosphate antiporter: members of the major facilitator superfamily**. *Curr Opin Struct Biol* 2004, **14**:413-419.

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

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

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

Akama H, Kanemaki M, Yoshimura M, Tsukihara T, Kashiwagi T, Yoneyama H, Narita S, Nakagawa A, Nakae T: **Crystal structure of the drug discharge outer membrane protein, OprM, of *Pseudomonas aeruginosa*: dual modes of membrane anchoring and occluded cavity end**. *J Biol Chem* 2004, **279**:52816-52819.

Aires JR, Nikaido H: **Aminoglycosides are captured from both periplasm and cytoplasm by the AcrD multidrug efflux transporter of *Escherichia coli.*** *J Bacteriol* 2005, 187:1923-1929.

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

Andersen C, Koronakis E, Bokma E, Eswaran J, Humphreys D, Hughes C, Koronakis V: **Transition to the open state of the TolC periplasmic tunnel entrance.** *Proc Natl Acad Sci U S A* 2002, **99**:11103-11108.

Baranwal S, Dey K, Ramamurthy T, Nair GB, Kundu M: **Role of active efflux in association with target gene mutations in fluoroquinolone resistance in clinical isolates of *Vibrio cholerae*.** *Antimicrob Agents Chemother* 2002, **46**:2676-2678.

Barrasa MI, Tercero JA, Lacalle RA, Jimenez A: **The ard1 gene from *Streptomyces capreolus* encodes a polypeptide of the ABC-transporters superfamily which confers resistance to the aminonucleoside antibiotic A201A.** *Eur J Biochem* 1995, **228**:562-569.

Begum A, Rahman MM, Ogawa W, Mizushima T, Kuroda T, Tsuchiya T: **Gene Cloning and Characterization of Four MATE Family Multidrug Efflux Pumps from *Vibrio cholerae* Non-O1.** *Microbiol Immunol* 2005, **49**:949-957.

Berger B, Wilson DB, Wolf E, Tonchev T, Milla M, Kim PS: **Predicting coiled coils by use of pairwise residue correlations.** *Proc Natl Acad Sci U S A* 1995, **92**:8259-8263.

Biswas GD, Anderson JE, Sparling PF: **Cloning and functional characterization of *Neisseria gonorrhoeae* tonB, exbB and exbD genes.** *Mol Microbiol* 1997, **24**:169-179.

Bokma E, Koronakis E, Lobedanz S, Hughes C, Koronakis V: **Directed evolution of a bacterial efflux pump: adaptation of the *E. coli* TolC exit duct to the *Pseudomonas* MexAB translocase.** *FEBS Lett.* 2006, **580**:5339-43.

Borges-Walmsley MI, Beauchamp J, Kelly SM, Jumel K, Candlish D, Harding SE, Price NC, Walmsley AR: **Identification of oligomerization and drug-binding domains of the membrane fusion protein EmrA.** *J Biol Chem* 2003, **278**:12903-12912.

Borges-Walmsley MI, Du D, McKeegan KS, Sharples GJ, Walmsley AR: **VceR regulates the vceCAB drug efflux pump operon of *Vibrio cholerae* by alternating between mutually exclusive conformations that bind either drugs or promoter DNA.** *J Mol Biol* 2005, **349**:387-400.

Borges-Walmsley MI, McKeegan KS, Walmsley AR: **Structure and function of efflux pumps that confer resistance to drugs.** *Biochem J* 2003, **376**:313-338.

Boulter JM, Wang DN: **Purification and characterization of human erythrocyte glucose transporter in decylmaltoside detergent solution.** *Protein Expr Purif* 2001, **22**:337-348.

Brown MH, Paulsen IT, Skurray RA: **The multidrug efflux protein NorM is a prototype of a new family of transporters.** *Mol Microbiol* 1999, **31**:394-395.

Brunger AT, Adams PD, Clore GM, DeLano WL, Gros P, Grosse-Kunstleve RW, Jiang JS, Kuszewski J, Nilges M, Pannu NS, et al.: **Crystallography & NMR system: A new software suite for macromolecular structure determination.** *Acta Crystallogr D Biol Crystallogr* 1998, **54 (Pt 5)**:905-921.

Bussiere DE, Muchmore SW, Dealwis CG, Schluckebier G, Nienaber VL, Edalji RP, Walter KA, Ladror US, Holzman TF, Abad-Zapatero C: **Crystal structure of ErmC', an rRNA methyltransferase which mediates antibiotic resistance in bacteria.** *Biochemistry* 1998, **37**:7103-7112.

Butler PJ, Ubarretxena-Belandia I, Warne T, Tate CG: **The *Escherichia coli* multidrug transporter EmrE is a dimer in the detergent-solubilised state.** *J Mol Biol* 2004, **340**:797-808.

Careaga CL, Falke JJ: **Thermal motions of surface alpha-helices in the D-galactose chemosensory receptor. Detection by disulfide trapping.** *J Mol Biol* 1992, **226**:1219-1235.

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

Chen J, Lu G, Lin J, Davidson AL, Quiocho FA: **A tweezers-like motion of the ATP-binding cassette dimer in an ABC transport cycle.** *Mol Cell* 2003, **12**:651-661.

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

Choudhury P, Kumar R: **Association of metal tolerance with multiple antibiotic resistance of enteropathogenic organisms isolated from coastal region of deltaic Sunderbans.** *Indian J Med Res* 1996, **104**:148-151.

Cohen C, Parry DA: **Alpha-helical coiled coils and bundles: how to design an alpha-helical protein.** *Proteins* 1990, **7**:1-15.

Colmer JA, Fralick JA, Hamood AN: **Isolation and characterization of a putative**

multidrug resistance pump from *Vibrio cholerae*. *Mol Microbiol* 1998, **27**:63-72.

Dalbey RE, Chen M: **Sec-translocase mediated membrane protein biogenesis**. *Biochim Biophys Acta* 2004, **1694**:37-53.

Davidson AL, Chen J: **ATP-binding cassette transporters in bacteria**. *Annu Rev Biochem* 2004, **73**:241-268.

Dawson RJ, Locher KP: **Structure of a bacterial multidrug ABC transporter.** *Nature* 2006, **443:**180-185.

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

DeFelice LJ: **Transporter structure and mechanism**. *Trends Neurosci* 2004, **27**:352-359.

Delepelaire P: **Type I secretion in gram-negative bacteria**. *Biochim Biophys Acta* 2004, **1694**:149-161.

Diederichs K, Diez J, Greller G, Muller C, Breed J, Schnell C, Vonrhein C, Boos W, Welte W: **Crystal structure of MalK, the ATPase subunit of the trehalose/maltose ABC transporter of the archaeon *Thermococcus litoralis***. *Embo J* 2000, **19**:5951-5961.

Dong J, Yang G, McHaourab HS: **Structural basis of energy transduction in the transport cycle of MsbA**. *Science* 2005, **308**:1023-1028.

Dridi L, Tankovic J, Petit JC: **CdeA of Clostridium difficile, a new multidrug efflux**

transporter of the MATE family. *Microb Drug Resist* 2004, **10**:191-196.

Dubochet J, McDowall AW, Menge B, Schmid EN, Lickfeld KG: **Electron microscopy of frozen-hydrated bacteria.** *J Bacteriol* 1983, **155:**381-390.

Eisenberg D, Luthy R, Bowie JU: **VERIFY3D: assessment of protein models with three-dimensional profiles**. *Methods Enzymol* 1997, **277**:396-404.

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

Elkins CA, Nikaido H: **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* 2002, 184:6490-8.

Elkins CA, Nikaido H: **Chimeric analysis of AcrA function reveals the importance of its C-terminal domain in its interaction with the AcrB multidrug efflux pump**. *J Bacteriol* 2003, **185**:5349-5356.

Elkins CA, Nikaido H: **3D structure of AcrB: the archetypal multidrug efflux transporter of *Escherichia coli* likely captures substrates from periplasm**. *Drug Resist Updat* 2003, **6**:9-13.

Elvin CM, Hardy CM, Rosenberg H: **Pi exchange mediated by the GlpT-dependent sn-glycerol-3-phosphate transport system in *Escherichia coli*.** *J Bacteriol* 1985, **161**:1054-1058.

Emsley P, Cowtan K: **Coot: model-building tools for molecular graphics**. *Acta Crystallogr D Biol Crystallogr* 2004, **60**:2126-2132.

Ermolova NV, Smirnova IN, Kasho VN, Kaback HR: **Interhelical packing modulates conformational flexibility in the lactose permease of *Escherichia coli*.** *Biochemistry* 2005, **44**:7669-7677.

Eswaran J, Hughes C, Koronakis V: **Locking TolC entrance helices to prevent protein translocation by the bacterial type I export apparatus.** *J Mol Biol* 2003, **327**:309-315.

Federici L, Du D, Walas F, Matsumura H, Fernandez-Recio J, McKeegan KS, Borges-Walmsley MI, Luisi BF, Walmsley AR: **The crystal structure of the outer membrane protein VceC from the bacterial pathogen *Vibrio cholerae* at 1.8 Å resolution.** *J Biol Chem* 2005, **280**:15307-15314.

Ferenci T, Boos W, Schwartz M, Szmelcman S: **Energy-coupling of the transport system of *Escherichia coli* dependent on maltose-binding protein.** *Eur J Biochem* 1977, **75**:187-193.

Fernandez-Recio J, Totrov M, Skorodumov C, Abagyan R: **Optimal docking area: a new method for predicting protein-protein interaction sites.** *Proteins* 2005, **58**:134-143.

Fiser A, Sali A: **Modeller: generation and refinement of homology-based protein structure models.** *Methods Enzymol* 2003, **374**:461-491.

Fleishman SJ, Harrington SE, Enosh A, Halperin D, Tate CG, Ben-Tal N: **Quasi-symmetry in the Cryo-EM structure of EmrE provides the key to modeling its transmembrane domain.** *J Mol Biol* 2006.

Furukawa H, Tsay JT, Jackowski S, Takamura Y, Rock CO: **Thiolactomycin resistance in *Escherichia coli* is associated with the multidrug resistance efflux pump encoded by emrAB**. *J Bacteriol* 1993, **175**:3723-3729.

Gandlur SM, Wei L, Levine J, Russell J, Kaur P: **Membrane topology of the DrrB protein of the doxorubicin transporter of *Streptomyces peucetius***. *J Biol Chem* 2004, **279**:27799-27806.

Gaudet R, Wiley DC: **Structure of the ABC ATPase domain of human TAP1, the transporter associated with antigen processing**. *Embo J* 2001, **20**:4964-4972.

Gerken H, Misra R: **Genetic evidence for functional interactions between TolC and AcrA proteins of a major antibiotic efflux pump of *Escherichia coli***. *Mol Microbiol* 2004, **54**:620-631.

Godsey MH, Baranova NN, Neyfakh AA, Brennan RG: **Crystal structure of MtaN, a global multidrug transporter gene activator**. *J Biol Chem* 2001, **276**:47178-47184.

Gottesman MM, Pastan I, Ambudkar SV: **P-glycoprotein and multidrug resistance**. *Curr Opin Genet Dev* 1996, **6**:610-617.

Graham LL, Harris R, Villiger W, Beveridge TJ: **Freeze-substitution of gram-negative eubacteria: general cell morphology and envelope profiles.** *J Bacteriol* 1991, **173**:1623-1633.

Green HM, Alberola-Ila J: **Development of ERK activity sensor, an in vitro, FRET-based sensor of extracellular regulated kinase activity**. *BMC Chem Biol* 2005, **5**:1.

Griffith JK, Baker ME, Rouch DA, Page MG, Skurray RA, Paulsen IT, Chater KF, Baldwin SA, Henderson PJ: **Membrane transport proteins: implications of sequence comparisons**. *Curr Opin Cell Biol* 1992, **4**:684-695.

Grkovic S, Brown MH, Schumacher MA, Brennan RG, Skurray RA: **The staphylococcal QacR multidrug regulator binds a correctly spaced operator as a pair of dimers**. *J Bacteriol* 2001, **183**:7102-7109.

Grkovic S, Brown MH, Skurray RA: **Regulation of bacterial drug export systems**. *Microbiol Mol Biol Rev* 2002, **66**:671-701.

Guilfoile PG, Hutchinson CR: **A bacterial analog of the mdr gene of mammalian tumor cells is present in *Streptomyces peucetius*, the producer of daunorubicin and doxorubicin**. *Proc Natl Acad Sci U S A* 1991, **88**:8553-8557.

Gutman N, Steiner-Mordoch S, Schuldiner S: **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* 2003, **278**:16082-16087.

Hagman KE, Lucas CE, Balthazar JT, Snyder L, Nilles M, Judd RC, Shafer WM: **The MtrD protein of *Neisseria gonorrhoeae* is a member of the resistance/nodulation/division protein family constituting part of an efflux system**. *Microbiology* 1997, **143 (Pt 7)**:2117-2125.

Hagman KE, Pan W, Spratt BG, Balthazar JT, Judd RC, Shafer WM: **Resistance of *Neisseria gonorrhoeae* to antimicrobial hydrophobic agents is modulated by the mtrRCDE efflux system**. *Microbiology* 1995, **141 (Pt 3)**:611-622.

Hase CC, Thai LS, Boesman-Finkelstein M, Mar VL, Burnette WN, Kaslow HR,

Stevens LA, Moss J, Finkelstein RA: **Construction and characterization of recombinant *Vibrio cholerae* strains producing inactive cholera toxin analogs.** *Infect Immun* 1994, **62**:3051-3057.

He GX, Kuroda T, Mima T, Morita Y, Mizushima T, Tsuchiya T: **An H(+)-coupled multidrug efflux pump, PmpM, a member of the MATE family of transporters, from *Pseudomonas aeruginosa*.** *J Bacteriol* 2004, **186**:262-265.

Heidelberg JF, Eisen JA, Nelson WC, Clayton RA, Gwinn ML, Dodson RJ, Haft DH, Hickey EK, Peterson JD, Umayam L, et al.: **DNA sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*.** *Nature* 2000, **406**:477-483.

Heldwein EE, Brennan RG: **Crystal structure of the transcription activator BmrR bound to DNA and a drug.** *Nature* 2001, **409**:378-382.

Heymann JA, Hirai T, Shi D, Subramaniam S: **Projection structure of the bacterial oxalate transporter OxlT at 3.4A resolution.** *J Struct Biol* 2003, **144**:320-326.

Higgins MK, Bokma E, Koronakis E, Hughes C, Koronakis V: **Structure of the periplasmic component of a bacterial drug efflux pump.** *Proc Natl Acad Sci U S A* 2004, **101**:9994-9999.

Hinrichs W, Kisker C, Duvel M, Muller A, Tovar K, Hillen W, Saenger W: **Structure of the Tet repressor-tetracycline complex and regulation of antibiotic resistance.** *Science* 1994, **264**:418-420.

Hirai T, Subramaniam S: **Structure and transport mechanism of the bacterial oxalate transporter OxlT.** *Biophys J* 2004, **87**:3600-3607.

Hopfner KP, Karcher A, Shin DS, Craig L, Arthur LM, Carney JP, Tainer JA:

Structural biology of Rad50 ATPase: ATP-driven conformational control in DNA double-strand break repair and the ABC-ATPase superfamily. *Cell* 2000, **101**:789-800.

Huang Y, Lemieux MJ, Song J, Auer M, Wang DN: **Structure and mechanism of the glycerol-3-phosphate transporter from *Escherichia coli***. *Science* 2003, **301**:616-620.

Hubbell WL, Cafiso DS, Altenbach C: **Identifying conformational changes with site-directed spin labeling**. *Nat Struct Biol* 2000, 7:735-739.

Huda MN, Chen J, Morita Y, Kuroda T, Mizushima T, Tsuchiya T: **Gene cloning and characterization of VcrM, a Na+-coupled multidrug efflux pump, from *Vibrio cholerae* non-O1**. *Microbiol Immunol* 2003, **47**:419-427.

Huda MN, Morita Y, Kuroda T, Mizushima T, Tsuchiya T: **Na+-driven multidrug efflux pump VcmA from *Vibrio cholerae* non-O1, a non-halophilic bacterium**. *FEMS Microbiol Lett* 2001, **203**:235-239.

Huffman JL, Brennan RG: **Prokaryotic transcription regulators: more than just the helix-turn-helix motif**. *Curr Opin Struct Biol* 2002, **12**:98-106.

Ip H, Stratton K, Zgurskaya H, Liu J: **pH-induced conformational changes of AcrA, the membrane fusion protein of *Escherichia coli* multidrug efflux system**. *J Biol Chem* 2003, **278**:50474-50482.

Jack DL, Storms ML, Tchieu JH, Paulsen IT, Saier MH, JR: **A broad-specificity multidrug efflux pump requiring a pair of homologous SMR-type proteins**. *J Bacteriol* 2000, **182**:2311-2313.

Johnson JM, Church GM: **Alignment and structure prediction of divergent protein families: periplasmic and outer membrane proteins of bacterial efflux pumps.** *J Mol Biol* 1999, **287**:695-715.

Kaatz GW, McAleese F, Seo SM: **Multidrug resistance in *Staphylococcus aureus* due to overexpression of a novel multidrug and toxin extrusion (MATE) transport protein.** *Antimicrob Agents Chemother* 2005, **49**:1857-1864.

Kaback HR, Sahin-Toth M, Weinglass AB: **The kamikaze approach to membrane transport.** *Nat Rev Mol Cell Biol* 2001, **2**:610-620.

Karpowich N, Martsinkevich O, Millen L, Yuan YR, Dai PL, MacVey K, Thomas PJ, Hunt JF: **Crystal structures of the MJ1267 ATP binding cassette reveal an induced-fit effect at the ATPase active site of an ABC transporter.** *Structure (Camb)* 2001, **9**:571-586.

Kazama H, Hamashima H, Sasatsu M, Arai T: **Characterization of the antiseptic-resistance gene qacE delta 1 isolated from clinical and environmental isolates of *Vibrio parahaemolyticus* and *Vibrio cholerae* non-O1.** *FEMS Microbiol Lett* 1999, **174**:379-384.

Kihara A, Akiyama Y, Ito K: **Dislocation of membrane proteins in FtsH-mediated proteolysis.** *Embo J* 1999, **18**:2970-2981.

Kihara A, Akiyama Y, Ito K: **FtsH is required for proteolytic elimination of uncomplexed forms of SecY, an essential protein translocase subunit.** *Proc Natl Acad Sci U S A* 1995, **92**:4532-4536.

Kobayashi N, Nishino K, Hirata T, Yamaguchi A: **Membrane topology of ABC-type macrolide antibiotic exporter MacB in *Escherichia coli*.** *FEBS Lett* 2003,

546:241-246.

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

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

Koteiche HA, Reeves MD, McHaourab HS: **Structure of the substrate binding pocket of the multidrug transporter EmrE: site-directed spin labeling of transmembrane segment 1**. *Biochemistry* 2003, **42**:6099-6105.

Langton KP, Henderson PJ, Herbert RB: **Antibiotic resistance: multidrug efflux proteins, a common transport mechanism?** *Nat Prod Rep* 2005, **22**:439-451.

Lee EH, Shafer WM: **The farAB-encoded efflux pump mediates resistance of gonococci to long-chained antibacterial fatty acids**. *Mol Microbiol* 1999, **33**:839-845.

Lemieux MJ, Reithmeier RA, Wang DN: **Importance of detergent and phospholipid in the crystallization of the human erythrocyte anion-exchanger membrane domain**. *J Struct Biol* 2002, **137**:322-332.

Lemieux MJ, Song J, Kim MJ, Huang Y, Villa A, Auer M, Li XD, Wang DN: **Three-dimensional crystallization of the *Escherichia coli* glycerol-3-phosphate transporter: a member of the major facilitator superfamily**. *Protein Sci* 2003, **12**:2748-2756.

Letoffe S, Delepelaire P, Wandersman C: **Protein secretion in gram-negative**

bacteria: assembly of the three components of ABC protein-mediated exporters is ordered and promoted by substrate binding. *Embo J* 1996, **15**:5804-5811.

Li XZ, Poole K: **Mutational analysis of the OprM outer membrane component of the MexA-MexB-OprM multidrug efflux system of *Pseudomonas aeruginosa***. *J Bacteriol* 2001, **183**:12-27.

Lim D, Poole K, Strynadka NC: **Crystal structure of the MexR repressor of the mexRAB-oprM multidrug efflux operon of *Pseudomonas aeruginosa***. *J Biol Chem* 2002, **277**:29253-29259.

Linton KJ, Cooper HN, Hunter IS, Leadlay PF: **An ABC-transporter from *Streptomyces longisporoflavus* confers resistance to the polyether-ionophore antibiotic tetronasin**. *Mol Microbiol* 1994, **11**:777-785.

Locher KP, Lee AT, Rees DC: **The E. coli BtuCD structure: a framework for ABC transporter architecture and mechanism**. *Science* 2002, **296**:1091-1098.

Lomovskaya O, Lewis K: **Emr, an *Escherichia coli* locus for multidrug resistance**. *Proc Natl Acad Sci U S A* 1992, **89**:8938-8942.

Lupas A, Van Dyke M, Stock J: **Predicting coiled coils from protein sequences**. *Science* 1991, **252**:1162-1164.

Mao W, Warren MS, Black DS, Satou T, Murata T, Nishino T, Gotoh N, Lomovskaya O: **On the mechanism of substrate specificity by resistance nodulation division (RND)-type multidrug resistance pumps: the large periplasmic loops of MexD from Pseudomonas aeruginosa are involved in substrate recognition.** *Mol Microbiol* 2002, 46:889-901.

Marger MD, Saier MH, Jr.: **A major superfamily of transmembrane facilitators that catalyse uniport, symport and antiport.** *Trends Biochem Sci* 1993, **18**:13-20.

Markham PN, Neyfakh AA: **Efflux-mediated drug resistance in Gram-positive bacteria.** *Curr Opin Microbiol* 2001, **4**:509-514.

Mazurkiewicz P, Konings WN, Poelarends GJ: **Acidic residues in the lactococcal multidrug efflux pump LmrP play critical roles in transport of lipophilic cationic compounds.** *J Biol Chem* 2002, **277**:26081-26088.

McGuffin LJ, Bryson K, Jones DT: **The PSIPRED protein structure prediction server.** *Bioinformatics* 2000, **16**:404-405.

Mey AR, Payne SM: **Analysis of residues determining specificity of *Vibrio cholerae* TonB1 for its receptors.** *J Bacteriol* 2003, **185**:1195-1207.

Mikolosko J, Bobyk K, Zgurskaya HI, Ghosh P: **Conformational flexibility in the multidrug efflux system protein AcrA.** *Structure* 2006, **14**:577-587.

Miroux B, Walker JE: **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* 1996, **260**:289-298.

Mizuguchi K, Deane CM, Blundell TL, Johnson MS, Overington JP: **JOY: protein sequence-structure representation and analysis.** *Bioinformatics* 1998, **14**:617-623.

Mokhonov V, Mokhonova E, Yoshihara E, Masui R, Sakai M, Akama H, Nakae T: **Multidrug transporter MexB of *Pseudomonas aeruginosa*: overexpression, purification, and initial structural characterization.** *Protein Expr Purif* 2005, **40**:91-100.

Mokhonov VV, Mokhonova EI, Akama H, Nakae T: **Role of the membrane fusion protein in the assembly of resistance-nodulation-cell division multidrug efflux pump in *Pseudomonas aeruginosa***. *Biochem Biophys Res Commun* 2004, **322**:483-489.

Morita Y, Kataoka A, Shiota S, Mizushima T, Tsuchiya T: **NorM of *Vibrio parahaemolyticus* is an Na(+)-driven multidrug efflux pump**. *J Bacteriol* 2000, **182**:6694-6697.

Murakami S, Nakashima R, Yamashita E, Yamaguchi A: **Crystal structure of bacterial multidrug efflux transporter AcrB**. *Nature* 2002, **419**:587-593.

Murakami S, Nakashima R, Yamashita E, Matsumoto T, Yamaguchi A: **Crystal structures of a multidrug transporter reveal a functionally rotating mechanism.** *Nature* 2006, **443:**173-179.

Murray DS, Schumacher MA, Brennan RG: **Crystal structures of QacR-diamidine complexes reveal additional multidrug-binding modes and a novel mechanism of drug charge neutralization**. *J Biol Chem* 2004, **279**:14365-14371.

Murshudov GN, Vagin AA, Lebedev A, Wilson KS, Dodson EJ: **Efficient anisotropic refinement of macromolecular structures using FFT**. *Acta Crystallogr D Biol Crystallogr* 1999, **55 (Pt 1)**:247-255.

Nehme D, Li XZ, Elliot R, Poole K: **Assembly of the MexAB-OprM multidrug efflux system of *Pseudomonas aeruginosa*: identification and characterization of mutations in mexA compromising MexA multimerization and interaction with MexB**. *J Bacteriol* 2004, **186**:2973-2983.

Newberry KJ, Brennan RG: **The structural mechanism for transcription activation by MerR family member multidrug transporter activation, N terminus.** *J Biol Chem* 2004, **279**:20356-20362.

Neyfakh AA: **Mystery of multidrug transporters: the answer can be simple.** *Mol Microbiol* 2002, **44**:1123-1130.

Nicholls A, Sharp KA, Honig B: **Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons.** *Proteins* 1991, **11**:281-296.

Nikaido H: **Prevention of drug access to bacterial targets: permeability barriers and active efflux.** *Science* 1994, **264**:382-388.

Nikaido H: **Multiple antibiotic resistance and efflux.** *Curr Opin Microbiol* 1998, **1**:516-523.

Nikaido H: **Preventing drug access to targets: cell surface permeability barriers and active efflux in bacteria.** *Semin Cell Dev Biol* 2001, **12**:215-223.

Nikaido H, Basina M, Nguyen V, Rosenberg EY: **Multidrug efflux pump AcrAB of *Salmonella typhimurium* excretes only those beta-lactam antibiotics containing lipophilic side chains.** *J Bacteriol* 1998, **180**:4686-4692.

Nishino K, Yamaguchi A: **Analysis of a complete library of putative drug transporter genes in *Escherichia coli*.** *J Bacteriol* 2001, **183**:5803-5812.

Olano C, Rodriguez AM, Mendez C, Salas JA: **A second ABC transporter is involved in oleandomycin resistance and its secretion by *Streptomyces antibioticus*.** *Mol Microbiol* 1995, **16**:333-343.

Otsuka M, Yasuda M, Morita Y, Otsuka C, Tsuchiya T, Omote H, Moriyama Y: **Identification of essential amino acid residues of the NorM Na+/multidrug antiporter in *Vibrio parahaemolyticus***. *J Bacteriol* 2005, **187**:1552-1558.

Patterson GH, Piston DW, Barisas BG: **Forster distances between green fluorescent protein pairs**. *Anal Biochem* 2000, **284**:438-440.

Paulsen IT: **Multidrug efflux pumps and resistance: regulation and evolution**. *Curr Opin Microbiol* 2003, **6**:446-451.

Paulsen IT, Brown MH, Littlejohn TG, Mitchell BA, Skurray RA: **Multidrug resistance proteins QacA and QacB from *Staphylococcus aureus*: membrane topology and identification of residues involved in substrate specificity**. *Proc Natl Acad Sci U S A* 1996, **93**:3630-3635.

Paulsen IT, Brown MH, Skurray RA: **Proton-dependent multidrug efflux systems**. *Microbiol Rev* 1996, **60**:575-608.

Paulsen IT, Nguyen L, Sliwinski MK, Rabus R, Saier MH, Jr.: **Microbial genome analyzes: comparative transport capabilities in eighteen prokaryotes**. *J Mol Biol* 2000, **301**:75-100.

Paulsen IT, Skurray RA: **Topology, structure and evolution of two families of proteins involved in antibiotic and antiseptic resistance in eukaryotes and prokaryotes--an analysis**. *Gene* 1993, **124**:1-11.

Perrakis A, Harkiolaki M, Wilson KS, Lamzin VS: **ARP/wARP and molecular replacement**. *Acta Crystallogr D Biol Crystallogr* 2001, **57**:1445-1450.

Piddock LJ: **Multidrug-resistance efflux pumps - not just for resistance.** *Nat Rev Microbiol* 2006, **4**:629-636.

Piddock LJ: **Clinically relevant chromosomally encoded multidrug resistance efflux pumps in bacteria.** *Clin Microbiol Rev* 2006, **19**:382-402.

Podlesek Z, Comino A, Herzog-Velikonja B, Zgur-Bertok D, Komel R, Grabnar M: ***Bacillus licheniformis* bacitracin-resistance ABC transporter: relationship to mammalian multidrug resistance.** *Mol Microbiol* 1995, **16**:969-976.

Pos KM, Schiefner A, Seeger MA, Diederichs K: **Crystallographic analysis of AcrB.** *FEBS Lett* 2004, **564**:333-339.

Putman M, van Veen HW, Konings WN: **Molecular properties of bacterial multidrug transporters.** *Microbiol Mol Biol Rev* 2000, **64**:672-693.

Ren Q, Kang KH, Paulsen IT: **TransportDB: a relational database of cellular membrane transport systems.** *Nucleic Acids Res* 2004, **32**:D284-288.

Roosild TP, Greenwald J, Vega M, Castronovo S, Riek R, Choe S: **NMR structure of Mistic, a membrane-integrating protein for membrane protein expression.** *Science* 2005, **307**:1317-1321.

Ross JI, Eady EA, Cove JH, Cunliffe WJ, Baumberg S, Wootton JC: **Inducible erythromycin resistance in staphylococci is encoded by a member of the ATP-binding transport super-gene family.** *Mol Microbiol* 1990, **4**:1207-1214.

Rotem D, Schuldiner S: **EmrE, a multidrug transporter from *Escherichia coli*, transports monovalent and divalent substrates with the same stoichiometry.** *J Biol Chem* 2004, **279**:48787-48793.

Rotem D, Steiner-Mordoch S, Schuldiner S: **Identification of tyrosine residues critical for the function of an ion-coupled multidrug transporter.** *J Biol Chem* 2006, **281**:18715-18722.

Rouch DA, Cram DS, DiBerardino D, Littlejohn TG, Skurray RA: **Efflux-mediated antiseptic resistance gene qacA from *Staphylococcus aureus*: common ancestry with tetracycline- and sugar-transport proteins.** *Mol Microbiol* 1990, **4**:2051-2062.

Rouquette-Loughlin C, Stojiljkovic I, Hrobowski T, Balthazar JT, Shafer WM: **Inducible, but not constitutive, resistance of gonococci to hydrophobic agents due to the MtrC-MtrD-MtrE efflux pump requires TonB-ExbB-ExbD proteins.** *Antimicrob Agents Chemother* 2002, **46**:561-565.

Rouquette-Loughlin CE, Balthazar JT, Shafer WM: **Characterization of the MacA-MacB efflux system in *Neisseria gonorrhoeae*.** *J Antimicrob Chemother* 2005, **56**:856-860.

Saidijam M, Psakis G, Clough JL, Meuller J, Suzuki S, Hoyle CJ, Palmer SL, Morrison SM, Pos MK, Essenberg RC, et al.: **Collection and characterisation of bacterial membrane proteins.** *FEBS Lett* 2003, **555**:170-175.

Saier MH, Jr., Paulsen IT: **Phylogeny of multidrug transporters.** *Semin Cell Dev Biol* 2001, **12**:205-213.

Scarborough GA: **Large single crystals of the *Neurospora crassa* plasma membrane H+-ATPase: an approach to the crystallization of integral membrane proteins.** *Acta Crystallogr D Biol Crystallogr* 1994, **50**:643-649.

Schmitt L, Benabdelhak H, Blight MA, Holland IB, Stubbs MT: **Crystal structure of**

the nucleotide-binding domain of the ABC-transporter haemolysin B: identification of a variable region within ABC helical domains. *J Mol Biol* 2003, 330:333-342.

Schuldiner S: **Structural biology: the ins and outs of drug transport.** *Nature* 2006, **443:**156-157.

Schumacher MA, Brennan RG: **Deciphering the molecular basis of multidrug recognition: crystal structures of the Staphylococcus aureus multidrug binding transcription regulator QacR**. *Res Microbiol* 2003, **154**:69-77.

Schumacher MA, Miller MC, Brennan RG: **Structural mechanism of the simultaneous binding of two drugs to a multidrug-binding protein**. *Embo J* 2004, **23**:2923-2930.

Schumacher MA, Miller MC, Grkovic S, Brown MH, Skurray RA, Brennan RG: **Structural mechanisms of QacR induction and multidrug recognition**. *Science* 2001, **294**:2158-2163.

Schumacher MA, Miller MC, Grkovic S, Brown MH, Skurray RA, Brennan RG: **Structural basis for cooperative DNA binding by two dimers of the multidrug-binding protein QacR**. *Embo J* 2002, **21**:1210-1218.

Schweizer HP: **Efflux as a mechanism of resistance to antimicrobials in *Pseudomonas aeruginosa* and related bacteria: unanswered questions**. *Genet Mol Res* 2003, **2**:48-62.

Seeger MA, Schiefner A, Eicher T, Verrey F, Diederichs K, Pos KM: **Structural asymmetry of AcrB trimer suggests a peristaltic pump mechanism.** *Science* 2006, **313:**1295-1298.

Senior AE, al-Shawi MK, Urbatsch IL: **The catalytic cycle of P-glycoprotein**. *FEBS Lett* 1995, **377**:285-289.

Shafer WM, Qu X, Waring AJ, Lehrer RI: **Modulation of *Neisseria gonorrhoeae* susceptibility to vertebrate antibacterial peptides due to a member of the resistance/nodulation/division efflux pump family**. *Proc Natl Acad Sci U S A* 1998, **95**:1829-1833.

Sharff A, Fanutti C, Shi J, Calladine C, Luisi B: **The role of the TolC family in protein transport and multidrug efflux. From stereochemical certainty to mechanistic hypothesis**. *Eur J Biochem* 2001, **268**:5011-5026.

Sharoni M, Steiner-Mordoch S, Schuldiner S: **Exploring the binding domain of EmrE, the smallest multidrug transporter**. *J Biol Chem* 2005, **280**:32849-32855.

Shaw KJ, Rather PN, Hare RS, Miller GH: **Molecular genetics of aminoglycoside resistance genes and familial relationships of the aminoglycoside-modifying enzymes**. *Microbiol Rev* 1993, **57**:138-163.

Smirnova IN, Kaback HR: **A mutation in the lactose permease of *Escherichia coli* that decreases conformational flexibility and increases protein stability**. *Biochemistry* 2003, **42**:3025-3031.

Smith PC, Karpowich N, Millen L, Moody JE, Rosen J, Thomas PJ, Hunt JF: **ATP binding to the motor domain from an ABC transporter drives formation of a nucleotide sandwich dimer**. *Mol Cell* 2002, **10**:139-149.

Soskine M, Adam Y, Schuldiner S: **Direct evidence for substrate-induced proton release in detergent-solubilized EmrE, a multidrug transporter**. *J Biol Chem* 2004,

279:9951-9955.

Soskine M, Mark S, Tayer N, Mizrachi R, Schuldiner S: **On parallel and antiparallel topology of an homodimeric multidrug transporter.** *J Biol Chem* 2006.

Soskine M, Steiner-Mordoch S, Schuldiner S: **Crosslinking of membrane-embedded cysteines reveals contact points in the EmrE oligomer**. *Proc Natl Acad Sci U S A* 2002, **99**:12043-12048.

Stegmeier JF, Polleichtner G, Brandes N, Hotz C, Andersen C: **Importance of the adaptor (membrane fusion) protein hairpin domain for the functionality of multidrug efflux pumps.** *Biochemistry* 2006, **45:**10303-10312.

Stenberg F, Chovanec P, Maslen SL, Robinson CV, Ilag LL, von Heijne G, Daley DO: **Protein complexes of the *Escherichia coli* cell envelope**. *J Biol Chem* 2005, **280**:34409-34419.

Su XZ, Chen J, Mizushima T, Kuroda T, Tsuchiya T: **AbeM, an H+-coupled *Acinetobacter baumannii* multidrug efflux pump belonging to the MATE family of transporters**. *Antimicrob Agents Chemother* 2005, **49**:4362-4364.

Tamura N, Murakami S, Oyama Y, Ishiguro M, Yamaguchi A: **Direct interaction of multidrug efflux transporter AcrB and outer membrane channel TolC detected via site-directed disulfide cross-linking.** *Biochemistry* 2005, **44:**11115-11121.

Tate CG: **Comparison of three structures of the multidrug transporter EmrE.** *Curr Opin Struct Biol* 2006, **16:**457-464.

Thanabalu T, Koronakis E, Hughes C, Koronakis V: **Substrate-induced assembly of a contiguous channel for protein export from *E.coli*: reversible bridging of an**

inner-membrane translocase to an outer membrane exit pore. *Embo J* 1998, **17**:6487-6496.

Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG: **The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools**. *Nucleic Acids Res* 1997, **25**:4876-4882.

Tikhonova EB, Wang Q, Zgurskaya HI: **Chimeric analysis of the multicomponent multidrug efflux transporters from gram-negative bacteria.** *J Bacteriol.* 2002, 184:6499-507.

Tikhonova EB, Zgurskaya HI: **AcrA, AcrB, and TolC of *Escherichia coli* form a stable intermembrane multidrug efflux complex**. *J Biol Chem* 2004, **279**:32116-32124.

Touze T, Eswaran J, Bokma E, Koronakis E, Hughes C, Koronakis V: **Interactions underlying assembly of the *Escherichia coli* AcrAB-TolC multidrug efflux system**. *Mol Microbiol* 2004, **53**:697-706.

Tsukagoshi N, Aono R: **Entry into and release of solvents by *Escherichia coli* in an organic-aqueous two-liquid-phase system and substrate specificity of the AcrAB-TolC solvent-extruding pump**. *J Bacteriol* 2000, **182**:4803-4810.

Tzeng YL, Ambrose KD, Zughaier S, Zhou X, Miller YK, Shafer WM, Stephens DS: **Cationic antimicrobial peptide resistance in *Neisseria meningitidis***. *J Bacteriol* 2005, **187**:5387-5396.

Ubarretxena-Belandia I, Baldwin JM, Schuldiner S, Tate CG: **Three-dimensional structure of the bacterial multidrug transporter EmrE shows it is an asymmetric homodimer**. *Embo J* 2003, **22**:6175-6181.

Ubarretxena-Belandia I, Tate CG: **New insights into the structure and oligomeric state of the bacterial multidrug transporter EmrE: an unusual asymmetric homo-dimer**. *FEBS Lett* 2004, **564**:234-238.

Vaccaro L, Koronakis V, Sansom MS: **Flexibility in a drug transport accessory protein: molecular dynamics simulations of MexA.** *Biophys J* 2006, **91**:558-564.

Vagin AA, Steiner RA, Lebedev AA, Potterton L, McNicholas S, Long F, Murshudov GN: **REFMAC5 dictionary: organization of prior chemical knowledge and guidelines for its use**. *Acta Crystallogr D Biol Crystallogr* 2004, **60**:2184-2195.

Vaguine AA, Richelle J, Wodak SJ: **SFCHECK: a unified set of procedures for evaluating the quality of macromolecular structure-factor data and their agreement with the atomic model**. *Acta Crystallogr D Biol Crystallogr* 1999, **55 (Pt 1)**:191-205.

Van den Berg B, Clemons WM, Jr., Collinson I, Modis Y, Hartmann E, Harrison SC, Rapoport TA: **X-ray structure of a protein-conducting channel**. *Nature* 2004, **427**:36-44.

van Veen HW: **Towards the molecular mechanism of prokaryotic and eukaryotic multidrug transporters**. *Semin Cell Dev Biol* 2001, **12**:239-245.

van Veen HW, Higgins CF, Konings WN: **Multidrug transport by ATP binding cassette transporters: a proposed two-cylinder engine mechanism**. *Res Microbiol* 2001, **152**:365-374.

van Veen HW, Venema K, Bolhuis H, Oussenko I, Kok J, Poolman B, Driessen AJ, Konings WN: **Multidrug resistance mediated by a bacterial homolog of the**

human multidrug transporter MDR1. *Proc Natl Acad Sci U S A* 1996, **93**:10668-10672.

Veal WL, Nicholas RA, Shafer WM: **Overexpression of the MtrC-MtrD-MtrE efflux pump due to an mtrR mutation is required for chromosomally mediated penicillin resistance in *Neisseria gonorrhoeae***. *J Bacteriol* 2002, **184**:5619-5624.

Vediyappan G, Borisova T, Fralick JA: **Isolation and characterization of VceC gain-of-function mutants that can function with the AcrAB multiple-drug-resistant efflux pump of *Escherichia coli*.** *J Bacteriol* 2006, **188:**3757-3762.

Venter JC, Adams MD, Myers EW, Li PW, Mural RJ, Sutton GG, Smith HO, Yandell M, Evans CA, Holt RA, et al.: **The sequence of the human genome**. *Science* 2001, **291**:1304-1351.

Verdon G, Albers SV, Dijkstra BW, Driessen AJ, Thunnissen AM: **Crystal structures of the ATPase subunit of the glucose ABC transporter from Sulfolobus solfataricus: nucleotide-free and nucleotide-bound conformations**. *J Mol Biol* 2003, **330**:343-358.

Wang DN, Safferling M, Lemieux MJ, Griffith H, Chen Y, Li XD: **Practical aspects of overexpressing bacterial secondary membrane transporters for structural studies**. *Biochim Biophys Acta* 2003, **1610**:23-36.

White DG, Goldman JD, Demple B, Levy SB: **Role of the acrAB locus in organic solvent tolerance mediated by expression of marA, soxS, or robA in *Escherichia coli***. *J Bacteriol* 1997, **179**:6122-6126.

Woebking B, Reuter G, Shilling RA, Velamakanni S, Shahi S, Venter H, Balakrishnan L, van Veen HW: **Drug-lipid A interactions on the *Escherichia coli* ABC transporter MsbA.** *J Bacteriol* 2005, 187:6363-9.

Wolf E, Kim PS, Berger B: **MultiCoil: a program for predicting two- and three-stranded coiled coils.** *Protein Sci* 1997, **6**:1179-1189.

Woolley RC, Vediyappan G, Anderson M, Lackey M, Ramasubramanian B, Jiangping B, Borisova T, Colmer JA, Hamood AN, McVay CS, et al.: **Characterization of the Vibrio cholerae vceCAB multiple-drug resistance efflux operon in *Escherichia coli*.** *J Bacteriol* 2005, **187**:5500-5503.

Yang Q, Wang X, Ye L, Mentrikoski M, Mohammadi E, Kim YM, Maloney PC: **Experimental tests of a homology model for OxlT, the oxalate transporter of Oxalobacter formigenes.** *Proc Natl Acad Sci U S A* 2005, **102**:8513-8.

Yernool D, Boudker O, Jin Y, Gouaux E: **Structure of a glutamate transporter homologue from *Pyrococcus horikoshii*.** *Nature* 2004, **431**:811-818.

Yerushalmi H, Lebendiker M, Schuldiner S: **EmrE, an *Escherichia coli* 12-kDa multidrug transporter, exchanges toxic cations and H+ and is soluble in organic solvents.** *J Biol Chem* 1995, **270**:6856-6863.

Yerushalmi H, Lebendiker M, Schuldiner S: **Negative dominance studies demonstrate the oligomeric structure of EmrE, a multidrug antiporter from *Escherichia coli*.** *J Biol Chem* 1996, **271**:31044-31048.

Yerushalmi H, Schuldiner S: **An essential glutamyl residue in EmrE, a multidrug antiporter from *Escherichia coli*.** *J Biol Chem* 2000, **275**:5264-5269.

Yin Y, He X, Szewczyk P, Nguyen T, Chang G: **Structure of the multidrug transporter EmrD from *Escherichia coli.*** *Science* 2006, **312:**741-744.

Yoneyama H, Maseda H, Kamiguchi H, Nakae T: **Function of the membrane fusion protein, MexA, of the MexA, B-OprM efflux pump in *Pseudomonas aeruginosa* without an anchoring membrane.** *J Biol Chem* 2000, **275**:4628-4634.

Yu EW, Aires JR, Nikaido H: **AcrB multidrug efflux pump of *Escherichia coli*: composite substrate-binding cavity of exceptional flexibility generates its extremely wide substrate specificity.** *J Bacteriol* 2003, **185**:5657-5664.

Yu EW, McDermott G, Zgurskaya HI, Nikaido H, Koshland DE, Jr.: **Structural basis of multiple drug-binding capacity of the AcrB multidrug efflux pump.** *Science* 2003, **300**:976-980.

Yuan YR, Blecker S, Martsinkevich O, Millen L, Thomas PJ, Hunt JF: **The crystal structure of the MJ0796 ATP-binding cassette: Implications for the structural consequences of ATP hydrolysis in the active site of an ABC transporter.** *J Biol Chem* 2001, **276**:32313-32321.

Zgurskaya HI, Nikaido H: **Cross-linked complex between oligomeric periplasmic lipoprotein AcrA and the inner-membrane-associated multidrug efflux pump AcrB from *Escherichia coli.*** *J Bacteriol* 2000, **182**:4264-4267.

Zhao Q, Li XZ, Mistry A, Srikumar R, Zhang L, Lomovskaya O, Poole K: **Influence of the TonB energy-coupling protein on efflux-mediated multidrug resistance in *Pseudomonas aeruginosa*.** *Antimicrob Agents Chemother* 1998, **42**:2225-2231.

Zhao Q, Poole K: **Mutational analysis of the TonB1 energy coupler of *Pseudomonas aeruginosa*.** *J Bacteriol* 2002, **184**:1503-1513.

Zheleznova EE, Markham PN, Neyfakh AA, Brennan RG: **Structural basis of multidrug recognition by BmrR, a transcription activator of a multidrug transporter**. *Cell* 1999, **96**:353-362.

Zheng L, Hoeflich KP, Elsby LM, Ghosh M, Roberts SG, Ikura M: **FRET evidence for a conformational change in TFIIB upon TBP-DNA binding**. *Eur J Biochem* 2004, **271**:792-800.

Zhu ZY, Sali A, Blundell TL: **A variable gap penalty function and feature weights for protein 3-D structure comparisons**. *Protein Eng* 1992, **5**:43-51.

Appendix A:

Articles published from this thesis

Federici L*, **Du D***, Walas F, Matsumura H, Fernandez-Recio J, McKeegan KS, Borges-Walmsley MI, Luisi BF, Walmsley AR: **The crystal structure of the outer membrane protein VceC from the bacterial pathogen Vibrio cholerae at 1.8 A resolution**. *J Biol Chem* 2005, **280**:15307-15314. **(*equal contributions)**

Borges-Walmsley MI, **Du D**, McKeegan KS, Sharples GJ, Walmsley AR: **VceR regulates the vceCAB drug efflux pump operon of Vibrio cholerae by alternating between mutually exclusive conformations that bind either drugs or promoter DNA**. *J Mol Biol* 2005, **349**:387-400.

Appendix B:

DNA and amino acid sequences of VceCAB pump of *Vibrio cholerae* vaccine strain CVD101

1. DNA sequence of VceC

ATGAAAAATAGCGTTCAAACGGTAGGTTTGTTGCATAAAGCAGCGCCTTATTTTACG
GCTTCTCTATTGAGTGCTTTCGTTTTAACGGGTTGCTCAGTACCCGATCATTACCCT
GATCTTGCGACGATGTGGGAAGCGAATGAGCTTAGCTCGACAAACACATTTTCTCAC
CAAGCTGAAATGGACTGGCCATCAGTGAATTGGTGGCAACGCTATCAAGATGCACAA
TTAAACCATCTGATTGAGGAAGCGTTGCAGCACTCACCATCGTTAGAGATGGCAATG
GCTCGGCTTAAAGGGGCACAAGGTTTTGCTCGTCAAGCAGGCGCAATTCGATCATTT
GATCTCGGGTTAGCGGCATCCGCCACGGAAAGTAAAGTCAGTGAACGTTATCAATCT
GCTACGCCTCCCGATGGTTGGAATGACTATGGCACTCTAACCTTGAATTTTCAGTAT
GACTTTGACTTTTGGGGAAAAAATCGCGCAGCGGTTGTTGCGGCGACTTCCGAGCTT
GCTGCCGCTGAGGCAGAGAGTGTGGCTGCGCGTTTGATGATCAGCACATCGATTGCT
AATGCTTATGCCGAATTAGCTCGTCTTTATGCCAATCAAGAGACCGTGCATGCCGCG
CTACAAGTGCGCAACAAAACCGTTGAGCTTTTAGAAAAACGTTATGCCAATGGCTTA
GAAACGCTAGGTTCTGTCAGCCAAGCGAAAGCGGTTGCTGCGAGTGTGGAAGCGGAG
TTGCTTGGTATTCAAGAATCGATCCAATTACAGAAAAATGCTTTGGCCGCTTTGGTG
GGGCAAGGGCCTGATCGCGCTGCATCGATTGAAGAACCCCATATCACCTTGACCTCT
CGTTATGGCTTGCCGTCGGAGGCTGGCGTTGGATTACTGGGGCATCGTGCTGATATC
ACCGCTGCGCGTTGGCGGGCAGAAGCCGCAGCACAGCAAGTCGGTATTGCGCAAGCA
CAATTCTATCCAGATGTCACTTTGTCAGCGTTTATTGGTTATCAGGCATTTGGTCTG
GATCATTTGTTTGATAGTGGTAATGACGCAGGTGCAATTGGGCCTGCGATCTACTTA
CCACTATTTACGGGTGGGCGTTTAGAGGGGCAGTTGACTTCGGCAGAAGCTCGTTAC
CAAGAAGCGGTCGCTCAATACAATGGCACGTTAGTGCAAGCCTTGCATGAAATCGCC
GATGTGGTGACCAGTAGCCAAGCTCTGCAAGCGCGAATTAATAAAACCGAACAAGCC
GTACAGCAAGCTGAACAAGCACTGCATATCGCGACGAACCGTTATCAAGGCGGTTTA
GCCACCTATCTCGATGTTTTGGTAGCTGAAGAGTCTTTACTCAATAACCAAAGAGCA
TTGGTCAATTTGCAGTCACGCGCTTTTAGTTTAGATCTTGCGCTGATCCATGCGCTA
GGCGGCGGTTTTGAAACAACAGAATCTTAA

2. Amino acid sequence of VceC

MKNSVQTVGLLHKAAPYFTASLLSAFVLTGCSVPDHYPDLATMWEANELSSTNTFSH
QAEMDWPSVNWWQRYQDAQLNHLIEEALQHSPSLEMAMARLKGAQGFARQAGAIRSF
DLGLAASATESKVSERYQSATPPDGWNDYGTLTLNFQYDFDFWGKNRAAVVAATSEL
AAAEAESVAARLMISTSIANAYAELARLYANQETVHAALQVRNKTVELLEKRYANGL

ETLGSVSQAKAVAASVEAELLGIQESIQLQKNALAALVGQGPDRAASIEEPHITLTS
RYGLPSEAGVGLLGHRADITAARWRAEAAAQQVGIAQAQFYPDVTLSAFIGYQAFGL
DHLFDSGNDAGAIGPAIYLPLFTGGRLEGQLTSAEARYQEAVAQYNGTLVQALHEIA
DVVTSSQALQARINKTEQAVQQAEQALHIATNRYQGGLATYLDVLVAEESLLNNQRA
LVNLQSRAFSLDLALIHALGGGFETTES

3. DNA sequence of VceA

TTGAAACAACAGAATCTTAATAGGTTATCTTCAATGAATTCAAATAATAGCAACACC
GAATTTGATGGTGTACGCGCAGCTCAATCGCGTAAAAAAGGGTTTCTCGGATTAGCG
GCCGCAATTGTGGTTGCCGGAGGGAGTTATGCACTCTATTGGCACTTTATCGGTTCA
CGCTACATTTCCACCGACAACGCTTATGCTGCGGCGGAAATCGCCGAAGTCACGCCG
GCCGTGGGCGGGATTATCGCGCAAGTGAATGTAGTGGATACCGAGTACGTCAAACAA
GGCGACGTACTTGTACAGCTGGATGATACCGATGCGAGATTAGCTTTGCTTCAAGCT
GAGGCGGATTTGGCACTCGCAAAACGCCGTGTACGTAGCTATCTCGCTAATGACGAA
GGTTTAAGTGCTATGGTGGAAGCACAAGAAGCCAATGAGCAGCGTGTGAAAGCGCAG
TTAAAAGCGGCGCAAGCGGACTTTGAACGAGCAAAAATCGATTTATCTCGTCGGGAA
GACTTAGTGCGTTCCGGCTCCGTTTCTGGTGAAGAGCTGACGAATGCCAAAACCGGT
TTTGCGCAGGCACAAGCAAACCTTAACGCGGCAAAAGCGGCTATGGCGCAAGCACAG
GCAACTAAGCTCTCGACCATTGGCTCACAAAAAGCTAATGCAGCGTTGACCGATAAC
ACGACCGTCGATAGCAACCCTGAAGTCTTACTGGCAAAAGCGCGTTATGAACAAGCG
AAAATCGATCTCGAGCGTACGGTGATTCGTGCGCCGATCAGTGGAATCGTCGCTAAA
CGCCAAGTACAAGTTGGGCGTCGTGTACAGGTGGGGATGCCTTTAATGACCGTGGTG
CCTACTGACCATATTTATGCCGATGCCAATTTTAAAGAAGTGGAATTGCGCGATGTG
AAAGTTGGACAACCCGTTACCTTAACGGCAGATCTGTATGGTGATGATGTGACCTAC
CATGGTGTGGTCGCTGGTTTTTCCGGCGGTACGGGCTCCGCATTTTCGATGATCCCT
GCACAAAACGCCACAGGCAACTGGATTAAAGTGGTGCAGCGTTTGCCGATCCGTATC
GAGTTGGATCCCAAAGATTTGCAAGCCTATCCACTGCAGGTAGGATTATCTATGGTT
GCCACGATAGATACTGCTGGAACTACCGATCCACAAACCTTAGTGCAGTACCGCGCC
GCCAAAGTATCAGAGCAAATCTAA

4. Amino acid sequence of VceA

MKQQNLNRLSSMNSNNSNTEFDGVRAAQSRKKGFLGLAAAIVVAGGSYALYWHFIGS
RYISTDNAYAAAEIAEVTPAVGGIIAQVNVVDTEYVKQGDVLVQLDDTDARLALLQA
EADLALAKRRVRSYLANDEGLSAMVEAQEANEQRVKAQLKAAQADFERAKIDLSRRE
DLVRSGSVSGEELTNAKTGFAQAQANLNAAKAAMAQAQATKLSTIGSQKANAALTDN
TTVDSNPEVLLAKARYEQAKIDLERTVIRAPISGIVAKRQVQVGRRVQVGMPLMTVV
PTDHIYADANFKEVELRDVKVGQPVTLTADLYGDDVTYHGVVAGFSGGTGSAFSMIP
AQNATGNWIKVVQRLPIRIELDPKDLQAYPLQVGLSMVATIDTAGTTDPQTLVQYRA
AKVSEQI*

5. DNA sequence of VceB

ATGAGTCATAACGCTGACAATGAGATGCAACCTCTGTCAGGATGGGCACTGTTTTTC
GGTGCCCTGTGTCTGGCAATGGCCAACTTTCTCGCTATTTTGGATACCACTATTGCT
AACGTCTCTGTCTCCAATATCGCTGGTAGCCTAGGAACTTCAACTAGCCAAGGGACT
TATGTGATTACCTCTTATGCGGTCGCAGAGGCGATCTCGGTACCTCTTACGGGTTGG
TTGGCTTCGCGATTTGGCTCTATTCGAGTCTTTGTCACCTGTTTTCTGTTATTCGGT
GTTTTTTCATTACTCTGTGGTTTAGCCAATAGCATGAGCACGCTGGTGATGTTTCGA
GTGTTGCTTGGTTTCGTGGGAGGGCCTTTAATGCCACTTTCACAAACCTTGATGATG
CGAATATTCCCCAAAAACAAAAGCCATGCTGCGATTGGTATCTGGTCGATGACCACG
CTCGTGGCACCTATCATGGGACCTATTCTTGGCGGCGTGTTGTGTGACCAACTCAGT
TGGCCATACATCTTTTTCATCAAAATGCCTTTCGCTATTGCCGCGGCGTTACTGTGT
TGGAAATTGCTGAAAAAATTTGAAACCAAAACGACGCATTCAAAAATTGATAAAGTG
GGGCTGGCGTTATTGGTCGTGTGGGTCGCCGCATTGCAATTAATGTTGGATGAAGGT
AAGGATCATGACTGGTTTGAGTCATCGCGCATTGTGTTTTTGGCGGTGATTGCTGTG
ATTGGTTTCATTGCATTCTTGATTTGGGAGCTAACGGAGCGAAACCCTGTCGTGGAT
TTAAAGGTATTCCGACATCGAGGTTACAGTATTAGCATGGTGACGTTNTCACTCGCG
TTCGGCGCTTTTTTTAGTATTTCTGTAGTAACACCTTTATGGTTGCAAATTNACATG
GGTTATACCGCCACCATTTCAGGCCACGCTACTGCCAGTATGGGGATTCTAGCGGTT
TTTTTGGCACCGATAGTGGCCAATCTGTCCTCGAAATTTGATCCTCGGCCCTTTGTT
TTTGCTGGCGTGATGTGGCTTGGGCTGTGGACTTTCATGCGTGGCTTTAACACAGTA
GATATGACGTTTTCGCAAATCAGTTGGCCTTTATTTTTCCAAGGTATCGGCATGCCG
TTGTTCTTTGTGCCATTGACTGCGATAGCCCTTGGTAGTGTCAAACCGCACGAAATG
GAATCAGCAGCAGGATTGATGAACTTTATCCGAACCTTGTCCGGTGCGTTTGCGACC
TCAATGATCAACACTTCATGGGAGCATGAAACACGTTATGTGCATGCTGAGTTGGCA
GGCTTAACTGATAAAGCGGGTGTCGCGGCACAAGCGATGCAAAGTTCAGGAATGAGT
GCCGAGCAGACGCGCTCTGCGATGGACTGGATATTGCAAAATCAGAGTGTGATGGTG
GCAACCAACCAGTTGTTTATCGTGATTGCACTGATCTTTGTATTTGCAGCCTGCATG
ATTTGGTTTGCGCCAAAACCCAAGCAAGCGGTGGATACCTCTGCTGTTCATTAA

6. Amino acid sequence of VceB

MSHNADNEMQPLSGWALFFGALCLAMANFLAILDTTIANVSVSNIAGSLGTSTSQGT
YVITSYAVAEAISVPLTGWLASRFGSIRVFVTCFLLFGVFSLLCGLANSMSTLVMFR
VLLGFVGGPLMPLSQTLMMRIFPKNKSHAAIGIWSMTTLVAPIMGPILGGVLCDQLS
WPYIFFIKMPFAIAAALLCWKLLKKFETKTTHSKIDKVGLALLVVWVAALQLMLDEG
KDHDWFESSRIVFLAVIAVIGFIAFLIWELTERNPVVDLKVFRHRGYSISMVTXSLA
FGAFFSISVVTPLWLQIXMGYTATISGHATASMGILAVFLAPIVANLSSKFDPRPFV
FAGVMWLGLWTFMRGFNTVDMTFSQISWPLFFQGIGMPLFFVPLTAIALGSVKPHEM
ESAAGLMNFIRTLSGAFATSMINTSWEHETRYVHAELAGLTDKAGVAAQAMQSSGMS
AEQTRSAMDWILQNQSVMVATNQLFIVIALIFVFAACMIWFAPKPKQAVDTSAVH*