Characterisation of *Shigella flexneri* polysaccharide co-polymerase (PCP) protein Wzz

Analysis of structure, function and protein interaction

Magdalene Papadopoulos

Submitted for the Degree of Doctor of Philosophy

Discipline of Microbiology and Immunology

The School of Molecular and Biomedical Science

The University of Adelaide

August 2010

CHAPTER ONE

INTRODUCTION

1.1 SHIGELLA

Shigella are Gram-negative facultative intracellular bacteria belonging to the family *Enterobacteriaceae*, and have long been identified as causative agents of bacillary dysentery, otherwise known as shigellosis (Sansonetti, 2001). Shigellosis is a potentially fatal invasive disease of the lower gastrointestinal tract transmitted via the faecal-oral route. *Shigellae* are now classified as a subtype of *E. coli* (Lan and Reeves, 2002; Parsot, 2005) and are divided into subgroups; *Shigella dysenteriae* (Group A, which is made up of 15 serotypes), *Shigella flexneri* (Group B, which consists of 14 classical serotypes and subserotypes), *Shigella boydii* (Group C, which has 20 serotypes), and *Shigella sonnei* (Group D, consisting of a single serotype) (Levine *et al.*, 2007).

1.1.1 SHIGELLOSIS

Shigellosis is a particularly contagious disease, with as few as 100 organisms required for transmission (DuPont *et al.*, 1989). Symptoms of shigellosis range from mild diarrhoea, to severe, bloody mucoid stools. In severe but rare cases, neurological symptoms including seizures (Ashkenazi *et al.*, 1987), lethargy, headaches and convulsions may be present (Ashkenazi *et al.*, 1990). Despite the fact that shigellosis is often a self limiting disease which is cleared in healthy individuals within 5 – 7 days (Niyogi, 2005), in some cases it may progress to serious life threatening conditions, including haemolytic uraemia (Koster *et al.*, 1978), intestinal perforation and rectal prolapse (Bennish, 1991). In developing countries, teaming shigellosis with factors such as hypothermia, hypoglycaemia, bronchopneumonia and

altered consciousness were found to be predictive for death (Bennish *et al.*, 1990; van den Broek *et al.*, 2005).

1.1.2 EPIDEMIOLOGY

Diarrhoeal disease has proven to be a significant global crisis, with an estimated 1 billion illness episodes and 2.5 million deaths in children occurring each year in the 1990s in developing countries (Kosek *et al.*, 2003). Shigellosis contributes to this disease burden, resulting in approximately 164.7 million cases a year; from these cases, 1.1 million deaths occurred in developing countries (Kotloff *et al.*, 1999). A total of 69% of cases reported in developing countries occurred in children and 61% of these cases resulted in fatality (Kotloff *et al.*, 1999). More recently, an extensive shigellosis study conducted in six Asian counties indicated that the annual incidence of treated shigellosis was 2.1 per 1000 per year, and 13.2 per 1000 in children under 5 years of age (von Seidlein *et al.*, 2006). In industrialized countries, the rate of incidence is much lower, estimated to be 3.7 per 100 000 in the United States in 1999 (Gupta *et al.*, 2004) and 3.2 per 100 000 in the Netherlands between 1996 and 2000 (van Pelt *et al.*, 2003). The spread of *Shigella* is assisted by overcrowded conditions and poor sewerage management (Hale, 1991). While *S. dysenteriae* predominantly causes epidemic dysentery, *S. flexneri* is the principal subgroup responsible for endemic shigellosis, causing more mortality than other *Shigella* strains in developing countries (Bennish, 1991; Jennison and Verma, 2004). In the United States, *Shigella* infection is the third most common cause of bacterial gastroenteritis, with 18.4% of isolates typed as *S. flexneri* (Gupta *et al.*, 2004). Shigellosis is also endemic in central Australia, with *S. flexneri* 6, *S. flexneri* 2a and *S. sonnei* mainly causing infection (Albert *et al.*, 1990) and this is predominantly acquired via person-to-person transmission (Ashbolt *et al.*, 2002).

1.1.3 *S. FLEXNERI* INVASION

S. flexneri is an infectious organism (DuPont *et al.*, 1989; Jennison and Verma, 2004) with the initial step of *S. flexneri* infection being invasion of the basolateral colonic epithelial cells, primarily achieved via entry into endocytic membranous epithelial cells (M cells). The model of pathogenesis is summarised in Figure 1.1. M cells are located in the follicle-associated epithelium (FAE), which reside over lymphoid follicles in the small and large intestines (Clark and Hirst, 2002). These cells sample luminal antigens and transport them to a self-produced pocket formed by their basolateral membrane (Gebert *et al.*, 1996; Jennison and Verma, 2004; Kraehenbuhl and Neutra, 2000), containing macrophages and lymphocytes, equipped to initiate a mucosal immune response (Neutra *et al.*, 1996). In the *Shigella* infection model, bacterial cells are internalised by M cell vacuoles and released into the intra-epithelial pocket (Jennison and Verma, 2004). Bacteria engulfed by resident macrophages evade killing mechanisms and induce apoptosis, causing the release of chemokines such as interleukin 1 (IL-1) and initiating an inflammatory response which recruits polymorphonuclear cells (PMNs) (Sansonetti and Phalipon, 1999). Epithelial cells undergoing invasion also produce pro-inflammatory cytokines such as interleukin 8 (IL-8), which results in the recruitment of further PMNs to the subepithelial area (Sansonetti *et al.*, 1999). The recruitment of PMNs causes the disruption of epithelial integrity, hence facilitating the influx of further luminal *Shigella* to the basolateral side of the FAE via an M cell-independent process (Perdomo *et al.*, 1994). Once the basolateral side of the membrane has been reached, *Shigella* are able to invade the epithelial cells (Mounier *et al.*, 1992). Rearrangement of the host cell cytoskeleton occurs, resulting in engulfment of the bacterium (Jennison and Verma, 2004; Sansonetti and Phalipon, 1999) and lysis of the encapsulating endocytic vacuole (Sansonetti *et al.*, 1986). Using the host cell actin assembly machinery, a propulsive force is created, which is termed actin based motility (ABM). ABM drives bacteria through the cytoplasm until contact with the cell membrane has been established, and form protrusions to

NOTE:
This figure is included in the print copy of the thesis held in the University of Adelaide Library.

Figure 1.1: Pathogenesis of *Shigella*

Shigella invade via M cells and infect resident macrophages, inducing cell death by apoptosis. *Shigella* move intra- and intercellulary via formation of F-actin tails, and *Shigella* are engulfed by adjacent cell by bacterial mediated endocytosis. *Shigella* lyse the vacuole in order to escape to the cytoplasm and spread to adjacent cells. Figure adapted from Tran (2008).

neighbouring epithelial cells (Monack and Theriot, 2001). *Shigella* are capable of lysing the membranes that contain them, and continue replicating and spreading intra- and inter-cellularly.

1.2 *S. FLEXNERI* VIRULENCE

S. flexneri expresses many virulence factors to elicit disease, and in this section, relevant virulent factors are described, including the large virulence plasmid, the Mxi-Spa Type III secretion system (TTSS) and its effector proteins, and the IcsA protein. Another crucial virulence factor is lipopolysaccharides (LPS) which is also described in detail the following section.

1.2.1 LARGE VIRULENCE PLASMID

In *S. flexneri*, the ability to convey pathogenicity is characterised by the capacity to invade and multiply within the colonic mucosa (Labrec *et al.*, 1964). This ability is principally attributable to the possession of a large virulence plasmid (Sansonetti *et al.*, 1982; Sasakawa *et al.*, 1993). The ~230 Kb plasmid, lacking in all *S. flexneri* avirulent strains (Sansonetti *et al.*, 1982), encodes many of the key virulence factors necessary to elicit invasiveness, most of which are contained within a ~31 Kb pathogenicity island (PAI) (Maurelli *et al.*, 1985; Sasakawa *et al.*, 1988). The PAI contains genes that encode the Mxi-Spa type III secretions system (TTSS), which translocates effector molecules from the bacterial cytoplasm to the membrane and cytoplasm of the host cell (Galan and Collmer, 1999; Jouihri *et al.*, 2003), and contains genes encoding the effector outer surface proteins (OSP) and Ipa secreted by the TTSS, and the Ipg effector protein chaperones. Sequencing of the large virulence plasmid indicate that multiple gene rearrangements of the plasmid have

taken place in *S. flexneri* evolutionary history, as insertion sequences (IS) elements account for 30% of the complete sequence of the plasmid (Zhang *et al.*, 2003).

1.2.2 MXI-SPA TYPE III SECRETION SYSTEM

TTSSs are found in many pathogenic Gram-negative species and are evolutionary and structurally related to the export system of flagella (Macnab, 1999). The Mxi-Spa TTSS is comprised of a collection of proteins that assemble into a structure spanning both the inner membrane (IM) and outer membrane (OM) (Cornelis, 2006) (Figure 1.2), and extending from the OM by 60nm in *Yersinia enterolictica*, 55nm in *Salmonella typhi* and 50nm in *Shigella* (Blocker *et al.*, 1999; Kubori *et al.*, 1998). More specifically, the TTSS complex assembles into a base structure, and a hollow external needle enclosing a ~25 A channel projected into the extracellular environment (Blocker *et al.*, 1999). The base structure consists of a cytoplasmic bulb and a membrane-spanning pair of stacked rings joined by a central rod (Blocker *et al.*, 1999; Schuch and Maurelli, 2001a).

The main proteins which assemble to form the TTSS are MxiH, MxiI, MxiD, MxiG, MxiJ and MxiM. MxiH is the main component of the needle structure (Blocker *et al.*, 1999; Blocker *et al.*, 2001; Tamano *et al.*, 2000), and is the first protein to be secreted through the TTSS (Veenendaal *et al.*, 2009). The MxiI protein is a minor periplasmic rod component (Zenk *et al.*, 2007). The lipoprotein MxiM is anchored to the inner face of the OM and represents a class of secretory proteins (Hardie *et al.*, 1996; Koster *et al.*, 1997), which are proteins promoting stability for secretins during assembly processes and promote secretin insertion into the OM (Schuch and Maurelli, 2001b). MxiM stabilises MxiD, a member of the secretin family (Schuch and Maurelli, 2001a), which has the capacity to multimerise into stacked OM rings localised to the outer rim of the upper ring of the needle complex (Zenk *et al.*, 2007). The inner membrane proteins MxiG and MxiJ form the base of the needle complex. Spa33, an essential C ring component necessary for TTSS needle formation, is

NOTE:
This figure is included in the print copy of the thesis held in the University of Adelaide Library.

Figure 1.2: The architecture of the Type III secretion system

Indicated in this figure are the Mxi proteins comprising the Type III secretion system (TTSS) and the Spa proteins comprising the ‘bulb’ region. Diagram adapted from the Blocker laboratory, University of Bristol.

localised beneath the TTSS and interacts with MxiG and MxiJ (Morita-Ishihara *et al.*, 2006). Spa40 and Spa24 are predicted components of the TTSS inner membrane export apparatus (Zenk *et al.*, 2007), and Spa32 controls the length of the needle most likely by switching the specificity of the TTSS from secretion of needle components, to secreting the Ipa proteins (Magdalena and Goldberg, 2002; Tamano *et al.*, 2002). Spa47 is the ATPase energiser of the system (Jouihri *et al.*, 2003), while Spa22 is proposed to control Ipa localisation within the TTSS. The lipid environment of the eukaryotic cell membrane triggers the formation of a pore complex around the tip of the needle structure (van der Goot *et al.*, 2004) and once contact between bacteria and host is established, the TTSS is induced and begins inserting the Ipa invasins into the cytoplasm of the host cell (Blocker *et al.*, 1999).

1.2.3 IPA, IPG PROTEINS AND OTHER EFFECTOR PROTEINS SECRETED BY THE TTSS

Also present on the large virulence plasmid are the *ipa* genes, encoding the Ipa invasins which are important effectors for *S. flexneri* entry and invasion (Buysse *et al.*, 1987). IpaA – D, are essential for *in vitro* epithelial cell invasion (Menard *et al.*, 1993; Sasakawa *et al.*, 1988) and characterisation of *ipa* mutants indicated that IpaB, IpaC, and IpaD are essential for entry into epithelial cells (Menard *et al.*, 1993; Sasakawa *et al.*, 1989). Ipa B and D proteins act as a secretion plug and appear to be located at the TTSS needle tip (Olive *et al.*, 2007). IpaB and IpaC proteins serve as the membrane pore section of the type III machinery. It has been shown that Ipa secretion is crucial for invasiveness, even if the bacteria is capable of full production of Ipa B, C and D proteins (Allaoui *et al.*, 1992; Allaoui *et al.*, 1993; Venkatesan *et al.*, 1992). Ipa proteins are synthesised and stored within the bacteria, and associate with chaperone proteins until secretion is initiated by contact with the host cell (Menard *et al.*, 1994). Once contact with epithelial cells has been engaged and detected, IpaB and IpaC promote invasion by interacting with $\alpha 5\beta 1$ integrin (Watarai *et al.*, 1996) and CD44

(Skoudy *et al.*, 2000). IpaB and IpaC independently bind to IpgC, a specific chaperone that is encoded by the gene located immediately upstream from *ipaB* (Menard *et al.*, 1994). IpaC is not secreted, but acts as chaperone (Menard *et al.*, 1994). Once secretion is achieved, the N-terminus of IpaC binds IpaB (Harrington *et al.*, 2003). IpaC and IpaB are hydrophobic and this characteristic facilitates insertion into the host cell membrane to form a pore (Blocker *et al.*, 1999). IpaC activates the host cell Rho GTPases, triggers actin polymerisation and filopodial extension in the vicinity of bacteria (Tran Van Nhieu *et al.*, 1999). IpaA is secreted into the cytosol of the host epithelial cell where it binds the cytoskeleton associated protein vinculin. This complex depolymerises actin filaments, organising an entry foci around the bacterium (Bourdet-Sicard *et al.*, 1999). Other effector proteins secreted by the TTSS are IcsB, IpaH, VirA and OspG. IcsB plays a crucial role in escaping from autophagy (Ogawa *et al.*, 2005), and OspG binds UbcH5, and is involved in modulating the inflammatory response of the host (Kim *et al.*, 2005; Okuda *et al.*, 2005). The IpaH proteins are also involved in modulating the inflammatory response of the host, and are encoded both on the large virulence plasmid and the chromosome (Ashida *et al.*, 2007). VirA induces destruction of local microtubule structures, and also promotes the actin-based motility of bacteria within the host cell cytoplasm (Yoshida *et al.*, 2006; Yoshida *et al.*, 2002). Another effector IpgD, is injected into epithelial cell by TTSS where it acts as a phosphoinositide phosphatase, uncoupling the plasma membrane from the actin cytoskeleton, allowing membrane extensions to form (Niebuhr *et al.*, 2002). IpaB is localised to the cytosol of macrophages and affinity purification illustrates the binding to IL-1β converting enzyme/caspase 1 (Chen *et al.*, 1996). Caspase activation triggers apoptosis in macrophages, initiating a strong inflammatory response and increasing the permeability of the epithelial barrier to *Shigella* entry as described earlier.

1.2.4 ICSA AND ACTIN-BASED MOTILITY

The ability to utilise ABM is a key determinant in establishing invasiveness and executing disease (Goldberg, 2001). Intracellular movement is reliant upon polarised actin polymerisation and when concentrated at one pole of the bacterium results in formation of F actin tails and leads to unidirectional propulsion (Cossart, 2000; Robbins *et al.*, 2001). IcsA, formerly referred to as VirG, is a 120 kDa outer membrane protein essential for this intra- and intercellular spread (Bernardini *et al.*, 1989; Goldberg *et al.*, 1993). IcsA is a member of the autotransporter family composed of 1101 amino acids, and is comprised of three domains; the N-terminal signal sequence, the α-domain, and the C-terminal β-domain (Goldberg *et al.*, 1993; Lett *et al.*, 1989; Suzuki *et al.*, 1995). The α-domain is the functionally active region of IcsA, while the C-terminus is predicted to form a β barrel that anchors the protein in the outer membrane, exposing the N-terminus to the bacterial surface (Robbins *et al.*, 2001). The α-domain contains six glycine-rich repeat regions, which have been shown to be essential in stimulating actin assembly of *S. flexneri*. These regions have been shown to be important in interacting with the neural Wiskott–Aldrich syndrome protein (N-WASP) and other host proteins such as vinculin (Egile *et al.*, 1999; May and Morona, 2008; Suzuki *et al.*, 1998; Suzuki *et al.*, 1996). The recruitment of N-WASP is necessary for F actin comet tail formation in mammalian cells (Egile *et al.*, 1999; Suzuki *et al.*, 1998). Furthermore, N-WASP initiates actin-related protein (Arp) 2/3 complex-mediated actin polymerisation (Egile *et al.*, 1999; Suzuki *et al.*, 2002).

1.3 LIPOPOLYSACCHARIDES

1.3.1 *SHIGELLA* LPS AND VIRULENCE

LPS is a virulence factor that is critical in establishing disease and bacterial survival in animal hosts, contributing to evasion of host immune responses, particularly the alternative

complement cascade (Raetz and Whitfield, 2002). *S. flexneri* rough strains display avirulence in animal models and are deficient in cell to cell spread and are incapable of producing a positive Sereny test (Okamura *et al.*, 1983). Rough mutants also demonstrate a significant deficiency in tissue culture monolayer plaque formation, with very small or no plaques being observed (Hong and Payne, 1997; Sandlin *et al.*, 1995; Van den Bosch *et al.*, 1997; Van den Bosch and Morona, 2003). A number of genes associated with LPS biosynthesis affect virulence of *S. flexneri* (Okada *et al.*, 1991a; Okada *et al.*, 1991b). Recent studies have also demonstrated a clear link between underacylated lipid A and reduced endotoxic potential (Ranallo *et al.*, 2010). MsbB, an acyltransferase involved in lipid A biosynthesis, is a critical virulence gene. *S. flexneri* 2a *msbB* mutants were attenuated in an acute mouse pulmonary challenge model, and double mutants in *msbB1* and *msbB2* resulted in bacteria being defected in the ability to invade, replicate, and spread within epithelial cells (Ranallo *et al.*, 2010). These studies show that loss of either *msbB* gene resulted in reduction of LPS endotoxicity.

1.3.2 MODAL CHAIN LENGTH AND VIRULENCE

O antigen (Oag) is a significant determinant in virulence, however the presence of Oag is not the sole determinant of LPS for providing virulence and evasiveness; rather, the average Oag chain length plays a critical role in establishing pathogenesis. In *S. flexneri*, the protein responsible for the regulation of the Oag modal chain length is Wzz (formerly known as Cld or Rol), a member of the polysaccharide co-polymerase family (see section 1.4.2). *S. flexneri* Wzz mutants are unable to form plaques on HeLa cell monolayers or form F-actin comet tails, indicating that regulation of wild-type Oag modal chain length is important in cell-to-cell spread (Morona *et al.*, 2003). Previous studies have indicated that short type LPS (11-17 repeat units) produced in wild-type invasive *S. flexneri* is necessary for maintaining unipolar IcsA localisation to ensure efficient ABM and intercellular spread (Morona and Van Den Bosch, 2003; Robbins *et al.*, 2001; Sandlin *et al.*, 1995). The chromosomally encoded wzz_{SF}

has been shown to be necessary for IcsA function, as when wzz_{PHS-2} is the sole determinant of LPS modal length, hence producing VL-type Oag chain lengths, the mutants are unable to form plaques on HeLa cells, have reduced virulence in a Sereny assay, reduced levels of IcsA on its cell surface and a significantly reduced ability to form F-actin comet tails (Van den Bosch *et al.*, 1997). It is believed that VL-type Oag chain can prevent and sterically hinder IcsA function in actin based motility (Morona and Van Den Bosch, 2003). The function of the VL-type Oag chains appears to be in serum resistance (Hong and Payne, 1997). Mutations in the *wzz* of *Salmonella typhimurium* (wzz_{ST} and wzz_{FepE}) result in strains which exhibit enhanced susceptibility to complement and are highly attenuated in a mouse model (Murray *et al.*, 2003).

1.3.3 STRUCTURE

Providing protection against host defences and affecting host-cell interactions, LPS is a surface glycolipid composed of three separate structural domains: 1) Lipid A, the hydrophobic anchor for LPS in the outer membrane, 2) core oligosaccharide, divided into the inner and outer core, and 3) the Oag polymer, an oligosaccharide repeat unit (Raetz and Whitfield, 2002) (Figure 1.3). There are four phenotypes of LPS morphology described - primarily, smooth LPS (S-LPS), consisting of complete LPS molecules with a non-random Oag chain length. LPS considered rough (R-LPS) consists of LPS molecules which lack the Oag repeat units. LPS deemed semi-rough (SR-LPS) contain only one Oag repeat attached to the lipid A-core region (Morona *et al.*, 1994; Naide *et al.*, 1965). The fourth LPS phenotype, smooth unregulated (SU-LPS) is observed as a loss of regulated modal length, and displays random chain length.

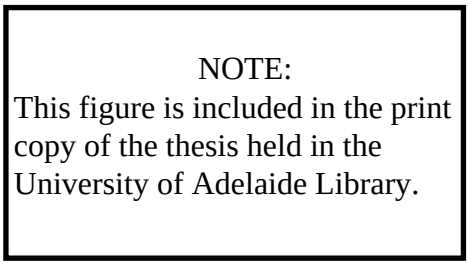


Figure 1.3: Schematic illustration of LPS

Indicated in this figure are the Kdo (2-keto-3-deoxy-D-manno-octulosonic acid), glucosamine (Gln), heptose (Hep) and galactose (Gal) residues comprising the lipid A-core region. The fatty acid residues are illustrated, and phosphorylated residues are indicated in magenta circles. Adapted from Yethon and Whitfield (2001).

1.3.4 LIPID A AND CORE

Functioning as the membrane anchor for LPS, lipid A is the most conserved element of this molecule. Lipid A, also referred to as 'endotoxin', is responsible for many biological effects exhibited by mammalian systems during sepsis, inducing toxic shock, generating fever and activating host lymphocytes and macrophages non-specifically (Bone, 1993; Raetz, 1993; Yethon and Whitfield, 2001). The innate immune response to lipid A via a protein kinase cascade leads to nuclear factor kappa B activity and cytokine production, which cause these biological effects (Muller *et al.*, 1993).

There are approximately 10^6 lipid A residues, 10^7 phospholipids and 10^5 undecaprenyl phosphate-sugar molecules in an *E. coli* cell (Galloway and Raetz, 1990). The typical lipid A backbone in *E. coli* and *S. flexneri* is a disaccharide of glucosamine residues linked β,1' - 6, and phosphorylated at positions 1 and 4' (Yethon and Whitfield, 2001). LpxA, C and D are soluble proteins whereas LpxB and LpxH are peripheral membrane proteins. LpxK, KdtA, LpxL and LpxM are integral inner membrane proteins. Their active sites are presumed to face the cytoplasmic surface of the inner membrane, given that their water-soluble co-substrates are cytoplasmic molecules. The biosynthesis of lipid A-core is initiated with the primary reaction of the sugar nucleotide UDP-N-acetylglucosamine fatty acetylation by the UDP-N-acetylglucosamine acetyltransferase, LpxA. *E. coli* LpxA requires the thioester *R*-3-hydroxymyristoyl acyl carrier protein (ACP) as its donor substrate (Anderson and Raetz, 1987; Wyckoff *et al.*, 1998) (Figure 1.4). The unfavourable equilibrium constant for the acylation of UDP-N-acetylglucosamine results in the deacetylation of this product by a zinc metalloenzyme, LpxC; at this stage, lipid A biosynthesis has been committed (Young *et al.*, 1995). Following deacetylation, a second β-hydroxymyristate moiety is incorporated by the UDP-3-O-(3-hydroxymyristoyl)-glucosamine N-acyltransferase LpxD, to generate UDP-2,3-diacylglucosamine. This product is cleaved at its pyrophosphate bond by the selective pyrophosphatase LpxH, which catalyses the effect of water on the α-phosphorus atom of the

NOTE:
This figure is included in the print copy of the thesis held in the University of Adelaide Library.

Figure 1.4: The pathway for lipid A synthesis

Indicated in this figure is the glucosamine disaccharide backbone of lipid A (blue), and the Kdo disaccharide (black). LpxH and LpxB are peripheral membrane proteins, and LpxA, LpxC and LpxD are cytoplasmic proteins and the later enzymes (initiated with LpxK) are integral inner membrane proteins (Raetz and Whitfield, 2002). Diagram adapted from Raetz *et al*. (2008).

UDP moiety to form 2,3-diacylglucosamine-1-phosphate (lipid X) (Raetz *et al.*, 2007). Lipid X is the direct precursor of the reducing sugar of lipid A, and a β, 1'-6 linked disaccharide is then formed by LpxB, which mediates the condensation of another molecule of UDP-2,3-diacylglucosamine with lipid X, releasing the UDP. The LpxK kinase then phosphorylates the 4' position of the disaccharide 1-phosphate generated by LpxB to form lipid IV_A (Raetz and Whitfield, 2002). *E. coli* LPS contains two 3-deoxy-D-*manno*-oct-2-ulosonic acid (Kdo) residues that are transferred to lipid IV_A by the bifunctional enzyme encoded by *waaA* (Clementz and Raetz, 1991; Raetz *et al.*, 2007). The labile sugar nucleotide CMP-Kdo is the Kdo donor (Meredith and Woodard, 2003; Raetz, 1990). The second Kdo residue is incorporated more rapidly than the first, and therefore the intermediate with a single Kdo residue does not have the chance to accumulate (Raetz *et al.*, 2007). The final steps of *E. coli* lipid A biosynthesis engage the addition of lauroyl and myristoyl residues by LpxL and LpxM to the distal glucosamine at the non-reducing end of lipid A, forming acyloxyacyl moieties (Raetz and Whitfield, 2002; Yethon and Whitfield, 2001).

All LPS core regions possess the Kdo residue that links the core to lipid A, however many core regions also possess L-*glycero*-D-*manno*-heptose (Hep) residues linked to the Kdo (Yethon and Whitfield, 2001). This region is classified as the 'inner core', i.e., the Kdo residues with or without the Hep residues, whilst the distal sugar region of the core is deemed the 'outer core', comprising of hexose, glucose and galactose residues. There are five classifications of outer core structures for *E. coli*, R1, R2, R3, R4 and K-12 and two core variants in *Salmonella enterica*, differing in the arrangement and content of sugar residues, and also in the position to which Oag is ligated (Figure 1.5). *E. coli* and *S. flexneri* predominantly possess R1 type core. The proteins involved in the synthesis of the core regions are encoded on the chromosomal *waa* locus (Raetz and Whitfield, 2002). The *waa* locus encodes 16 genes arranged in the order of *gmhD, waaF, waaC, waaL, waaV, waaW, waaY, waaT, waaO, waaP, waaG, waaQ* and *waaA.* Synthesis of inner core is catalysed by WaaA. Transfer of the first two Kdo residues of the inner core is catalysed by WaaA Kdo

NOTE:
This figure is included in the print copy of the thesis held in the University of Adelaide Library.

Figure 1.5: Structures of different core types and synthesis of R1 type core

A) Indicated in this figure are the enzymes associated with core synthesis and modification of the R1-type core region. The glycosyltransferases associated with forming the inner core background are indicated in orange, and enzymes which modify the structure are shown in blue. Denoted in green are the outer core glycosyltransferases, and the ligase enzyme is shown in pink. B) The five different core types identified in *E. coli*. The heptose (hep) residues are indicated in red, and the position of the ligated Oag is shown in purple. Diagrams adapted from Raetz and Whitfield (2002).

transferase, WaaA. *gmhD* encodes an ADP-L-glycero-D-mannoheptose-6-epimerase, a sugar biosynthesis protein which is required for the conversion of ADP-D-glycero-D-mannose heptose to ADP-L-glycero-D-mannose-heptose (Coleman, 1983), the donor unit required for the addition of heptose to the core (Sirisena *et al.*, 1994; Vimont *et al.*, 1997).

The products of *waaC* and *waaF* encode the first and second heptosyltransferase, respectively (Sirisena *et al.*, 1992). The proteins associated with inner core assembly phosphoryl transfer reactions are WaaP and WaaY. WaaP phosphate transferase attaches a phosphate to Hep I (Muhlradt, 1969; Yethon *et al.*, 1998) and WaaY adds a phosphate to HepII. The WaaQ functions as a HepIII transferase adding a HepIII to the HepII (Figure 1.5).

S. flexneri has the R1 type core, and the biosynthesis of the outer core continues with the WaaG protein, encoding a UDP-glucosyl-transferase, which is able to attach the first glucose, Glc I, to heptose II (Creeger and Rothfield, 1979). WaaO, an α-1,3 glucosyltransferase, uses UDP hexose to synthesise its core region and attach a second glucose unit to the first glucose unit, and WaaT further adds an α-1,2-Galacose residue to this second glucose (Heinrichs *et al.*, 1998). WaaW adds another α-1,2 Galactose residue to the first galactose position, and WaaV adds β-glucose to the second glucose (at this juncture, Oag is ligated to the core) (Figure 1.5).

1.3.5 O ANTIGEN GENETICS AND BIOSYNTHESIS

The Oag component of LPS is the most variable region and is the immunological dominant surface epitope in *S. flexneri*, *E. coli* and *S. enterica* (Yethon and Whitfield, 2001). Differences in the content of oligosaccharide units, the linkages between units and the number of sugars present in the Oag allow for many variations in Oag, and result in the immense serological diversity exhibited by bacterial species. Oaf diversity is important in allowing bacteria to present differences in surface structures which may offer selective advantage in specific niches (Liu *et al.*, 2008). Diversity is acquired via insertion elements, resulting in

inactivation of genes or introducing new genes into the Oaf gene cluster, hence forming new Oag. Also, mutations in Oag genes contribute to diversity (Liu *et al.*, 2008). The result of this diversity generates 12 traditionally accepted serotypes of *S. flexneri* (Figure 1.6). Oag modification also has contributed to this diversity; bacteriophages are known to contain Oag modification genes (Guan and Verma, 1998). These phages can alter the host bacterial Oag by acetylation, glycosylation or alteration of the mode of linkages between the repeat units (Guan and Verma, 1998). Studies have also shown that bacteriophage-encoded glucosylation of *Shigella* Oag shortens the LPS molecule and enhances TTSS function showing that LPS glucosylation promotes bacterial invasion and evasion of innate immunity (West *et al.*, 2005).

In *S. flexneri* Y, Oag is composed of a tetrasaccharide, consists of repeating tetrasaccharide units comprised of three L-Rhamnose residues and one residue of N-acetylglucosamine (Allison and Verma, 2000) (Figure 1.6). There are α linkages between rhamnose residues, and β linkages between rhamnose and N-acetylglucosamine. Serotypes differ by the addition of either glucosyl or *O*-acetyl groups to different sugars within the tetrasaccharide repeat unit via the linkages (Allison and Verma, 2000). The cluster of genes which encode for Oag biosynthesis and processing is located at 44 min near *his* on the chromosome (Formal *et al.*, 1970; Morona *et al.*, 1995), between the *galF* and *gnd* genes in *S. flexneri* and *E. coli*. The cluster consists of three sets of genes, the first set are involved in the synthesis of nucleotide sugar precursors of the Oag, while the second encode a collection of glycosyl transferase proteins that sequentially transfer the sugars to form an oligosaccharide on a lipid carrier. The third set encodes phosphoryl transfer reactions Oag processing proteins. This first set of genes in the Oag cluster begins with *rmlBDAC*, which encode enzymes involved in the synthesis of the nucleotide sugar dTDP-rhamnose. RmlA is a glucose-1-phosphate thymidyl transferase, while RmlB is a dTDP-glucose-4,6-dehydratase. Both RmlA and B act to convert glycose-1-phosphate to the dTDP-6-deoxy-D-xylo-4-hexulose intermediate. RmlC and D are a TDP-6-deoxy-D-glucose-3,5-epimerase and dTDP-6-deoxy-L-mannose dehydrogenase, respectively, and act sequentially on dTDP-6-deoxy-D-xylo-4-

NOTE:
This figure is included in the print copy of the thesis held in the University of Adelaide Library.

Figure 1.6 Oag structure

A) The Oag structure consists of the tetrasaccharide repeat unit, composed of rhamnose residues and an N-acetylglucosamine residue. O-acetyl and glucosyl groups added to different sugars in the Oag unit differ between serotypes. Serotypes have group-specific and type-specific antigen determinants, illustrated in Arabic and Roman numerals, respectively. Rhamnose (Rha), N-acetylglucosamine (GlcNAc) are indicated. Glucosyl groups are indicated by purple circles, and O acetyl groups are indicated by blue squares. This diagram is adapted from Allison and Verma (2000). B) The O ag biosynthesis operon of *S. flexneri*, and genes and open reading frames are indicated by arrows. Figure adapted from Daniels (1999).

hexulose to give dTDP-L-Rhamnose. UDP-N-acetylglucosamine is derived via exogenous means by the proteins encoded by the *glmU* region of the chromosome.

Once synthesized, the repeating unit structures of O polysaccharide are assembled on a membrane-bound carrier, undecaprenyl phosphate by the N-acetyl-glucosamine-1-phosphate transferase, WecA (Meier-Dieter *et al.*, 1992). In *Enterobacteriaceae*, *wecA* is situated outside of the Oag gene cluster and is part of the Enterobacterial Common Antigen (ECA) gene cluster (Alexander and Valvano, 1994; Lehrer *et al.*, 2007; Meier and Mayer, 1985; Samuel and Reeves, 2003; Yao and Valvano, 1994). Following this, further sequential incorporation of the remaining sugars to synthesise the Oag repeat unit occurs, via the action rhamnosyl transferases RfbF and RrfG, adding the remaining three rhamnose residues in order to complete the tetrasaccharide repeat unit.

1.3.6 O ANTIGEN PROCESSING

The Oag processing genes are located in the same cluster of Oag biosynthesis genes. There are two major polysaccharide polymerisation systems, Wzy-dependent and -independent processes (Raetz and Whitfield, 2002). The Wzy-dependent pathway occurs in most Oag synthesis pathway, especially that of heteropolymeric Oag (Guo *et al.*, 2008). Oag biosynthesis occurs on the cytoplasmic face of the inner membrane; once synthesised, the Oag unit is ‘flipped’ across the inner membrane by a flippase encoded by the *wzx* gene (Figure 1.7). Strains that lack Wzx result in an accumulation of undecaprenyl linked O units on the cytoplasmic face of the membrane (Liu *et al.*, 1996). Wzx proteins are hydrophobic membrane proteins which have 12 transmembrane segments exhibiting similar motifs, with high numbers of alpha-helical transmembrane segments (Cunneen and Reeves, 2008) and share structural features with bacterial permeases (Macpherson *et al.*, 1995). It has been theorised that Wzx transits undecaprenyl phosphate linked Oag units to the periplasmic face of the membrane via a proton or electrochemical gradient as an energy source (Guo *et al.*,

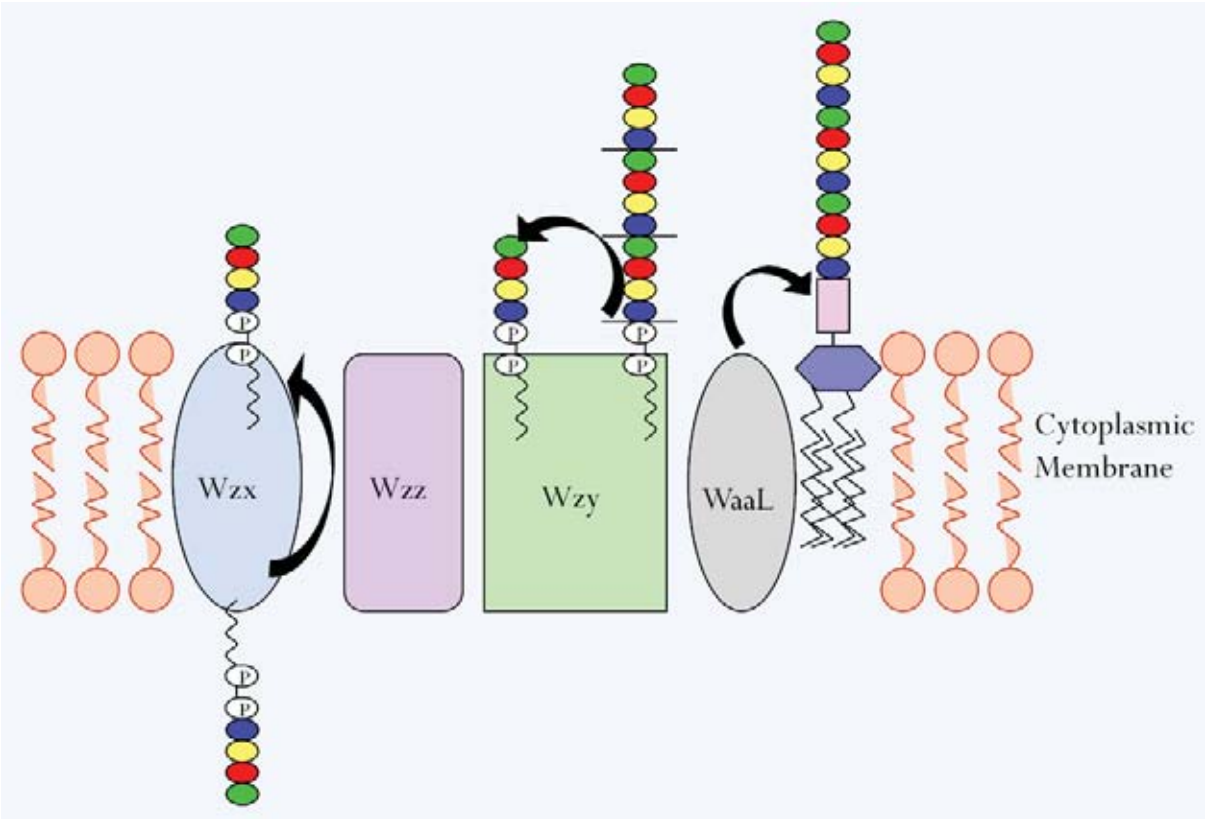


Figure 1.7: Model for Oag processing

A model for the events occurring during Oag processing. The individual undecaprenyl-linked Oag repeat units are flipped across the membrane by Wzx, and are polymerised by Wzy in the periplasm. The growing nascent chain is transferred from the undecaprenyl carrier to the non-reducing terminus of the new subunit. Oag chain modal length is determined by Wzz. The polymer is ligated to lipid A-core by WaaL and transported to the outer membrane.

2008). A lack of Wzx specificity for the chemical structure of the Oag has been demonstrated, however Wzx seems to recognise the primary undecaprenyl linked sugar (Feldman *et al.*, 1999; Guo *et al.*, 2008). Once the O units are translocated to the periplasmic face, they are linked by the Oag polymerase, Wzy (Collins and Hackett, 1991). The reaction involves transfer of nascent polymer from its undecaprenyl phosphate carrier to the non-reducing end of the new undecaprenyl phosphate linked Oag repeat unit (Bray and Robbins, 1967). This results in a chain length increase by one new repeat unit added at a time to the reducing end of the nascent polymer (Figure 1.7). *wzy* mutants produce SR-LPS consisting of a lipid A-core capped with a single O unit, as further O units are not adjoined to the polymer chain (Collins and Hackett, 1991). Wzy proteins are predicted to be integral membrane proteins with 11–13 transmembrane domains, and like the Wzx proteins, they exhibit little primary sequence similarity (Morona *et al.*, 1994; Raetz and Whitfield, 2002). Wzy proteins also exhibit a high specificity for the Oag repeat unit structure (Yi 2006), and recent studies have been able to chemically reconstitute polysaccharide synthesis *in vitro* (Woodward *et al.*, 2010). The subsequent ligation step, catalysed by the ligase WaaL, is common to all polymerization pathways, and results in the transfer of the Oag repeat unit to the nascent lipid A-core. The polysaccharide co-polymerase (PCP) protein Wzz regulates this number of O polysaccharides (section 1.4).

The biosynthesis and processing pathway for Oag is similar to that for the synthesis of other bacterial polysaccharides, such as capsular polysaccharides (CPS) and exopolysaccharides (EPS). In capsular biosynthesis, undecaprenyl-linked repeat units are assembled at the interface between the cytoplasm and the inner membrane (Whitfield, 2006). Synthesised undecaprenyl-linked repeats are flipped across the membrane in a process requiring Wzx, providing the substrates for Wzy-dependent polymerisation, as the polymer grows by transfer of the growing chain to the incoming undecaprenyl-linked unit. Polymerisation requires phosphorylation of C-terminal tyrosine residues in the Wzc oligomer (Grangeasse, 2002; Paiment *et al.*, 2002), and subsequent dephosphorylation by the Wzb

phosphatase (Vincent *et al.*, 1999). The polymer is translocated by Wza, which is believed to act as a channel (Reid and Whitfield, 2005).

1.3.7 LPS EXPORT

Once assembled, LPS needs to be exported from the inner membrane across the periplasm, and assembled into the outer leaflet of the outer membrane. It has been shown that MsbA participates in the transportation of newly synthesized core-lipid A molecules from the cytoplasmic to the periplasmic surface of the inner membrane (Doerrler *et al.*, 2004). Seven lipopolysaccharides transport (Lpt) proteins mediate the final steps in LPS assembly (Ruiz *et al.*, 2008; Sperandeo *et al.*, 2009). Four proteins mediating LPS transport (LptB, LptC, LptF, and LptG) are proposed to form a complex that associates ATP hydrolysis with the release of LPS from the inner membrane (Ruiz *et al.*, 2008; Sperandeo *et al.*, 2007; Sperandeo *et al.*, 2008), and this purified complex has been shown to exhibit ATP hydrolytic activity (Narita and Tokuda, 2009). Two proteins LptD and LptE (Imp and RlpB, respectively) form a complex which mediates the correct insertion of LPS into the outer leaflet of the outer membrane and LptE is believed to receive LPS from the periplasm, as it binds LPS specifically (Chng *et al.*, 2010b). LptA, a periplasmic protein, is theorized to assist exportation of LPS across the periplasm, perhaps coordinating with the functions of the two complexes (Ma *et al.*, 2008; Sperandeo *et al.*, 2007; Suits *et al.*, 2008). Recent data indicates that the Lpt proteins directly interact and may form a physical trans-envelope complex (Chng *et al.*, 2010a).

1.4 POLYSACCHARIDE CO-POLYMERASES (PCPS)

1.4.1 THE POLYSACCHARIDE CO-POLYMERASE (PCP) FAMILY

Members of the polysaccharide co-polymerase (PCP) family are regulators in the biosynthesis of cell surface polysaccharides (Morona *et al.*, 2009; Morona *et al.*, 2000; Tocilj *et al.*, 2008). Many of the PCPs have coiled coil-regions with a correlation in size and the resulting function in the degree of polymerisation or determining modal chain length (Morona *et al.*, 2000; Purins *et al.*, 2008). PCPs are categorised into three distinct classes, and are grouped based on characteristics including coiled-coil prediction profiles, sizes, the type of polysaccharide synthesised and sequence similarity (Cuthbertson *et al.*, 2009; Morona *et al.*, 2009). PCPs are inner membrane proteins, and have two transmembrane (TM) regions, TM1 and TM2, located close to the N-terminus and C-terminus, respectively. The area between TM1 and TM2 is a large hydrophilic region located on the periplasmic face of the inner membrane.

The three PCP classes are PCP1, PCP2 and PCP3 (Cuthbertson *et al.*, 2009; Morona *et al.*, 2009). The PCP1 proteins are associated with either regulation of LPS Oag, or biosynthesis of the ECA. Members of this class include the chromosomally encoded *S. flexneri* Wzz (Wzz_{SF}), the plasmid encoded *S. flexneri* Wzz_{PHS-2}, *Salmonella typhimurium* Wzz (Wzz_{ST}), *Escherichia coli* K-12 FepE, and *E. coli* Wzz_{ECA} (Morona *et al.*, 2009). Wzz_{SF} and Wzz_{ST} exhibit a high degree of sequence identity, greater than 70%, however generally the PCP1 proteins exhibit low sequence homology e.g., Wzz_{SF} and Wzz_{PHS-2} share ~22% identity (Morona *et al.*, 2009; Morona *et al.*, 2000). PCP2 proteins are involved in the synthesis of high molecular mass polysaccharides, such as CPS and EPS. Both PCP1 and PCP2 classes are associated with the Wzy-dependent mechanism of polysaccharide polymerization and processing, while the PCP3 group proteins are involved in polymerization utilising ABC-2 type transporters (Morona *et al.*, 2000; Reizer *et al.*, 1992). The PCP2 family is further assorted into PCP2a and PCP2b classifications; PCP2a proteins exhibit a

characteristic carboxy terminal protein tyrosine kinase (PTK) binding domain, a larger hydrophilic region between TM1 and TM2 than that displayed by the PCP1 proteins, while PCP2b proteins exhibit a smaller hydrophilic region between TM1 and TM2 and are associated with PTK binding cytoplasmic proteins deemed 'C components' (Morona *et al.*, 2000). There is low to moderate sequence identity between the proteins of this family, ranging from 25 to 55% (Morona *et al.*, 2000). PCP2a proteins include *E. coli* Wzc, involved in the biosynthesis of Colanic acid CPS, *Bradyrhizobium japonicum* ExoP, involved in EPS synthesis, and *Sinorhizobium meliloti* ExoP, involved in succinoglycan EPS synthesis . PCP2b proteins include *Streptococcus salivarius* CpsC and the ATP domain CpsD, involved in EPS polysaccharide biosynthesis, and *Streptococcus pneumoniae* Cps19aC and the ATP domain CPS19aD involved in CPS19A polysaccharide synthesis (Morona *et al.*, 2009). The PCP3 proteins are involved in CPS production, are smaller than the PCP2a proteins and lack the PTK-binding domain. Members of this family include *Neiserria meningitidis* CtrB and *E. coli* KpsE(K5).

A significant feature of PCP proteins is the conserved motif '$PX_2PX_4SPKX_1X_{10}GGMXGAG$' proximal to and partly overlapping the TM2 region, rich in proline and glycine residues (Becker *et al.*, 1995; Morona *et al.*, 2000). Within this proline-glycine conserved motif lies GXXXG/A, a motif shown to be involved in mediating interactions between TM regions and membrane protein folding (Senes *et al.*, 2004; Tocilj *et al.*, 2008). This conserved region has been shown to significantly influence the LPS Oag modal chain length distribution (Daniels and Morona, 1999).

1.4.2 MODAL CHAIN LENGTH REGULATION BY PCPS

In *S. flexneri*, the number of Oag repeats attached to the lipid A core is non-randomly distributed and contains approximately 11-17 repeats, otherwise known as short (S-type) LPS (Morona *et al.*, 1995). The protein controlling this non-random distribution of chain length is

Wzz (Wzz_{SF}). In a number of *S. flexneri* strains, there is an additional *wzz* (termed wzz_{PHS-2}) encoded on the small plasmid pHS-2 (Stevenson *et al.*, 1995). Wzz_{PHS-2} is responsible for LPS with Oag modal length of ≥ 90 repeats (VL-type). There is much variation in the resulting Oag modal chain lengths due to regulation by the diverse Wzz proteins in *Enterobacteriacea*.

Previous studies assessing the relationship between Wzz structure and function have been conducted; Franco *et al.* (1998) attempted to mimic the phenotypic variation observed in the different Wzz proteins by examining effects of amino acid changes which may be held responsible for determining a particular Oag modal chain length (Franco *et al.*, 1998). Franco *et al.* (1998) predicted that individual amino acid variations within *E. coli* and *S. flexneri* Wzz proteins may be held accountable for the alteration of the resulting Oag modal chain length, and explored the involvement of chain length determination of the region around residue 220, which exhibits variation (Franco *et al.*, 1998). Chimeric *wzz* genes were constructed, replacing this segment in Wzz from an intermediate (I –type, 10-18 RUs) Oag modal chain length *E coli* O2:K1:H6 with a Wzz segment from an S-type (7-16 RUs) producing strain O2:K1:H5, resulting in a decrease in Oag modal chain length from intermediate to S type (Franco *et al.*, 1998). Other notable residues present in Wzz proteins which confer S type LPS and are absent in I-type LPS strains are G221 and I224; using site directed mutagenesis, intermediate-type strains were altered at either of these positions to assess whether these individual residues are involved in chain length determination. The additional glycine failed to alter the resulting Oag length, however the isoleucine to valine mutation alone was sufficient to shift the intermediate type Oag from I-type to S-type (Franco *et al.*, 1998). *E. coli* strains which exhibit long–type (L-type, 16-25 RUs) Oag modal chain length differ substantially in the amino terminal region of Wzz, prompting the construction of a chimeric *wzz* featuring the initial 424 base pairs of a long type *wzz* gene cloned into an intermediate *wzz* gene, resulting in a chimeric Wzz protein (Franco *et al.*, 1998). This chimeric protein conferred a typical L-type Oag chain length of 16-25. *S. enterica* Wzz, producing L-type Oag

modal length of 19-30 repeat units, shares several similar residues in the amino terminal region as does L- type-producing Wzz from *E. coli*, and these residues were substituted (M77I, Q83S, D90E, and L91I) in the intermediate *E. coli* Wzz in order to ascertain whether the point mutations would shift the modal length from I-type to L-type. The results indicated that only D90E shifted the mutations midway between long and intermediate mode, that L91I reduced the modal length, and the first two alterations had no effect. Taken together, these studies indicate that the resulting function is a complex and intricate consequence of many critically positioned amino acids (Franco *et al.*, 1998). This led to the suggestion that Wzz function is not likely the result of one particular region of the protein, but rather the result of the overall structure (Franco *et al.*, 1998).

Klee *et al.* (1997) made the observation that although *E. coli* Wzz and *S. typhimurium* Wzz differed at nine amino acid positions, the major differences were amino acids 267 and 270 (Klee *et al.*, 1997); in *S. flexneri*, this residue is a basic lysine, whereas in *S. enterica*, this position is held by a polar asparagine residue (Klee *et al.*, 1997). Site-directed mutagenesis studies targeting a number of *S. flexneri* conserved residues, singularly or in combination have also been conducted (Daniels and Morona, 1999). Daniels and Morona (1999) constructed a K267N substitution via site-directed mutagenesis, which resulted in an increase in wild-type Wzz_{SF} 10-17 RUs Oag modal chain length to 13-20 RUs, therefore indicating that this residue is important in Oag chain length variation, but does not singularly drive Oag modal length to L-type (Daniels and Morona, 1999).

The TM regions exhibit a high level of conservation, and have previously been suggested to be involved in protein-protein interacting role (Bastin *et al.*, 1993), hence the glycine and proline residues in TM2 were also investigated by Daniels and Morona (1999). Proteins containing the substitution of glycine 305, 306, 309 and 311 to alanines were constructed, with the observations that single residue alterations did not alter the Oag chain length, however the dual change of G305A/G311A resulted in a significant reduction in the modal length, from S type (11-17) down to Very Short (VS) type (3-8 RUs) (Daniels and

Morona, 1999). The three proline residues were also substituted with alanines, with the interesting observation that two of the three residues (P286A and P292A) affected the regulation of Oag significantly, with P286A resulting in a reduced level of Oag chains, while the P292A mutation resulted in a complete loss of Wzz function (Daniels and Morona, 1999). Hybrid Wzz proteins between Wzz_{SF} and Wzz_{ST} were also constructed; both coding regions possess *Bgl*II sites, and it was used to fuse the proteins to create $Wzz_{ST/SF}$ (which had the N-terminal region of Wzz_{ST}, and the C-terminus of Wzz_{SF}) and $Wzz_{SF/ST}$ (which had the N-terminal region of Wzz_{SF} and the C-terminal region of Wzz_{ST}). $Wzz_{SF/ST}$ resulted in production of LPS with a modal length of 17–26 repeats, closer to that displayed by Wzz_{ST} (Daniels and Morona, 1999). $Wzz_{ST/SF}$ resulted in LPS with Oag modal chain length of 14–19 RUs, similar but longer than that produced by Wzz_{SF} (Daniels and Morona, 1999). These data indicate that residues involved in determination of chain modal length may be located in the C-terminal region of Wzz.

Daniels and Morona (1999) reported that Wzz_{SF} forms a dimer *in vivo* (Daniels and Morona, 1999), and appeared to implicate the TM1 in Wzz_{SF}-Wzz_{SF} interaction. This study also indicated that not only can dimerisation occur, but also that Wzz may oligomerise up to a hexameric size. The ability of Wzz to form large complexes raises the possibility of complex formation with other enzymes involved in Oag processing, such as WaaL or Wzy (Bastin *et al.*, 1993; Morona *et al.*, 1995).

1.4.3 3D STRUCTURES OF PCP PROTEINS

Recently, the periplasmic domain structures of several PCP proteins including *Salmonella typhimurium* Wzz*, E. coli* O157 FepE and WzzE have been solved, showing that these structures show extreme similarities at the protomer and oligomer level (Tocilj *et al.*, 2008). These protomers are elongated and consist of two structural components, a trapezoidal α/β base domain close to the membrane, and an extended α-helical hairpin, containing a ~100-

A-long helix forming anti parallel coiled coil interactions with two helices that fold back towards the membrane (Tocilj *et al.*, 2008) (Figure 1.8). The protomers self assemble into bell shaped oligomers displaying comparable structural features, with Wzz_{ST} forming pentameric oligomers, WzzE assembling into octameric oligomers and FepE assembling into nonameric structures (Tocilj *et al.*, 2008) (Figure 1.9). Recent studies from Larue *et al.* (2009) reported that Wzz_{ST}, Wzz_{FepE} and Wzz_{K40} favour hexameric structures (Larue *et al.*, 2009) and previous studies on the oligomeric status of *S. flexneri* Wzz (Wzz_{SF}) via cross-linking with formaldehyde indicated that Wzz_{SF} has the ability to form hexamers and high order oligomers, suggesting that oligomerisation is important in function (Daniels and Morona, 1999). Studies assessing cross-complementation of the Wzx flippase have also been conducted; in these studies, the authors demonstrated that Wzx from ECA can complement the Oag *wzx* deficiency in *E. coli* O16, as long as the gene cluster encoding the other genes associated with Oag processing is deleted (Marolda *et al.*, 2006). Also, reconstituting the expression of ECA Wzy or ECA Wzz significantly reduced the aforementioned complementation by Wzx ECA (Marolda *et al.*, 2006). The study also showed that Oag-associated Wzx from *E. coli* O16 and Wzx from *E. coli* O7 can cross complement deficiencies in the O16 and O7 clusters, as long as their corresponding Wzz and Wzy proteins are not co-expressed, providing the first genetic evidence that proteins involved in processing Oag may function as a complex (Marolda *et al.*, 2006).

1.5 PROPOSED MODELS OF WZZ MECHANISM

Little is known of the mode of action Wzz exerts upon Oag to maintain the wild-type modal chain length. Previous studies by Bastin *et al.* (1993) have lead to a model in which Wzz interacts with Wzy, acting as a molecular timer, thus allowing polymerization by Wzy to continue for a set amount of time, resulting in the consistent addition of repeat units during polymerisation. An alternative model proposed by Morona *et al.* (1995) suggested that Wzz

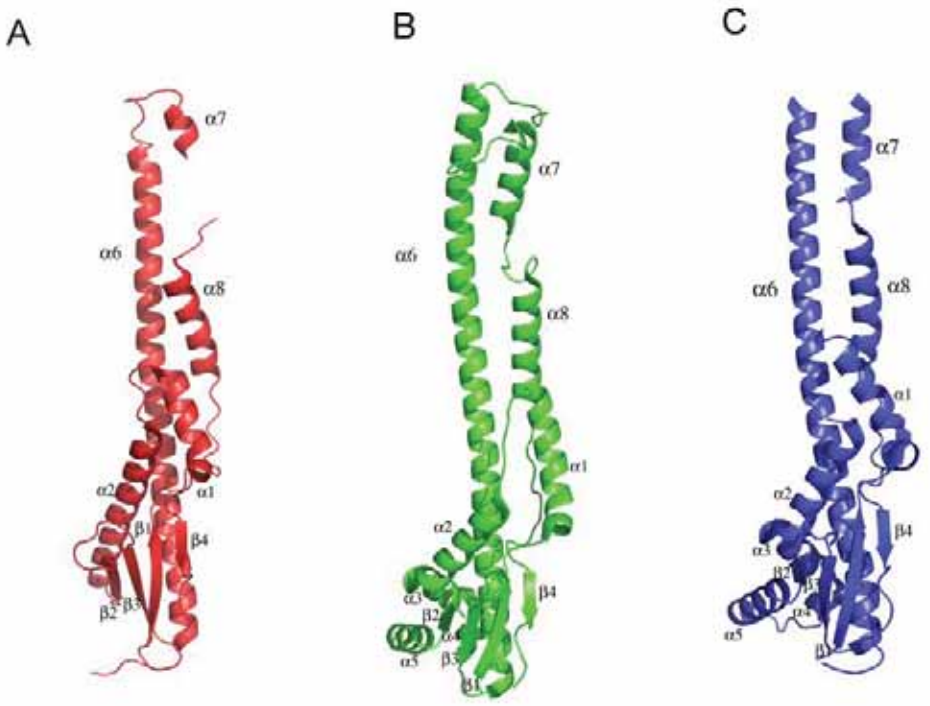
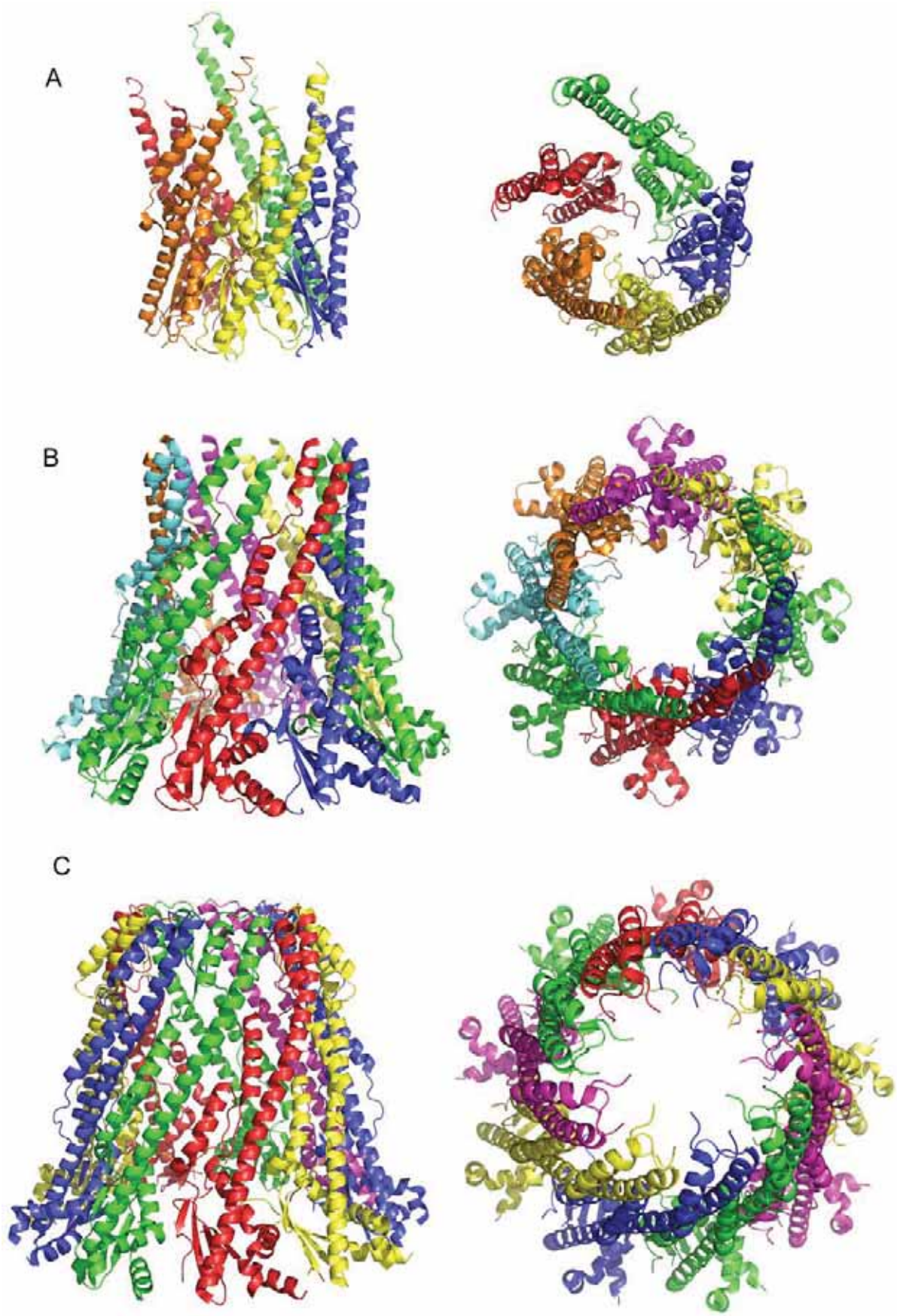


Figure 1.8: 3D structure of PCP monomers

The 3D structures for PCP monomers in ribbon representation, showing *S. typhimurium* $WzzB_{ST}$ (PDB 3b8p) in red, *E. coli* FepE (PDB 3b8n) in green and *E. coli* WzzE (PDB 3b8n) in blue. The α helices and β sheets are also indicated.

Figure 1.9: 3D structures for PCP oligomers

The 3D structural images for PCP proteins shown in ribbon representation with each monomer shown in a different colour. A) *S. typhimurium* $WzzB_{ST}$ (PDB 3b8p), B) *E. coli* WzzE (PDB 3b8n) and C) *E. coli* FepE (PDB 3b8n). LHS, side view; RHS, view looking down top and towards cytoplasmic membrane.


A
B
C

acts as a molecular chaperone, assisting the interaction between WaaL and Wzy, resulting in a modal length as a consequence of the ratio of Wzy and WaaL. Other data published have provided some support in favour of this alternative hypothesis, indicating a crucial role of WaaL in the process of Oag chain length regulation (Amor and Whitfield, 1997). Published data from Daniels (1998) indicate that the ratio of Wzy and Wzz was important in determining Oag modal chain length, which is supportive of the latter model (Daniels *et al.*, 1998). Previous studies have shown that in a strain deficient in a transcription factor (RfaH) which reduced *wzy* expression, complementation of *wzy* restored modal chain length, thus providing further evidence to suggest that Wzy levels in the cell may affect Oag modal chain length (Carter *et al.*, 2007). With recent developments in solving the PCP 3D structure and oligomeric arrangement, a new model has been proposed by Tocilj *et al.* (2008) in which the Wzz oligomers act as molecular scaffolds for multiple Wzy polymerase molecules, and the growing Oag chain is transferred from one Wzy to another Wzy molecule (Morona *et al.*, 2009; Tocilj *et al.*, 2008).

Previous targeted mutagenesis analysis has been conducted on Wzz_{SF} and although mutations targeting the TM regions exhibited dramatic changes in the resulting LPS Oag chain length, mutations targeting the periplasmic domain generally did not yield major changes in the resulting Oag modal chain length. Hence, an alternative approach to Wzz_{SF} mutagenesis was considered appropriate to investigate the relationship between Wzz structure and function by increasing the likelihood of acquiring Wzz mutants displaying phenotypic changes.

This thesis aims to investigate the structure and function of Wzz_{SF} by constructing and characterising a library of in-frame linker mutants, and to explore the involvement of TM2 regions in Wzz-Wzz protein interactions. In addition, the cellular location of Wzz has not been defined. This study therefore also aims to identify the cellular localisation of Wzz by the construction and visualisation of fluorescently tagged Wzz fusion proteins. Lastly, as the

interacting protein partners of Wzz have yet to be elucidated, the final aim of this project is to assess if interactions occur between Wzz and Wzy.

1.6 STUDY RATIONALE

Aims:

1) Investigate the structure-function relationship of Wzz_{SF} by random mutagenesis
2) Explore involvement of TM2 region in Wzz:Wzz interactions
3) Investigate the cellular localisation of Wzz_{SF}
4) Investigate protein interactions between Wzy and Wzz

CHAPTER TWO

MATERIALS AND METHODS

2.1 BACTERIAL STRAINS USED IN THIS STUDY

Bacterial strains and plasmids used and/or created in this study are summarised in Table 2.1 and Table 2.2, respectively.

2.2 BACTERIAL GROWTH CONDITIONS

2.2.1 LIQUID MEDIA GROWTH CONDITIONS

Bacteria were cultured at 37°C for 16-18 hours (h) in either Luria Bertani (LB, 0.5% (w/v) NaCl, 0.5% (w/v) Yeast Extract (Difco), 1% (w/v) Bacto Tryptone Peptone (BD)) or Mueller-Hinton broth (beef extract, 2g/L, casein, 17.5 g/L and starch, 1.5g/L, Difco) for strains carrying pSCRhaB2-based plasmids. Where appropriate, antibiotics were included at the following concentrations: ampicillin (Amp), 100 μg/ml; chloramphenicol (Cml), 25 μg/ml; kanamycin (Kan), 50 μg/ml; rifampicin (Rif), 8 μg/ml; trimethoprim (Tp), 50 μg/ml, streptomycin (Sm), 100 μg/mL and tetracycline (Tet), 10 μg/ml.

2.2.2 SOLID MEDIA GROWTH CONDITIONS

Bacteria were stored at −80°C in a suspension of glycerol (30% (w/v), Invitrogen) and peptone (1% (w/v), Difco) in Wheaton vials. Fresh cultures from these glycerols were prepared by streaking a loopful of the suspension onto LB agar (LB broth, 1.5% (w/v) Bacto agar (BD)) or MH agar (Mueller-Hinton broth, 1.5% (w/v) Bacto agar (BD)) with antibiotics as appropriate, followed by incubation at 37°C for 16-18 h to achieve adequate growth.

Table 2.1 Bacterial strains used in this study

Strain	Description	Characteristics	Source/reference
E1315	*endA hsdR supE44 thi-1 recA1 gyrA relAΔ (lacZYA-argF)* U169 (φ80d*lacΔ*(*lacZ*)M15)	Cloning strain	Gifco
Top10F'	F'*lac*I^{q} Tet^{R} *mcrA* Δ (*mrr-hsd* RMS-*mcrBC*) ϕ80 *lacZ* ΔM15 Δ*lacX*74 *deoR recA1 araD*139	Cloning strain, Tet^{R}	Stratagene
RMA2162	*S. flexneri* Y serotype, virulence plasmid negative	Parent strain to RMA2163	Lab collection
RMA2163	RMA2162 *wzz*::kan^{R}	*wzz* deficiency, Kan^{R}	Lab collection
RMA2741	RMA2163 conjugated with F'*lac*I^{q}Tet^{R}	Kan^{R}, Tet^{R}	Lab collection
PE638	*S. flexneri* Y *rpoB* (Rif^{R})	Parent strain to RMM109, Rif^{R}	Morona *et al*., 1994
RMM109	PE638Δ*wzy*	*wzy* deficient, Rif^{R}	Morona *et al*., 1994
Chapter 3 strains			
RMA3048	Top10F' (pRMCD30)	Tet^{R}, Amp^{R}	Daniels, 1999
MPRMA1	RMA2162 (pRMCD104)	Amp^{R}	This study
MPRMA2	RMA2162 (pRMCD106)	Amp^{R}	This study
MPRMA3	RMA2162 (pRMCD108)	Amp^{R}	This study
MPRMA4	RMA2162 (pRMCD113)	Amp^{R}	This study
MPRMA5	RMA2163 (pRMCD104)	Kan^{R}, Amp^{R}	This study
MPRMA6	RMA2163 (pRMCD106)	Kan^{R}, Amp^{R}	This study
MPRMA7	RMA2163 (pRMCD108)	Kan^{R}, Amp^{R}	This study
MPRMA8	RMA2163 (pRMCD113)	Kan^{R}, Amp^{R}	This study
MPRMA9	RMA2162 (pRMCD77)	Amp^{R}	This study
MPRMA10	RMA2162 (pRMCD78)	Amp^{R}	This study
MPRMA11	RMA2162 (pRMCD80)	Amp^{R}	This study
MPRMA12	RMA2162 (pRMCD122)	Amp^{R}	This study
MPRMA13	RMA2163 (pRMCD77)	Kan^{R}, Amp^{R}	This study
MPRMA14	RMA2163 (pRMCD78)	Kan^{R}, Amp^{R}	This study
MPRMA15	RMA2163 (pRMCD80)	Kan^{R}, Amp^{R}	This study
MPRMA16	RMA2163 (pRMCD122)	Kan^{R}, Amp^{R}	This study
MPRMA17	RMA2162 (pRMCD116)	Amp^{R}	This study
MPRMA18	RMA2162 (pRM117)	Amp^{R}	This study
MPRMA19	RMA2162 (pRMCD119)	Amp^{R}	This study
MPRMA21	RMA2163 (pRMCD116)	Kan^{R}, Amp^{R}	This study
MPRMA22	RMA2163 (pRMCD117)	Kan^{R}, Amp^{R}	This study
MPRMA23	RMA2163 (pRMCD119)	Kan^{R}, Amp^{R}	This study
MPRMA38	Top10F' (Wzz$_{i92}$)	Tet^{R}, Amp^{R}	This study
MPRMA39	Top10F' (Wzz$_{i128}$)	Tet^{R}, Amp^{R}	This study
MPRMA40	Top10F' (Wzz$_{i163}$)	Tet^{R}, Amp^{R}	This study
MPRMA41	Top10F' (Wzz$_{i231}$)	Tet^{R}, Amp^{R}	This study
MPRMA42	Top10F' (Wzz$_{i290}$)	Tet^{R}, Amp^{R}	This study
MPRMA43	Top10F' (Wzz$_{i191}$)	Tet^{R}, Amp^{R}	This study
MPRMA44	Top10F' (Wzz$_{i161}$)	Tet^{R}, Amp^{R}	This study
MPRMA45	Top10F' (Wzz$_{i32}$)	Tet^{R}, Amp^{R}	This study

MPRMA46	Top10F' (Wzz_{i279})	Tet^R, Amp^R	This study
MPRMA47	Top10F' (Wzz_{i255})	Tet^R, Amp^R	This study
MPRMA48	Top10F' (Wzz_{i131})	Tet^R, Amp^R	This study
MPRMA49	Top10F' (Wzz_{i81})	Tet^R, Amp^R	This study
MPRMA50	Top10F' (Wzz_{i219})	Tet^R, Amp^R	This study
MPRMA51	Top10F' (Wzz_{i247})	Tet^R, Amp^R	This study
MPRMA52	Top10F' (Wzz_{i138})	Tet^R, Amp^R	This study
MPRMA53	Top10F' (Wzz_{i80})	Tet^R, Amp^R	This study
MPRMA55	Top10F' (Wzz_{i199})	Tet^R, Amp^R	This study
MPRMA56	Top10F' (Wzz_{i66})	Tet^R, Amp^R	This study
MPRMA57	RMA2741 (pQE-30)	Tet^R, Kan^R, Amp^R	This study
MPRMA58	RMA2741 (pRMCD30)	Tet^R, Kan^R, Amp^R	This study
MPRMA59	RMA2741 (Wzz_{i92})	Tet^R, Kan^R, Amp^R	This study
MPRMA60	RMA2741 (Wzz_{i128})	Tet^R, Kan^R, Amp^R	This study
MPRMA61	RMA2741 (Wzz_{i163})	Tet^R, Kan^R, Amp^R	This study
MPRMA62	RMA2741 (Wzz_{i231})	Tet^R, Kan^R, Amp^R	This study
MPRMA63	RMA2741 (Wzz_{i290})	Tet^R, Kan^R, Amp^R	This study
MPRMA67	RMA2741 (Wzz_{i191})	Tet^R, Kan^R, Amp^R	This study
MPRMA68	RMA2741 (Wzz_{i161})	Tet^R, Kan^R, Amp^R	This study
MPRMA69	RMA2741 (Wzz_{i32})	Tet^R, Kan^R, Amp^R	This study
MPRMA70	RMA2741 (Wzz_{i279})	Tet^R, Kan^R, Amp^R	This study
MPRMA71	RMA2741 (Wzz_{i255})	Tet^R, Kan^R, Amp^R	This study
MPRMA72	RMA2741 (Wzz_{i131})	Tet^R, Kan^R, Amp^R	This study
MPRMA73	RMA2741 (Wzz_{i81})	Tet^R, Kan^R, Amp^R	This study
MPRMA74	RMA2741 (Wzz_{i219})	Tet^R, Kan^R, Amp^R	This study
MPRMA75	RMA2741 (Wzz_{i247})	Tet^R, Kan^R, Amp^R	This study
MPRMA76	RMA2741 (Wzz_{i138})	Tet^R, Kan^R, Amp^R	This study
MPRMA77	RMA2741 (Wzz_{i80})	Tet^R, Kan^R, Amp^R	This study
MPRMA78	RMA2741 (Wzz_{i199})	Tet^R, Kan^R, Amp^R	This study
MPRMA79	RMA2741 (Wzz_{i66})	Tet^R, Kan^R, Amp^R	This study

Chapter 4 strains			
MPRMA108	E1315 (pGEMT-Easy::*FLAG-wzz$_{SF}$*)	Amp^R	This study
MPRMA109	E1315 (pBAD-Wzz$_{SF}$)	Cml^R	This study
MPRMA112	RMA2163 (pBAD33)	Kan^R, Cml^R	This study
MPRMA113	RMA2163 (pBAD-Wzz$_{SF}$)	Kan^R, Cml^R	This study
MPRMA115	RMA2741 (pBAD33)	Tet^R, Cml^R	This study
MPRMA118	RMA2741 (pBAD-Wzz$_{SF}$)	Kan^R, Tet^R, Cml^R	This study
MPRMA119	RMA2741 (pBAD-Wzz$_{SF}$) (pRMCD30)	Kan^R, Tet^R, Cml^R, Amp^R	This study
MPRMA122	RMA2741 (pRMCD30) (pBAD33)	Kan^R, Tet^R, Cml^R, Amp^R	This study
MPRMA123	E1315 (pGEM-T Easy::*FLAG-Wzz$_{ST}$*)	Amp^R	This study
MPRMA124	E1315 (pGEM-T Easy::FLAG-Wzz$_{G305A/G311A}$)	Amp^R	This study
MPRMA125	E1315 (pBAD-Wzz$_{G305A/G311A}$)	Cml^R	This study
MPRMA126	E1315 (pBAD-Wzz$_{ST}$)	Cml^R	This study
MPRMA127	RMA2741 (pRMCD30) (pBAD-Wzz$_{G305A/G311A}$)	Kan^R, Tet^R, Cml^R, Amp^R	This study
MPRMA128	RMA2741 (pRMCD30) (pBAD-Wzz$_{ST}$)	Kan^R, Tet^R, Cml^R, Amp^R	This study
MPRMA129	RMA2741 (pQE-30) (pBAD-Wzz$_{SF}$)	Kan^R, Tet^R, Cml^R, Amp^R	This study
MPRMA131	RMA2741 (pBAD-Wzz$_{ST}$)	Kan^R, Tet^R, Cml^R	This study
MPRMA132	RMA2741 (pBAD-Wzz$_{G305A/G311A}$)	Kan^R, Tet^R, Cml^R	This study
MPRMA133	RMA2162 (pBAD-Wzz$_{G305AG311A}$)	Cml^R	This study
MPRMA134	RMA2162 (pBAD-Wzz$_{ST}$)	Cml^R	This study
MPRMA142	RMA2162 (pCDFDuet-1)	Sm^R	This study
MPRMA143	RMA2163 (pCDFDuet-1)	Kan^R, Sm^R	This study
MPRMA144	RMA2162 (pCDFDuet-1) (Wzz$_{i191}$)	Sm^R, Amp^R	This study
MPRMA146	RMA2162 (pCDFDuet-1) (Wzz$_{i255}$)	Sm^R, Amp^R	This study
MPRMA147	RMA2162 (pCDFDuet-1) (Wzz$_{i219}$)	Sm^R, Amp^R	This study
MPRMA148	RMA2162 (pCDFDuet-1) (Wzz$_{i247}$)	Sm^R, Amp^R	This study
MPRMA149	RMA2162 (pCDFDuet-1) (Wzz$_{i231}$)	Sm^R, Amp^R	This study
MPRMA150	RMA2163 (pCDFDuet-1) (Wzz$_{i191}$)	Kan^R, Sm^R, Amp^R	This study
MPRMA152	RMA2163 (pCDFDuet-1) (Wzz$_{i255}$)	Kan^R, Sm^R, Amp^R	This study
MPRMA153	RMA2163 (pCDFDuet-1)	Kan^R, Sm^R,	This study

	(Wzz_{i219})	Amp^R	
MPRMA154	RMA2163 (pCDFDuet-1) (Wzz_{i247})	Kan^R, Sm^R, Amp^R	This study
MPRMA155	RMA2163 (pCDFDuet-1) (Wzz_{i231})	Kan^R, Sm^R, Amp^R	This study
MPRMA156	RMA2162 (pCDFDuet-1) (pQE-30)	Sm^R, Amp^R	This study
MPRMA157	RMA2162 (pCDFDuet-1) (pRMCD30)	Sm^R, Amp^R	This study
MPRMA158	RMA2163 (pCDFDuet-1) (pQE-30)	Kan^R, Sm^R, Amp^R	This study
MPRMA159	RMA2163 (pCDFDuet-1) (pRMCD30)	Kan^R, Sm^R, Amp^R	This study
MPRMA196	RMA2162 (pCDFDuet-1) (Wzz_{i128})	Sm^R, Amp^R	This study
MPRMA197	RMA2163 (pCDFDuet-1) (Wzz_{i128})	Kan^R, Sm^R, Amp^R	This study
MPRMA198	RMA2162 (pCDFDuet-1) (Wzz_{i131})	Sm^R, Amp^R	This study
MPRMA199	RMA2163 (pCDFDuet-1) (Wzz_{i131})	Kan^R, Sm^R, Amp^R	This study
Chapter 5 strains			
MPRMA174	E1315 (pGEM-T Easy::*Sph*I*mcherry*)	Amp^R	This study
MPRMA190	E1315 (pQMCherry)	Amp^R	This study
MPRMA207	Top10F'(pQMCherry-Wzz_{SF})	Amp^R	This study
MPRMA217	RMA2741 (pQMCherry)	Amp^R	This study
MPRMA218	RMA2741 (pQMCherry-Wzz_{SF})	Amp^R	This study
Chapter 6 strains			
MPRMA171	E1315 (pGEM-T Easy::*strepII-wzy*)	Amp^R	This study
MPRMA183	E1315 (pStrepII-Wzy)	Tp^R	This study
MPRMA187	RMM109 (pSCRhaB2)	Rif^R, Tp^R	This study
MRRMA188	RMM109 (pStrepII-Wzy)	Rif^R, Tp^R	This study
MPRMA194	PE638 (pSCRhaB2)	Rif^R, Tp^R	This study
MPRMA221	Top10F' (pGEM-T Easy::*speIgfp*$^+$)	Tet^R, Amp^R	This study
MPRMA223	Top10F' (pGEM-T Easy::*spegfp*$^+$*sfo*)	Tet^R, Amp^R	This study
MPRMA225	Top10F' (pGEM-T Easy::*gfp*$^+$*strepII-wzy*)	Tet^R, Amp^R	This study
MPRMA227	Top10F' (pGEM-T Easy::*gfp*$^+$ *wzy*)	Tet^R, Amp^R	This study
MPRMA229	E1315 (*Kpn*I-*gfp*$^+$*strepIIwzy*-*Xba*I cloned into pGEM-T Easy)	Amp^R	This study
MPRMA231	E1315 (*Kpn*I-*gfp*$^+$*wzy*-*Xba*I cloned into pGEM-T Easy)	Amp^R	This study

MPRMA233	E1315 (pGFP^{+}StrepII-Wzy)	Tp^{R}	This study
MPRMA234	E1315 (pGFP^{+}Wzy)	Tp^{R}	This study
MPRMA239	RMM109 (pGFP^{+}StrepIIWzy)	Rif^{R}, Tp^{R}	This study
MPRMA240	RMM109 (pGFP^{+}Wzy)	Rif^{R}, Tp^{R}	This study

Table 2.2 Plasmids used in this study

Plasmid	Description	Resistance	Source/reference
Chapter 3 plasmids			
pGEM-T Easy	Cloning vector	Amp^R	Promega
pQE-30	Expression vector with N-terminal His_6 tag	Amp^R	Qiagen
pRMCD30	pQE-30 with wzz_{SF}	Amp^R	Daniels and Morona, 1999
$pWzz_{i32}$	pQE-30 with wzz_{SF} and an in-frame linker insertion at aa position 33	Amp^R	This study
$pWzz_{i66}$	pQE-30 with wzz_{SF} and an in-frame linker insertion at aa position 67	Amp^R	This study
$pWzz_{i80}$	pQE-30 with wzz_{SF} and an in-frame linker insertion at aa position 81	Amp^R	This study
$pWzz_{i81}$	pQE-30 with wzz_{SF} and an in-frame linker insertion at aa position 82	Amp^R	This study
$pWzz_{i92}$	pQE-30 with wzz_{SF} and an in-frame linker insertion at aa position 93	Amp^R	This study
$pWzz_{i128}$	pQE-30 with wzz_{SF} and an in-frame linker insertion at aa position 129	Amp^R	This study
$pWzz_{i131}$	pQE-30 with wzz_{SF} and an in-frame linker insertion at aa position 132	Amp^R	This study
$pWzz_{i138}$	pQE-30 with wzz_{SF} and an in-frame linker insertion at aa position 139	Amp^R	This study
$pWzz_{ii161}$	pQE-30 with wzz_{SF} and an in-frame linker insertion at aa position 162	Amp^R	This study
$pWzz_{i163}$	pQE-30 with wzz_{SF} and an in-frame linker insertion at aa position 164	Amp^R	This study
$pWzz_{i191}$	pQE-30 with wzz_{SF} and an in-frame linker insertion at aa position 192	Amp^R	This study
$pWzz_{i199}$	pQE-30 with wzz_{SF} and an in-frame linker insertion at aa position 200	Amp^R	This study
$pWzz_{i219}$	pQE-30 with wzz_{SF} and an in-frame linker insertion at aa position 220	Amp^R	This study
$pWzz_{i231}$	pQE-30 with wzz_{SF} and an in-frame linker insertion at aa position 232	Amp^R	This study

pWzz$_{i247}$	pQE-30 with *wzz$_{SF}$* and an in-frame linker insertion at aa position 248	Amp^R	This study
pWzz$_{i255}$	pQE-30 with *wzz$_{SF}$* and an in-frame linker insertion at aa position 256	Amp^R	This study
pWzz$_{i279}$	pQE-30 with *wzz$_{SF}$* and an in-frame linker insertion at aa position 280	Amp^R	This study
pWzz$_{i290}$	pQE-30 with *wzz$_{SF}$* and an in-frame linker insertion at aa position 291	Amp^R	This study
Chapter 4 plasmids			
pRMCD78	*wzz$_{SF}$* in pET-17b	Amp^R	Daniels and Morona, 1999
pRMCD80	*wzz$_{ST}$* in pET-17b	Amp^R	Daniels and Morona, 1999
pRMCD104	Chimeric *wzz$_{ST/SF}$* in pET-17b	Amp^R	Daniels and Morona, 1999
pRMCD106	Chimeric *wzz$_{SF/ST}$* in pET-17b	Amp^R	Daniels and Morona, 1999
pRMCD108	*wzz$_{K267N}$* in pET-17b	Amp^R	Daniels and Morona, 1999
pRMCD113	*wzz$_{G305A/G311A}$* in pET-17b	Amp^R	Daniels and Morona, 1999
pRMCD116	*wzz$_{P292A}$* in pET-17b	Amp^R	Daniels and Morona, 1999
pRMCD117	*wzz$_{P286A}$* in pET-17b	Amp^R	Daniels and Morona, 1999
pRMCD119	*wzz$_{K31A}$* in pET-17b	Amp^R	Daniels and Morona, 1999
pRMCD122	*wzz$_{M32T/I35C}$* in pET-17b	Amp^R	Daniels and Morona, 1999
pBAD33	Arabinose-inducible pBAD promoter vector	Cml^R	Guzman *et al*., 1995
pGEM-T Easy::*FLAGwzz$_{SF}$*	PCR amplified *Sac*I-*FLAG-wzz$_{SF}$*-*Sma*I fragment with primers SacFLAGWzzF and SmaWzzR cloned in pGEM-T Easy	Amp^R	This study
pBAD-Wzz$_{SF}$	*FLAG-wzz$_{SF}$* cloned in pBAD33 via *Sac*I and *Sma*I sites	Cml^R	This study
pGEM-T Easy::*FLAG-wzz$_{ST}$*	PCR amplified *Sac*I-*FLAG-wzz$_{ST}$*-*Sma*I fragment cloned in pGEM-T Easy	Amp^R	This study
pBAD-Wzz$_{ST}$	*FLAG-wzz$_{ST}$* cloned in pBAD33 via *Sac*I and *Sma*I sites	Cml^R	This study

pGEM-T Easy::FLAG-$wzz_{G305A/G311A}$	PCR amplified *Sac*I-*FLAG-wzz*$_{G305A/G311A}$-*Sma*I fragment with primers SacFLAGWzzF and SmaWzzR cloned in pGEM-T Easy	Amp^R	This study
pBAD-$Wzz_{G305A/G311A}$	*FLAG-wzz*$_{G305A/G311A}$ cloned in pBAD33 via *Sac*I and *Sma*I sites	Cml^R	This study
pCDFDuet-1	Vector with T7 promoter, CDF ori	Sm^R	Novagen
pBAD-NTF	Construct with FLAG tag	Amp^R	Marolda *et al*., 2004
Chapter 5 plasmids			
pRSET-B	Plasmid with *mCherry*		Shaner *et al*., 2004
pGEM-T Easy::*sph*I*mcherry*	pGEMT-Easy with PCR amplified m*Cherry*, amplified with primers mCherryF and mCherryR	Amp^R	This study
pQMCherry	m*Cherry* cloned in pQE-30 via *Bam*HI and *Sph*I	Amp^R	This study
pQMCherry-Wzz_{SF}	*wzz*$_{SF}$ cloned into pQMCherry via *Sac*I and *Sma*I sites	Amp^R	This study
Chapter 6 plasmids			
pSCRhaB2	P_{rhaB}, pBBR1 ori	Tp^R	Cardona and Valvano, 2005
pGEM-T Easy::*strepII-wzy*	PCR amplified *Nde*I-*strepII-wzy*-*Hind*III fragment with primers NdeStrepWzyF and HindWzyR cloned in pGEM-T Easy	Amp^R	This study
pStrepII-Wzy	pSCRhaB2 with *strepII-wzy* cloned in via *Nde*I and *Bam*HI	Tp^R	This study
pWH1012gfp+13	Plasmid containing *gfp*$^+$	Amp^R, Kan^R	Scholz *et al*., 2000
pGEM-T Easy::*speIgfp*$^+$	PCR amplified *gfp*$^+$ from pWH1012gfp+13 with primers SpeIgfpF and SpeIgfpR fragment cloned into pGEM-T Easy	Amp^R	This study
pGEM-T Easy::*speIgfp*$^+$*sfo*	PCR amplified *gfp*$^+$ from pWH1012gfp+13 with primers SpeIgfpF and SfoIgfpR fragment cloned into pGEM-T Easy	Amp^R	This study
pGEM-T Easy::*gfp*$^+$*strepII-wzy*	*Spe*I*gfp*$^+$*Spe*I fragment cloned into pGEM-T Easy::*strepII-wzy*	Amp^R	This study
pGEM-T Easy::*gfp*$^+$ *wzy*	*Spe*I*gfp*$^+$*Sfo*I fragment cloned into pGEM-T Easy::*strepII-wzy*	Amp^R	This study
pGEM-T Easy::*Kpn*I*gfp*$^+$*strep II-wzy-Xba*I	PCR amplified *Kpn*I-*gfp*$^+$*strepIIwzy-Xba*I fragment with primers KpngfpF and XbagfpR, cloned into pGEM-T Easy	Amp^R	This study

pGEM-T Easy::*Kpn*I*gfp*$^{+}$*wzy*-*Xba*I	PCR amplified *Kpn*I-*gfp*$^{+}$*wzy*-*Xba*I fragment with primers KpngfpF and XbagfpR, cloned into pGEM-T Easy	Amp^{R}	This study
pGFP^{+}StrepII-Wzy	*gfp*$^{+}$*strepIIwzy* fragment cloned into pSCRhaB2 via *Kpn*I and *Xba*I	Tp^{R}	This study
pGFP^{+}Wzy	*gfp*+*wzy* fragment cloned into pSCRhaB2 via *Kpn*I and *Xba*I	Tp^{R}	This study
pCB6-GFP-WH1	N-WASP domain fused to GFP	Amp^{R}	M. Way, EMBL, Heidelberg, Germany

2.3 CHEMICALS AND REAGENTS

Chemicals and reagents were obtained from the following suppliers: Sigma-Aldrich, Roche, Difco, BD, Ajax, Promega, Qiagen, Invitrogen, New England Biolabs. Key chemicals used in this study were L-arabinose (Sigma, cat. number A3256), L-Rhamnose monohydrate (Sigma, cat. number 83650), isopropyl-β-D-thiogalactopyranoside (IPTG, Biovectra, cat. number 1882), n-dodecyl B-D maltoside (DDM, Sigma cat. number D4641-1G), bis-acrylamide solution (BioRad, cat. number 161-0142), ammonium persulphate (BioRad, cat. number 161-0700), N,N,N',N'-Tetramethyl-ethylenediamine (TEMED, Sigma cat. number T22500), formaldehyde (Sigma, cat. number F8775), proteinase K (Invitrogen, cat. number 25530015).

2.4 DNA PREPARATION AND MANIPULATION

2.4.1 RESTRICTION ENZYME DIGESTS AND LIGATION REACTIONS

Restriction digests of plasmid DNA and ligation reactions were performed as specified by the manufacturer (NEB). Prior to ligation reactions, shrimp alkaline phosphatase (SAP) was used to prevent the re-ligation of digested plasmids as specified by the manufacturer (Roche). Briefly, 5 µL of digested plasmid was mixed with 1 µL of SAP and 1 µL of SAP buffer, with 3 µL of MQ water and incubated at 37°C for 30 minutes. The enzyme was deactivated by incubation at 65°C for 15 minutes, and 1 µL of this reaction was used in ligation reactions. Ligation using the pGEMT-Easy vector kit was also performed as specified by the manufacturer (Promega). Blue-white colour selection of transformants was conducted with X-gal in the LB agar, at a concentration of 80 µg/ml in N,N-dimethylformamide (DMF).

2.4.2 AGAROSE GEL ELECTROPHORESIS

Samples for electrophoresis were mixed with loading dye (Table 2.3) and separated on a 1% (w/v) agarose gel using TBE buffer at 110 V for 70 minutes. DNA sizes were determined by comparison of simultaneously electrophoresed *Eco*RI-restricted bacteriophage SPP1 known standards. The SPP1 sizes (in Kb) are; 8.51, 7.35, 6.11, 4.84, 3.59, 2.81, 1.95, 1.86, 1.51, 1.39, 1.16, 0.98, 0.72, 0.48, 0.36 and 0.09. SPP1 molecular weight standards were prepared in house. Gels were post stained with ethidium bromide solution, which later on in the project was replaced by Gel Red (Biotium). Following a 15-30 minute stain, the gels were rinsed in RO water and visualised under UV light with a 312 nm UV transilluminator (Spectroline) and photos were taken using a Tracktel GDS-2 gel documentation system. Towards the latter segment of the project, a BioRad Molecular Imager Gel Doc XR+ System was used to analyse and photograph agarose gels.

2.4.3 DNA PREPARATION USING A KIT

Plasmid DNA was extracted using the QIAprep Spin Miniprep kit (Qiagen). PCR amplified DNA and various restriction enzyme digests were purified using QIAquick PCR purification kit (Qiagen). Gel extracted DNA was purified using the PureLink Gel Extraction kit (Invitrogen).

2.4.4 PREPARATION OF BOILED CELL LYSATES FOR PCR

Bacterial lysates for PCR amplification were prepared by resuspending single colonies of bacteria in 100 µL of sterile MQ water using a sterile plastic pipette tip. Samples were incubated at 100°C for 5 minutes, then centrifuged at 13,000 rpm for 2 minutes (Heraeus Sepatech Biofuge 15). Two microlitres of the supernatant was used as a template for PCR.

Table 2.3 Buffers

Procedure	Reagent	Composition
Bacterial culture	LB broth	0.5% (w/v) NaCl, 0.5% (w/v) Yeast Extract (Difco), 1% (w/v), Bacto Tryptonepeptone (BD)
	LB agar	LB broth, 1.5% (w/v) Bacto agar (BD)
	Mueller-hinton (MH) broth	Beef extract 2g/L, casein 17.5 g/L, and starch 1.5 g/L (Difco)
	MH agar	MH broth, 1.5% (w/v) Bacto agar (BD)
	SOC media	2% (w/v) Tryptone (Oxoid), 0.5% (w/v) Yeast Extract (Oxoid), 0.04M NaCl, 0.01M KCl, 0.01M $MgCl_2$, 0.01M $MgSO_4$)
	-70ºC glycerol media	30% (w/v) glycerol (Invitrogen), 1% (w/v) Bacto peptone (Difco)
Preparing competent cells	Solution α	30 mM $K(CH_3COO)$, 100 mM KCl, 10 mM $CaCl_2$, 50 mM $MnCl_2$, 15% glycerol (v/v)
	Solution β	10 mM 3-N-morpholino propanesulfonic acid (MOPS, Sigma), 75 mM $CaCl_2$, 10 mM KCl, 15% (v/v) glycerol
Agarose gel electrophoresis	Loading buffer	0.1% (w/v) bromophenol blue (Sigma), 20% (v/v) glycerol, 0.1 mg/mL RNase (Qiagen)
	10x TBE	0.5 M Tris-HCl, 5 M boric acid, 0.001 M EDTA
	Agarose gel	1% (w/v) DNA grade agarose (Quantum Scientific) in 1x TBE buffer
SDS-PAGE	2 x sample buffer	4% (w/v) SDS, 20% (v/v) glycerol, 10% (v/v) β-mercaptoethanol, 0.04% (w/v) bromophenol blue, 0.125 M Tris-HCl, pH 6.8

	5 x Running buffer	0.5% (w/v) SDS, 1 M glycine, 0.125 M Tris-HCl
Western immunoblotting	1x TTBS buffer	0.016 M Tris, 0.05% (v/v) Tween 20 (Sigma), 0.12 M NaCl
	1x TBS buffer	0.016 M Tris, 0.12 M NaCl
	Transfer buffer	5 % (w/v) methanol, 0.025 M Tris, 0.2 M glycine
	Ponceau S stain	0.1% (w/v) Ponceau S (Sigma), 5% (v/v) glacial acetic acid
LPS silver staining	Lysing buffer	2% (w/v) SDS, 4% (v/v) β-mercaptoethanol (Sigma), 10% (v/v) glycerol, 0.1% (w/v) bromophenol blue (Sigma), 0.66 M Tris-HCl, pH 7.6
	Proteinase K solution	2.5 mg/ml in lysing buffer
	Fixing solution	5% (v/v) glacial acetic acid, 40% (v/v) ethanol
	Oxidising solution	5 % (v/v) glacial acetic acid, 40% (v/v) ethanol, 0.7% (w/v) periodic acid
	Staining solution	2 ml NH_3OH, 0.12 g sodium hydroxide, 1 g $AgNO_3$, made up to 150 ml with MQ or 28 ml sodium hydroxide (0.1 M), 2 ml ammonium hydroxide 30% (w/v) and 5 ml of $AgNO_3$ (20% (w/v) in MQ
	Developing solution	50 mg citric acid, 1 L MQ warmed to 42ºC, 500 µL formaldehyde
	Stopping solution	4% (v/v) acetic acid
Solubilisation buffers		2% (w/v) Triton X-100, 50 mM Tris pH 7.5 2% (w/v) Triton X-100/1 mM $MgCl_2$, 50 mM Tris pH 7.5 1.5% (w/v) Nonidet P40, 50 mM Tris pH 7.5 1% (w/v) perfluoro-octanoic acid (PFO), 50 mM Tris pH 7.5 1% (w/v) Zwittergent, 50 mM Tris pH 7.5

His_6-tagged protein purification	Solubilisation buffer	1% D-dodecyl β-D maltoside (DDM), 50 mM Tris, pH 7.5
	NaH_2PO_4 equilibration buffer	0.05 M NaH_2PO_4, 0.3 M NaCl, 0.02% DDM, pH 8.0
	Wash buffer A	Equilibration buffer with 20 mM imidazole
	Wash buffer B	Equilibration buffer with 30 mM imidazole
	Elution buffer	Equilibration buffer with 250 mM imidazole
Formaldehyde cross-linking	Wash buffer	10 mM K_2PO_4/KH_2PO_4 pH 6.8, in MQ
Microscopy	PBS wash buffer	Saline (0.85% (w/v) NaCl)
	Formaldehyde fixing solution	1 mL 3.7% (w/v) formaldehyde (Sigma) in saline (0.85% (w/v) NaCl)
	Poly-L-lysine solution	0.01% (w/v) Poly-L-lysine (Sigma) in saline (0.85% (w/v) NaCl)
	Mowiol/PPD solution	Mowiol 4-88 (Calbiochem) supplemented with 20 μg/ml *p*–phenylenylenediamine (Sigma) (at a ratio of 1:10, PPD:Mowiol)

2.5 POLYMERASE CHAIN REACTION (PCR)

2.5.1 GENERAL PCR

Oligonucleotides used in this study are listed in Table 2.4. PCR reactions were performed in 50 μL volume using 200 μM dNTPs (Sigma), 1x buffer (NEB), 100 μM oligonucleotide primers and 0.25 U *Taq* polymerase (NEB). Reactions were performed using an Eppendorf Mastercycler Gradient PCR thermocycler. Unless stated otherwise, PCR conditions involved 25 cycles of denaturation (95°C, 30 sec), annealing (ranging from 55-70°C, 30 sec) and extension (72°C, 1 min/kb product). Phusion High-Fidelity DNA Polymerase (Finnzymes) was used to amplify PCR products for cloning, as described by the manufacturer.

2.5.2 DNA SEQUENCING

Sequencing was performed using the ABI Prism Big Dye Terminator version 3.1. Sequencing reactions were carried out using 400 ng of plasmid DNA template, 2 μL 5 x BDT buffer, 4 μL Big Dye terminator mix, 1 μl of a single primer. Sequencing reactions were performed under the following cycling conditions; 96°C for 2 minutes, and 25 cycles of 96°C for 10 seconds, 50°C for 5 seconds, 60°C for 4 minutes. Following amplification, Big Dye Terminator-labeled DNA was precipitated by the addition of 75 μL of 0.2 mM magnesium sulphate and room temperature incubation for 15 minutes. Samples were then centrifuged (13,000 rpm, 15 min, 4°C, Eppendorf centrifuge 5415-R) and then washed in 100 μL of 70% (v/v) ethanol, and centrifuged again as before. Pellets were dried in a 37°C heating block to remove the remaining alcohol, and samples were sent to be analysed by the Australian Genomic Research Facility, Level 5, Gehrmann Laboratories, Research Road, The University

Table 2.4 Oligonucleotides used in this study

Primer	Sequence 5' – 3'[b]	Application	nt position[a]
#2197	AGGGTA*GAGCTC*AGAGTAGAAAAT	Forward primer amplifying wzz_{SF}	nt16276-16299
#2198	GTTA*CCCGGG*GAGCAGGTGTGA	Reverse primer amplifying wzz_{SF}	nt17447-17426
*Not*I miniprimer	TGCGGCCGCA	Primer homologous to 15-bp insertion using MGS	N/A
SacFLAGF	AGGGTA*GAGCTC*AGGAGATATCTTATGGACT ACAAGGACGACGACGACAAGAGAGTAGAAA ATAATAATGTTTC	Forward primer for *FLAG-*wzz_{SF}	nt16288-16310
SacFLAGSTF	AGGGTA*GAGCTC*AGGAGATATCTTATGGACT ACAAGGACGACGACGACAAGACAGTGGATA GTAATACGTC	Forward primer for *FLAG-*wzz_{ST}	nt1839-1858
SmaFLAGSTR	CAAT*CCCGGG*TTACAAGGCTTTTGGCTTATA GC	Reverse primer for *FLAG-*wzz_{ST}	nt2797-2816
mCherryF	*GGATCC*GTGAGCAAGGGCGAGGAG	Forward primer amplifying *mcherry*	nt4-21
mCherryR	*GCATGC*CAGATCCCGACCCATTTGCTGTCCA CCAGTCTTGTACAGCTCGTCCATG	Reverse primer amplifying *mcherry*	nt708-690
WzystrepF	AGGGTA*CATATG*AGGAGAATATACAAAATG*G CTAGC*TGGAGCCACCCGCAGTTCGAAAAA*GG CGCC*AATAATATTAATAAAATTTTTATTACA	Forward primer amplifying *wzy* and inserting *strepII*	nt8953-8979
WzyHindR	CAAT*AAGCTT*TTATTTTGCTCCAGAAGTGAG	Reverse primer amplifying *wzy*	nt10095-10078
SpegfpF	*ACTAGT*GCTAGCAAAGGAGAAGAAC	Forward primer amplifying gfp^+	nt4-22 of *gfp+*
SpegfpR	*ACTAGT*TTTGTAGAGCTCATCCATG	Reverse primer amplifying gfp^+	n717-694 of *gfp+*
SfogfpR	*GGCGCC*TTTGTAGAGCTCATCCATG	Reverse primer amplifying gfp^+	n717-694 of *gfp+*
KpnIgfpF	*GGTACC*AGGAGAATATACAAAATGGCTAGT GC	Forward primer amplifying 5' region flanking gfp^+wzy	N/A
XbaIgfpR	*TCTAGA*TTATTTTGCTCCAGAAGTGAG	Reverse primer amplifying gfp^+wzy region	N/A
M13F	GTTTTCCCAGTCACGAC	Used for sequencing inserts in pGEM-T Easy	nt2956-2972
M13R	CAGGAAACAGCTATGAC	Used for sequencing inserts in pGEM-T Easy	nt176-192
mCherryF1	GAGATCAAGCAGAGGCTGAAG	Used for sequencing *mcherry* and flanking region	nt493-513
Wzz129R	GGCAATAGCCACAATGACG	Used for sequencing *wzz* and flanking regions	nt16413-16395

[a] wzz_{SF} and wzy_{SF} genbank reference number X71970.1, wzz_{ST} (Z17278), pGEM-T Easy sequence available at http://www.promega.com/vectors/pgemtez.txt, mcherry (AY678264.1)
[b] Restriction sites underlined and italicised

of Queensland using an AB3730xl 96-capillary sequencer. Raw sequence data was analysed using DNAMAN (Lynnon Corporation).

2.6 PREPARATION OF COMPETENT CELLS

2.6.1 CHEMICALLY COMPETENT CELLS

Chemically competent cells were prepared from 16-18 h cultures, diluted 1/20 in 10 mL LB broth, and incubated with aeration at 37°C to an OD600 of ~0.8. Cultures were centrifuged at 4500 rpm for 10 minutes (Sigma 3K15) and pellets were resuspended in 5 mL solution α (Table 2.3). Bacterial suspensions were centrifuged at 13,000 rpm for 1 minute (Eppendorf centrifuge 5415-R), and the pellets were resuspended in 1 mL of solution β and incubated on ice for 1 h. Cells were dispensed into 1.5 mL reaction tubes and placed at -80°C for storage.

2.6.2 ELECTROCOMPETENT CELLS

Electrocompetent cells were prepared from 16-18 h cultures diluted in 10 ml LB broth and incubated with aeration at 37°C to an OD600 of ~0.8. Cultures were centrifuged at 4500 rpm for 10 minutes (Sigma 3K15) and pellets were washed twice in 5 mL of cold 10% (v/v) glycerol, centrifuged, and resuspended in 1 mL of 10% (v/v) glycerol. Aliquots were dispensed in 1.5 mL reaction tubes and stored at -80°C.

2.7 BACTERIAL TRANSFORMATION

2.7.1 HEAT SHOCK TRANSFORMATION

Chemically competent cells were thawed on ice for 5 minutes and mixed with 10 μL DNA and left on ice for 30 minutes. Cells were heat shocked by placing the cells at 42°C for

2 minutes and rapidly returning to site on ice for a further 30 minutes. LB was added to the cells (~600 μL) and incubated at 37°C for 45 minutes. Cells were then plated on LB plates containing antibiotics where appropriate, and the plates were incubated at 37°C for 16-18 h.

2.7.2 ELECTROPORATION

Electrocompetent cells were thawed on ice for 5 minutes, and mixed with 2 μL DNA. The mixture was left on ice for a further 5 minutes, and transferred to a sterile electroporation cuvette (0.2 cm gap, BioRad). Cells were electroporated at 2.5 kV (BioRad Gene Pulser, 25 μF, Capacitance extender 960 μF, Pulse Controller 200 Ω). SOC medium containing 0.2% glucose (w/v) was then rapidly added to the transformation mixture, and transformants were incubated at 37°C for 45 minutes, then plated on LB containing the appropriate antibiotics and incubated at 37°C for 16-18 h. When electroporating the vector (or constructs based on) pSCRhaB2, the transformation mixture was incubated for 1 hour at 37°C, and incubated at room temperature for 16-18 h prior to being plated on MH agar, containing the appropriate antibiotics (Cardona and Valvano, 2005).

2.7.3 CONJUGATION

Conjugation was performed using 16-18 hour cultures of the donor and recipient strains. Both cultures were centrifuged at 4500 rpm for 10 minutes (Sigma 3K15) and washed twice in 10 mL LB, and resuspended in 10 mL of LB. The donor and recipient strains were mixed in a 1:10 ratio, and centrifuged as above. The pellet was resuspended in 150 μL LB and spread onto a sterile cellulose acetate membrane filter (0.45 μm, type HA, Millipore) placed on an LB plate and incubated at 37°C for 4 h. The filter was then removed and transferred to 10 mL of LB broth. The cell suspension was diluted and spread on agar with antibiotics selecting for the desired conjugated strain, and incubated at 37°C for 16-18 h.

2.8 IN-FRAME LINKER MUTAGENESIS

A library of in-frame linker mutants was created using the Mutagenesis Generation System® (MGS) as directed by the manufacturer (Finnzymes). Briefly, plasmid pRMCD30 (Table 2.2), a pQE-30 (Qiagen) based construct with the *wzz*$_{SF}$ ORF expressed as a His_6 tagged protein His_6-Wzz_{SF}, was incubated with the kanamycin resistance-conferring Mu *entranceposon* DNA sequence element (harbouring *Not*I sites very close to its ends) to allow its random formation of transposition complexes. Transformants were selected for the presence of the *entranceposon* via kanamycin resistance, and were pooled together into three separate pools by applying 1.5 mL of LB broth onto transformation plates, using a sterile spreader to resuspend the colonies, and a sterile Pasteur pipette to harvest the pools into Wheaton vials. Each pool was subjected to *Not*I restriction digest, in order to extract the *entranceposon* from the plasmid, and were re-ligated, resulting in a 15 bp insertion containing the *Not*I restriction site. These plasmids were then electroporated into Top10 F' and transformants were plated on LB + Amp. Individual colonies were isolated, assessed for sensitivity to kanamycin, and boiled cell lysates were prepared for PCR screening to identify approximate sites of insertion using the *Not*I miniprimer, and forward and reverse primers #2197 and #2198 (Table 2.4). The precise position of the insertion within the coding region was determined by cycle sequencing (section 2.6.3).

2.9 3D STRUCTURAL IMAGES

Mutants constructed in section 2.5.5 post sequencing (section 2.6.2) were mapped onto 3D images of $WzzB_{ST}$ (PDB number 3b8p), FepE (PDB number 3b8n), and WzzE (PDB number 3b8o) (Tocilj *et al.*, 2008). The 3D images of the Wzz_i mutant insertion locations were created using Pymol ™ software (DeLano Scientific LLC 2008).

2.10 PROTEIN TECHNIQUES

2.10.1 SDS POLYACRYLAMIDE GEL ELECTROPHORESIS (SDS-PAGE)

Whole cell lysate samples were prepared by harvesting 5 x 10^8 cells in 1 x sample buffer (Table 2.3) by centrifugation at 13000 rpm for 3 minutes (Eppendorf centrifuge 5415-R). Samples were solubilised at 100ºC for 5 min, then proteins were separated on an SDS 15% polyacrylamide gels in PAGE running buffer at 200 V using a 14.5 cm Vertical Gel Electrophoresis Unit (Sigma). The BenchMark™ Pre-stained protein ladder molecular weight marker was used to identify sizes of resulting bands (Invitrogen, cat. number 10748-010). The BenchMark marker sizes are: 181.8 kDa, 115.5 kDa, 82.2 kDa, 64.2 kDa (pink orientation band), 48.8 kDa, 37.1 kDa, 25.9 kDa, 19.4 kDa, 14.8 kDa and 6.0 kDa (http://tools.invitrogen.com/content/sfs/manuals/10748010.pdf).

2.10.2 WESTERN TRANSFER AND IMMUNOBLOTTING

Proteins were transferred to 0.45 μm nitrocellulose membranes (GE Water and Process Technologies) for 2 h at 200 mA (14.5 cm Vertical Gel Electrophoresis Unit, Sigma). The nitrocellulose membrane was blocked in TTBS containing 5% (w/v) skim milk at room temperature, followed by a 16 h incubation with primary antibody in TTBS containing 2.5% (w/v) skim milk. The primary and secondary antibodies used in this project are listed in table 2.5. The membrane was washed 3 x in TTBS and incubated with secondary antibody (Table 2.5) (in TTBS containing 1% (w/v) skim milk) for 2 h at room temperature, washed again in TTBS for 3 x 5 minutes, and finally washed in 1 x TBS for 3 x 5 minutes. Detection was achieved with BM chemiluminescence ELISA substrate (POD) reagents (Roche) or Chemiluminescent Peroxidase substrate-3 (Sigma) as described by the manufacturer.

Table 2.5 Antibodies used in this study

Primary antibodies	**Concentration**	**Dilution**	**Source/reference**
Mouse monoclonal anti-His_6	0.2 mg/mL	1:1000	Novagen (cat. number 70796)
Rabbit polyclonal anti-Wzz_{SF}	-	1:1000	Daniels and Morona, 1999
Mouse monoclonal anti-FLAG M2	1 mg/mL	1:2000	Sigma-aldrich (cat. number F1804)
Mouse monoclonal anti-StrepII	200 μg/mL	1:1000	Novagen/Merck (cat. number 71590-3)
Mouse monoclonal (x2) anti-GFP	0.4 mg/ml	1:1000	Roche (cat. number 11814460001)
Secondary antibodies	**Concentration**	**Dilution**	**Source/reference**
HRP-conjugated goat anti-mouse IgG	1 mg/mL	1:30,000	KPL (cat. number 074-1806)
HRP-conjugated goat anti-rabbit IgG	1 mg/mL	1:30,000	KPL (cat. number 474-1506)

2.10.3 OVER-EXPRESSION AND PURIFICATION OF HIS_6-TAGGED PROTEIN

2.10.3.1 IPTG INDUCTION OF HIS_6-TAGGED PROTEINS

The pQE-30-based constructs in this study contained an N-terminal His_6 tag, and expression of His-tagged proteins was achieved with IPTG (BioVectra). Cultures carrying these constructs were grown for 16 h at 37°C with aeration and the appropriate antibiotics, then subcultured 1/50 in 10 mL LB and grown for a further 2 h at 37°C. Cultures were then induced with 0.5 mM IPTG for 1-1.5 h at 37°C with aeration. Strains prepared for microscopy were induced with 0.5 mM IPTG for 1.5 or 3 h.

2.10.3.2 HIS_6-TAGGED PROTEIN PURIFICATION WITH NI-NTA RESIN

For purifying His_6-tagged protein, overnight 16-18 h cultures were diluted 1/50 in 100 mL and subcultured at 37°C with aeration for 2 h, then induced with 0.5 mM IPTG for a further 1-1.5 h. For the co-purification assays which included both pBAD33 and pQE-30-based plasmids, only 0.2% (w/v) L-arabinose was used for induction conditions. Cells were centrifuged at 8000 rpm for 15 minutes at 4°C (Beckman centrifuge J2-21M) and washed in 50 mM Tris, pH 7.5, resuspended in 30 mL of 10 mM HEPES/1 mM $MgCl_2$ and passed once through a chilled French press chamber (FRENCH pressure cell press, SLM Aminco Instruments). Intact cells and inclusion bodies were removed with centrifugation at 4500 rpm for 15 minutes (Sigma 3K15) and the supernatant was ultracentrifuged at 35000 rpm for 1 hour at 4°C (Beckman Coulter Optima L-100 Ultracentrifuge). Membrane pellets were resuspended in 400 μL of 1% DDM (Table 2.3) and solubilised at 25°C with vigorous shaking for 30 minutes. The soluble fraction was then ultracentrifuged at 40000 rpm for 1 hour at 4°C (Optima TLX ultracentrifuge, TLA 100.4 rotor). Supernatants were collected and applied to Pierce polypropylene columns (Pierce, cat. number 29922) containing 200 μL of Ni-NTA resin (QIAGEN, cat. number 32010) that had been pre-washed in equilibration buffer (Table

2.3). The column was gently agitated at RT for 1 – 1.5 h, and was secured on a metal stand with clamps. The Ni-NTA was washed twice in 1 mL of wash buffer A, and once with 1 mL wash buffer B (Table 2.3). 3-4 elution fractions were collected to remove the bound His_6-tagged protein from the resin using 200 μL elution buffer (Table 2.3). The eluted protein samples were either stored at 4°C for short periods (<4 days) or added to 50% glycerol for storage at -20°C.

2.10.4 OVER-EXPRESSION OF FLAG-TAGGED PROTEINS

Cultures of strains carrying pBAD33-based constructs were grown for 16-18 h at 37°C with aeration, and diluted 1/50 in 10 mL LB containing the appropriate antibiotics and 0.2% (w/v) glucose, and grown to an OD_{600} of 0.2-0.4. Cultures were then centrifuged at 4500 rpm for 10 minutes (Sigma centrifuge 3K15) and washed twice in 10 mL of LB to remove glucose. Cells were resuspended in 10 mL of LB supplemented with 0.2% (w/v) arabinose and the appropriate antibiotics, and grown for a further 1-1.5 h at 37ºC with aeration.

2.10.5 OVEREXPRESSION OF STREPII AND GFP^+-TAGGED PROTEINS

Cultures carrying constructs based on vector pSCRhaB2 were grown in MH broth (supplemented with antibiotics and 0.2% (w/v) glucose) at 37ºC with aeration for 16-18 h and diluted 1/50 in MH broth and incubated for a further 2 h at 37ºC with aeration. The cells were washed twice in 10 mL MH broth, by centrifugation at 4500 rpm for 10 minutes (Sigma 3K15). Cells were resuspended in 10 mL MH broth supplemented with antibiotics and 0.2% (w/v) rhamnose (Sigma, cat. number 83650) to induce expression of StrepII or GFP^+-tagged proteins and incubated for 1.5-2 h at 37°C with aeration. Alternatively, cultures were induced at 25°C as follows: overnight cultures (prepared as described above) were diluted 1/50 in 100

mL of MH broth and induced with 0.2% (w/v) rhamnose at 25°C with aeration for 8 h. Cells were centrifuged at 8000 rpm for 15 minutes at 4°C (Beckman centrifuge J2-21M) and washed in 50 mM Tris, pH 7.5, resuspended in 30 mL of 10 mM HEPES/1 mM $MgCl_2$ and passed once through a chilled French press chamber (FRENCH pressure cell press, SLM Aminco Instruments). Intact cells and inclusion bodies were removed with centrifugation at 4500 rpm for 15 minutes (Sigma 3K15) and the supernatant was ultracentrifuged at 35000 rpm for 1 hour at 4°C (Beckman Coulter Optima L-100 Ultracentrifuge). Membrane pellets were resuspended in 400 µL of 1% (w/v) DDM (Table 2.3) and solubilised at 25°C with vigorous shaking for 30 minutes. A positive control used for Western immunoblotting with anti-GFP was CV-1 lysate containing GFP-WH1 protein (pCB6-GFP-WH1, (N-WASP WH1 domain fused to GFP, Prof. Michael Way, EMBL, Heidelberg, Germany)).

2.10.6 FORMALDEHYDE CROSS-LINKING

Cultures were grown for 16-18 h, subcultured and induced for His-tagged protein expression as described in section 2.10.3.1. The formaldehyde protocol was based on experiments conducted by Prossnitz *et al.* (1988). Cells were harvested from each strain (5 x 10^8) and centrifuged at 7000 rpm for 8 minutes (Eppendorf 5417R), washed in 1.2 mL of chilled 10 mM K_2PO_4/KH_2PO_4 pH 6.8 (Table 2.3), and incubated at 25°C with 0.5% (w/v) formaldehyde (Sigma) for 1 hour. Cells were washed again in buffer and resuspended in 80 µL sample buffer (Table 2.3). Samples for SDS-PAGE were heated at either 60°C or 100°C for 5 minutes prior to electrophoresis.

2.11 LIPOPOLYSACCHARIDE TECHNIQUES

2.11.1 PREPARATION OF LPS SAMPLES

Cultures were diluted 1/50 in LB and incubated at 37ºC for approximately 3 h with aeration. Cells were standardised by OD_{600} to harvest 1 x 10^9 cells, and was pelleted by centrifugation at 15000 rpm for 10 minutes (Heraeus Sepatech Biofuge 15) and resuspended in 50 µL of lysing buffer (Table 2.3). After 10 minutes incubation at 100ºC, 10µL of 2.5 mg/mL proteinase K (Invitrogen) solution was added and samples were incubated for ~16 h at 56ºC. LPS samples were stored at -20ºC.

2.11.2 ANALYSIS OF LPS BY SILVER-STAINED SDS-PAGE

Silver-staining was performed using the method described by Tsai and Frasch (1982) with minor changes. LPS samples were incubated at 100ºC for 5 min prior to loading 8-10 µL on SDS 15% (w/v) polyacrylamide gels using Sigma Slab vertical gel electrophoresis unit (28572-00), with glass plates of dimensions 16.5 cm x 22 cm. Samples were electrophoresed at 12 mA for ~18 h or until the dye front was eluted from the gel. Gels were placed in a shallow glass container and 200 mL of fixing solution (40% (v/v) ethanol, 10% (v/v) acetic acid, Table 2.3) was applied. The gels were fixed for 2.5 h with gentle agitation on a rotating shaker Stuart Orbital Shaker SSL1 (Crown Scientific, cat. number SSL1). The fixing solution was discarded, and oxidising solution (0.7% (v/v) periodic acid in 40% (v/v) ethanol, 10% (v/v) acetic acid, Table 2.3) was applied for 5 mins. After 1-1.5 h of washing in MQ water (changed at 15 min intervals), the gel was stained for 10 min in staining solution (sodium hydroxide (0.1 M), 2 ml ammonium hydroxide (30% (w/v)) and 5 ml of silver nitrate (20% (w/v), Table 2.3) and washed again as described above for 1 hour. The gel was developed with pre-warmed (42°C) formaldehyde solution (50 mg/ml citric acid, and 0.05% (w/v)

formaldehyde, Table 2.3), and the reaction was stopped with the application of 5% acetic acid in MQ water (Table 2.3).

2.12 COLICIN SENSITIVITY ASSAY

The double layer colicin sensitivity assay was carried out as described by Renato Morona (1982) and Masi *et al.* (2007). Briefly, 16 hour cultures of colicin-producing strains were grown without aeration at 37°C and were used to streak parallel lines of growth across 20 mL LB agar plates. Plates were incubated at 37°C for 16 h and the strains were killed with chloroform added to a piece of Whatman paper placed inside the lid of the upturned plate by incubation at room temperature for 30 min. Plates were then overlaid with 20 ml LB agar (including Amp and 0.5 mM IPTG for pQE-30 based constructs) or 20 ml MH agar (including trimethoprim and 0.2% (w/v) rhamnose for pSCRhaB2-based plasmids), allowed to set, and the strains to be tested streaked across and perpendicular to the original streak line. Plates were incubated at 37°C for 16 h and zones of inhibitory growth by colicin-sensitive bacteria documented. For Chapter 3 graphs, the ratio of the zone of clearance for each strain was obtained by dividing the clearance size by the size of the colicin E2 streak. An average ratio was obtained for each sample and Class, and was graphed (Y axis) against the Wzz_i mutant classes (X axis). Statistical significance of the ratios were calculated by performing a one way analysis of variance (ANOVA, Tukey's multiple comparison test) with GraphPad Prism version 5.03 (GraphPad Software, Inc, 1992-2010).

2.13 FLUORESCENT MICROSCOPY TECHNIQUES

For fluorescence microcopy, bacterial culture was induced either with IPTG (section 2.10.3) or rhamnose (section 2.10.5) as described, and 10 mL of culture were centrifuged at 13000 rpm for 3 minutes (Eppendorf 40 centrifuge 5415-R), washed once in filter sterilized 1

x PBS in MQ (Table 2.3), spun again and resuspended in 1 mL 3.7% (w/v) formaldehyde (Sigma, cat. number F8775) in saline (0.85% (w/v) NaCl) and incubated at room temperature for 20 minutes. The cells were centrifuged, washed again in 1 x PBS and resuspended in 1 mL of 1 x PBS. Coverslips were treated with 0.01% (w/v) Poly-L-lysine (Sigma) and rinsed with 1 x PBS. Cells were adhered to the coverslips by centrifugation in a 24-well flat bottomed tray at 3000 rpm for 10 minutes (Heraeus Labofuge 400R), washed with 1 x PBS and the coverslips were mounted onto glass microscope slides with Mowiol 4-88 (Calbiochem) supplemented with 20 μg/ml *p*–phenylenylenediamine (Sigma) prepared fresh each experiment (at a ratio of 1:6, PPD:Mowiol) and sealed with nail polish. Bacteria were visualised using 100 x 41 phase contract and oil immersion objective lens under an Olympus IX-70 microscope connected to a Hamamatsu ORCA-ER camera and controller, an ASI MFC-2000 automatic focus fine-tuning instrument, a Sutter Lambda 10-C filterwheel and controller, and a Vincent Uniblitz-2 VMM-D3 three channel shutter drive. The filters utilised were Texas Red, for red fluorescence, and FITC for green fluorescence. The resulting images were observed using Metamorph v6.3r7 software (Molecular Devices).

CHAPTER THREE

IN-FRAME LINKER MUTAGENESIS OF WZZ_{SF}

3.1 INTRODUCTION

Previous mutagenesis studies have yielded data on critical and non-critical structural regions of Wzz and homologous proteins (e.g., FepE, Wzz_{PHS-2}, $WzzB_{EC}$) that affect its function in Oag modal chain length regulation (see section 1.4.2). Franco *et al.* (1998) investigated the association of LPS Oag modal chain length and residues in the region of amino acid 220 in both *S. flexneri* and *E. coli*. Their results indicated that various residues such as isoleucine 224 appear to be involved in determining Oag chain length, as the I224V mutation in the *S. dysenteriae* Wzz was found to slightly shorten Oag modal chain length from I-type (10-18 RUs) LPS to S-type (7-16 RUs) (Franco *et al.*, 1998). Similarly, four residues (M77, Q83, D90 and L91) in *S. enterica* Wzz which are also common to L-type *E. coli*, were singularly substituted with the I-type *E. coli* Wzz residues, to deduce whether these changes decrease the LPS Oag modal chain length. The results indicated that only a D90E substitution shifted Oag modal chain length from L to I type LPS, and that a L91I substitution also reduced the LPS Oag modal chain length. Other studies have indicated the significance of the TM region conserved residues, illustrating that glycine to alanine alterations at residues G305 and G311 considerably reduce the LPS Oag modal chain length from S-type to VS type (3-8 RUs) (Daniels and Morona, 1999). These single substitution or double substitution mutations, however informative, mostly result in subtle changes to the LPS Oag modal chain length, and other alterations to residues that are located in the Wzz periplasmic region failed to alter the LPS Oag modal chain length phenotype. Hence, in order to be able to instill significant alterations on the resulting LPS Oag modal chain length for the purpose of

identifying structure-function relationships in Wzz, in-frame linker mutagenesis was undertaken to create a library of Wzz_{SF} mutants containing 5-aa insertions within the protein.

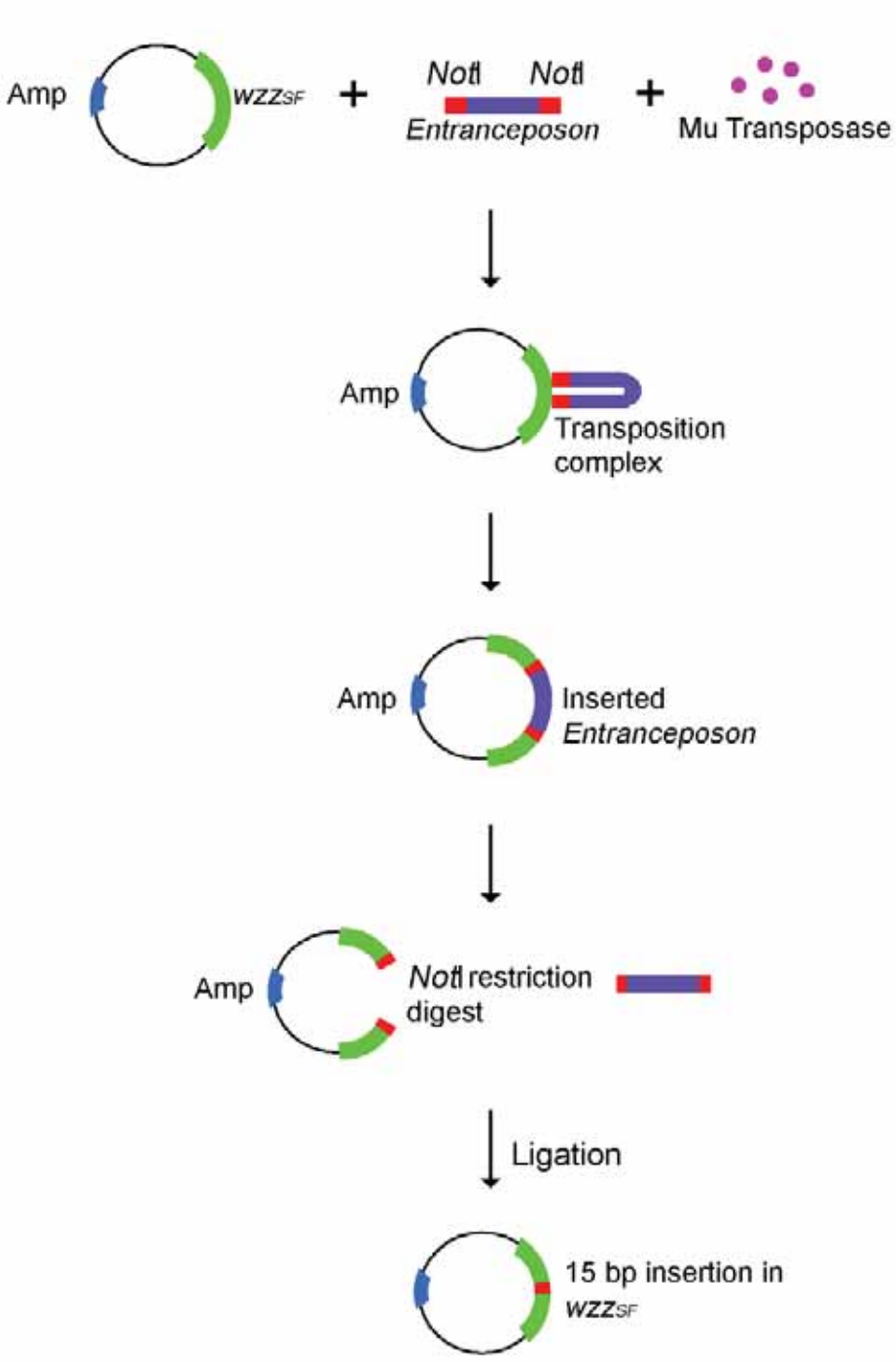
3.2 LINKER MUTAGENESIS

The Mutagenesis Generation System® (Finnzymes) was employed to produce a collection of in-frame *wzz*$_{SF}$ mutant constructs, as described in section 2.8. The process of the linker mutagenesis is illustrated in Figure 3.1. The strain Top10F' was used for cloning. Plasmid pRMCD30, a construct based on pQE-30 incorporating ampicillin resistance, an N-terminal His 6x tag and containing the *wzz*$_{SF}$ coding region with flanking *Sac*I and *Sma*I sites, was incubated with a Mu transposase and a 1,131 bp insertion (*entranceposon*) containing a kanamycin resistance cartridge having two *Not*I restriction sites at the 5' and 3' end of the insertion. The *entranceposon* is randomly inserted via the formation of a transposition complex assisted by the enzymatic action of the Mu transposase. Transformants from separate transformations were pooled and subcultured in LB overnight and plasmid DNA was extracted from each pool. A large scale restriction digest was conducted with *Not*I to remove the *entranceposon* (Figure 3.1). The restricted plasmids were re-ligated to allow the inclusion of a 15-bp insertion encompassing a *Not*I restriction site. This preparation was transformed into Top10F' and transformants were harvested to create pooled libraries MP1A, MP2A, MP3A and MP4A. Bacteria from these pools were plated out on LB + Amp to isolate single colonies, and sensitivity to kanamycin was assessed by patching on LB + Amp + Kan to ensure the absence of the *entranceposon*.

Over 90 isolates were screened via PCR using the *Not*I miniprimer (5'-TGCGGCCGCA-3'), which is complimentary to 10-bp of the 15-bp insertion, and primers which were specific to the 5' and 3' ends of *wzz*$_{SF}$ (#2197 forward primer and #2198 reverse primer, Table 2.4) to identify insertions within the *wzz*$_{SF}$ coding region. These were selected, and digests were conducted using a variety of restriction enzymes to further assess the

Figure 3.1 Construction of *wzz*$_{SF}$ mutant library

Plasmid pRMCD30 (Daniels and Morona, 1999), a pQE-30 based construct encoding His_6-tagged *wzz*$_{SF}$, was mutagenised with the *entranceposon* (section 2.8), allowing random insertion via the formation of a transposition complex. The *entranceposon* was eliminated from the plasmid by *Not*I restriction digestion, and re-ligation, and resulted in constructs containing only a 15-bp insertion within the gene. The Amp resistance gene is illustrated in the figure in blue, *wzz*$_{SF}$ is indicated in green, the *entranceposon* in purple and the *Not*I sites in red.


Amp
wzzSF
NotI
NotI
Entranceposon
Mu Transposase
Amp
Transposition
complex
Amp
Inserted
Entranceposon
Amp
NotI restriction
digest
Ligation
15 bp insertion in
wzzSF

location of the 15-bp insertion. Twenty five transformants tested positive by these methods, and DNA sequencing was utilised to determine the actual location of the 15-bp mutation. From these potential mutants, 18 unique mutants were selected for further analyses, as the remainder were found to be either duplicates or were mapped immediately outside of the *wzz*$_{SF}$ coding region (data not shown). Table 3.1 lists the nucleotide and amino acid sequences of the 18 15-bp insertion mutations (termed *wzz*$_i$ mutants). With the exception of i32, which is located directly at the commencement of TM1, the *wzz*$_i$ insertions are distributed throughout the Wzz$_{SF}$ periplasmic region (Figure 3.2B).

Due to the nature of the mutagenesis process, each insertion has a unique 5-aa sequence which depended on the *wzz*$_{SF}$ sequence that the 15-bp insertion had incorporated into, and the reading frame of the codon disruption; however, there exists a consensus sequence for each of these different reading frames. Frame 1 has a consensus sequence of CGRXX, frame 2 has a sequence of XRPH/QX, and frame 3 has a consensus sequence of XAAAX. Eight of the 18 mutants are in reading frame 1, 8 are in reading frame 2 and two of the 18 mutants are in frame 3 (Table 3.1). These 18 Wzz$_i$ mutants were further characterised as described below.

3.3 LPS MUTANT PROFILES

The pRMCD30 plasmids encoding mutant Wzz$_{SF}$ proteins containing various insertions (labelled Wzz$_i$ plasmids) were transformed into the strain RMA2741, an *S. flexneri* serotype Y strain with *lacI*^{q} (to repress expression from the *T5/lac* promoter in pQE-30) and a *wzz*$_{SF}$::Km^R mutation, but lacking both pHS-2 (encoding Wzz$_{pHS-2}$) and the large virulence plasmid (Table 2.1). The effect of each mutant Wzz$_i$ protein on LPS Oag modal chain length was assessed by SDS PAGE and silver staining. The strains were cultured overnight (16-18 h), subcultured and induced with IPTG to express the His$_6$-tagged Wzz$_i$ proteins (section 2.10.3). LPS samples were prepared and electrophoresed on a SDS 15% polyacrylamide gel

Table 3.1 The nucleotide and 5-aa sequence of the insertions in Wzz_i mutants, corresponding reading frames and resulting LPS Oag modal chain length conferred.

Wzz_i mutants							**Frame**	**Modal chain length (RUs)**	**Class**
$_i$32	M32	T	A	A	A	M	T33	2	Random	I
	ATG	ACT	GCG	GCC	GCA	ATG	ACA			
$_i$66	D66	C	G	R	T	D	V67	1	Random	I
	GAC	TGC	GGC	CGC	ACC	GAC	GTG			
$_i$80	I80	C	G	R	I	I	Y81	1	10-19	IV
	ATC	TGC	GGC	CGC	ATT	ATC	TAT			
$_i$81	Y81	C	G	R	I	Y	G82	1	10-19	IV
	TAT	TGC	GGC	CGC	ATC	TAT	GGT			
$_i$92	Q92	V	R	P	Q	Q	E93	2	8-13	III
	CAG	GTG	CGG	CCG	CAG	CAG	GAG			
$_i$128	Q128	L	R	P	H	Q	Q129	2	8-14	V
	CAG	CTG	CGG	CCG	CAC	CAG	CAA			
$_i$131	P131	L	R	P	Q	P	L132	2	14-22	V
	CCA	TTG	CGG	CCG	CAA	CCA	TTG			
$_i$138	G138	C	G	R	I	G	Q139	1	16-25	III
	GGG	TGC	GGC	CGC	ATT	GGG	CAA			
$_i$161	Q161	D	A	A	A	Q	E162	3	Random	I
	CAA	GAT	GCG	GCC	GCA	CAA	GAG			
$_i$163	L163	V	R	P	Q	L	E164	2	Random	I
	CTA	GTG	CGG	CCG	CAG	CTA	GAA			
$_i$191	E191	L	R	P	Q	E	Q192	2	2-8	II
	GAG	CTG	CGG	CCG	CAG	GAG	CAG			
$_i$199	Q199	I	A	A	A	Q	I200	3	Random	I
	CAG	ATT	GCG	GCC	GCA	CAG	ATT			
$_i$219	Q219	M	R	P	Q	Q	T220	2	2-10	II
	CAG	ATG	CGG	CCG	CAG	CAG	ACT			
$_i$231	L231	V	R	P	H	L	G232	2	2-6	II
	CTA	GTG	CGG	CCG	CAT	CTA	GGG			
$_i$247	P247	C	G	R	S	P	L248	1	2-10	II
	CCG	TGC	GGC	CGC	AGT	CCG	TTG			
$_i$255	Y255	C	G	R	N	Y	Q256	1	2-8	II
	TAT	TGC	GGC	CGC	AAC	TAT	CAG			
$_i$279	Y279	C	G	R	S	Y	V280	1	Random	I
	TAT	TGC	GGC	CGC	AGC	TAT	GTG			
$_i$290	D290	C	G	R	S	D	S291	1	Random	I
	GAT	TGC	GGC	CGC	AGC	GAT	AGT			

and silver stained as described in section 2.11. The resulting LPS Oag modal chain lengths were scored for each Wzz_i mutant and five different phenotypic classes were identified (Table 3.1), and representative samples from each class are illustrated in Figure 3.2. Seven of the 18 mutants, categorised into Class I, had lost the ability to confer wild-type LPS Oag modal chain length of 11-17 RUs, hence the resulting profile displayed random LPS Oag modal chain length distribution, and is represented by Wzz_{i32} (Figure 3.2A, lane 3). Five mutants displayed significantly reduced LPS Oag modal chain length, bringing the average Oag modal chain length from a wild-type length of 10-17 RUs down to a range between 2-10 Oag RUs, designated as very short (VS) type Oag (Table 3.1). This class (Class II) is represented by mutant Wzz_{i219} (Figure 3.2A, lane 4). The third mutant class observed (and designated Class III), displayed slightly reduced Oag modal chain length (Shorter type Oag, 8-14 RUs), and is represented by Wzz_{i92} (Figure 3.2A, lane 5). Two mutants conferred a near wild-type LPS Oag modal chain length between 11-19 RUs, termed Class IV and represented by mutant Wzz_{i81} in Figure 3.2A (lane 6). The last phenotypic class, designated Class V, increased the resulting Oag modal chain length from wild-type modal length to 16-25 RUs, represented by Wzz_{i131} (Figure 3.2A, lane 7). The approximate locations of the mutations in Wzz_{SF} are presented in Figure 3.2B, which also illustrates structural elements in Wzz_{SF} and the approximate location of the insertions, and TM1 and TM2.

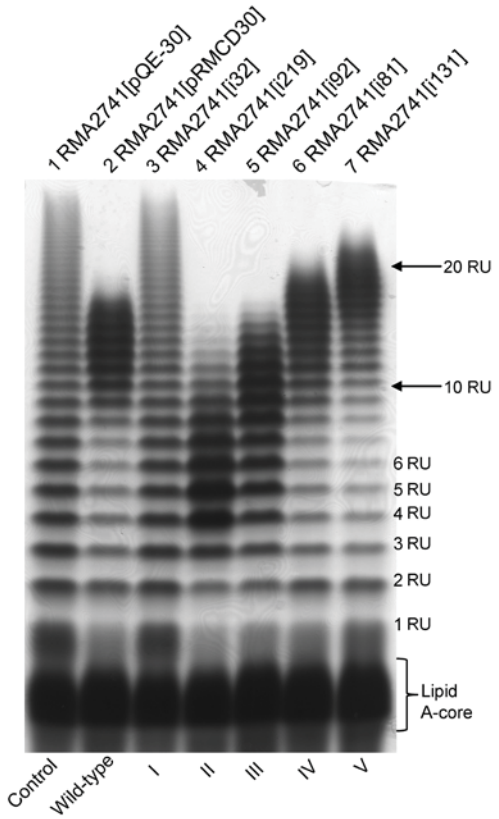
3.4 EXPRESSION OF WZZ_i MUTANT PROTEINS

The *S. flexneri* strains carrying plasmids encoding the mutant Wzz_i proteins were analysed for the ability to express each protein to determine whether the insertion mutations affected protein production which may have contributed to the observed LPS phenotype. Western immunoblotting using affinity purified polyclonal anti-Wzz_{SF} was performed on whole cell lysates of IPTG induced log phase cultures of these strains (section 2.10.1 and 2.10.2). Western immunoblotting indicated that the majority of the Wzz_i mutant proteins

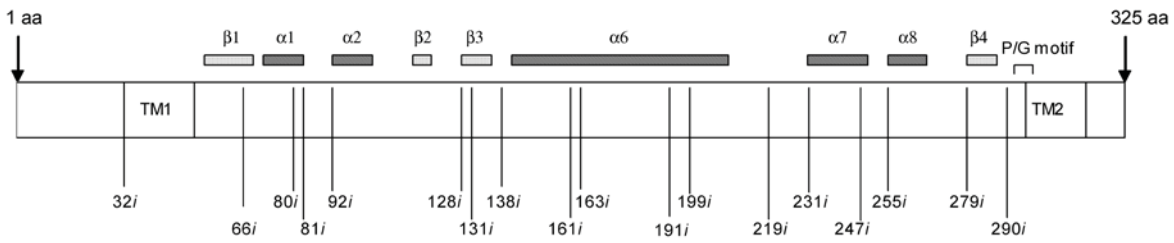
Figure 3.2 The LPS Oag modal chain length conferred by the different classes of Wzz_i mutants expressed in *S. flexneri* RMA2741 and schematic representation of the insertion locations in Wzz_{SF}

A) Strains were grown in LB + Amp, induced with IPTG for 1.5 h and LPS samples were prepared, electrophoresed on a SDS 15% polyacrylamide gel and silver stained as described in methods (section 2.11). Strains in each lane are as follows: 1) RMA2741 (pQE-30), 2) RMA2741 (pRMCD30), 3) RMA2741 (Class I, i32), 4) RMA2741 (Class II, i219), 5) RMA2741 (Class III, i92), 6) RMA2741 (Class IV, i81), 7) RMA2741 (Class V, i131). The LPS Oag repeat units (RUs) are numbered on the right hand side. B) A schematic representation of the locations of each insertion within Wzz_{SF}, illustrating the insertions within Wzz_{SF} that were designated titles according to the last uninterrupted amino acid preceding the 5-aa insertion. The proline rich motif is indicated (P/G motif). Secondary structural features (indicating α helices and β strands) are based on the 3D structure of $WzzB_{ST}$ (PDB number 3b8p). The transmembrane regions are also indicated (TM1 and TM2).

A


1 RMA2741[pQE-30]
2 RMA2741[pRMCD30]
3 RMA2741[i32]
4 RMA2741[i219]
5 RMA2741[i92]
6 RMA2741[i81]
7 RMA2741[i131]
20 RU
10 RU
6 RU
5 RU
4 RU
3 RU
2 RU
1 RU
Lipid
A-core
Control
Wild-type
I
II
III
IV
V

B


1 aa
325 aa
β1
α1
α2
β2
β3
α6
α7
α8
β4
P/G motif
TM1
TM2
32i
66i
80i
81i
92i
128i
131i
138i
161i
163i
191i
199i
219i
231i
247i
255i
279i
290i

could be detected, however four strains (two from Class I, i163 and i279; Figure 3.3 lane 10 and 16), and two from Class II, (i191 and i255; Figure 3.3 lanes 13 and 17) did not produce detectable levels of Wzz_{SF} related protein. Class I (i161), Class II (i231), Class II (i247) and Class III (i138) produced detectable levels of Wzz_i protein at a diminished level compared to wild-type (Figure 3.3, lanes 14, 11, 21 and 22).

3.5 COLICIN SENSITIVITY ASSAY

To further characterise differences between the LPS phenotype conferred by the Wzz_i mutants, *S. flexneri* strains carrying the plasmids encoding Wzz_i mutants were subjected to a colicin sensitivity assay to determine susceptibility to the lethal action of colicin E2. The double layer assay was performed as described previously (Morona 1982; Masi *et al.,* 2007). Strain RMA2782 (*E. coli* E95 strain carrying a plasmid producing colicin E2), was grown overnight and was used to create a streak of growth across LB agar. After the strain was chloroform killed and the plates were overlaid with another agar layer (containing Amp and IPTG, section 2.12), *S. flexneri* strains carrying pQE-30, pRMCD30 or one of the Wzz_i encoding plasmids were streaked perpendicularly across the original streak, allowed to grow for 16 h and the zones of inhibitory growth by colicin-sensitive strains were documented by measuring the resulting zones of killing. Due to small variations of the dimensions of the colicin E2 streak, the sizes of the zones of killing were divided by the colicin E2 streak size to acquire the resulting ratio of the zone of clearance (section 2.12). The average clearance zone ratio for each Class was graphed (Y axis) against the Classes (X axis), as were clearance zone average ratios for the controls (RMA2741 (pQE-30) and RMA2741 (pRMCD30)). The average clearance zone ratios for RMA2741 (pQE-30), Class I and Class II ratios were the largest (3.1, 2.9 and 3.0 respectively, Figure 3.4), which indicated that these strains were sensitive to colicin activity. The Class III clearance zone ratio was 2.5 (Figure 3.4). The clearance zone average ratio for RMA2741 (pRMCD30) was 1.7 (Figure 3.4), and the Class

Figure 3.3 Detection of Wzz_i production in *S. flexneri*

RMA2741 strains harbouring Wzz_i plasmids were grown in LB + Amp, induced with IPTG for 1.5 h and whole cell lysates were prepared and electrophoresed on a SDS 15% polyacrylamide gel (section 2.10.2). Affinity purified Wzz_{SF} polyclonal antibodies (Daniels and Morona, 1999) were used to detect Wzz-related protein at a dilution of 1:1000 (section 2.10.2) and sizes were determined using the Benchmark prestained molecular weight marker (Invitrogen). Each lane contained approximately 1 x 10^8 bacterial cells. Lanes are as follows: 1) RMA2741 (pQE-30), 2) RMA2741 (pRMCD30), 3) RMA2741 (Wzz_{i80}), 4) RMA2741 (Wzz_{i199}), 5) RMA2741 (Wzz_{i66}), 6) RMA2741 (pQE-30), 7) RMA2741 (pRMCD30), 8) RMA2741 (Wzz_{i92}), 9) RMA2741 (Wzz_{i128}), 10) RMA2741 (Wzz_{i163}), 11) RMA2741 (Wzz_{i231}), 12) RMA2741 (Wzz_{i290}), 13) RMA2741 (Wzz_{i191}), 14) RMA2741 (Wzz_{i161}), 15) RMA2741 (Wzz_{i32}), 16) RMA2741 (Wzz_{i279}) 17) RMA2741 (Wzz_{i255}), 18) RMA2741 (Wzz_{i131}), 19) RMA2741 (Wzz_{i81}), 20) RMA2741 (Wzz_{i219}), 21) RMA2741 (Wzz_{i247}) and 22) RMA2741 (Wzz_{i138}).

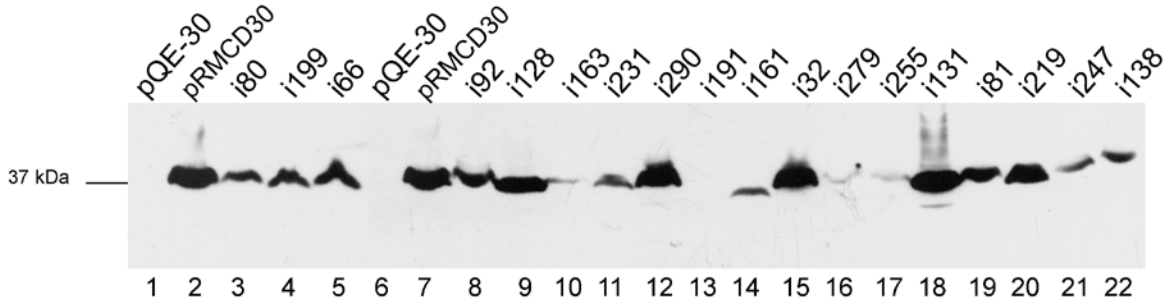
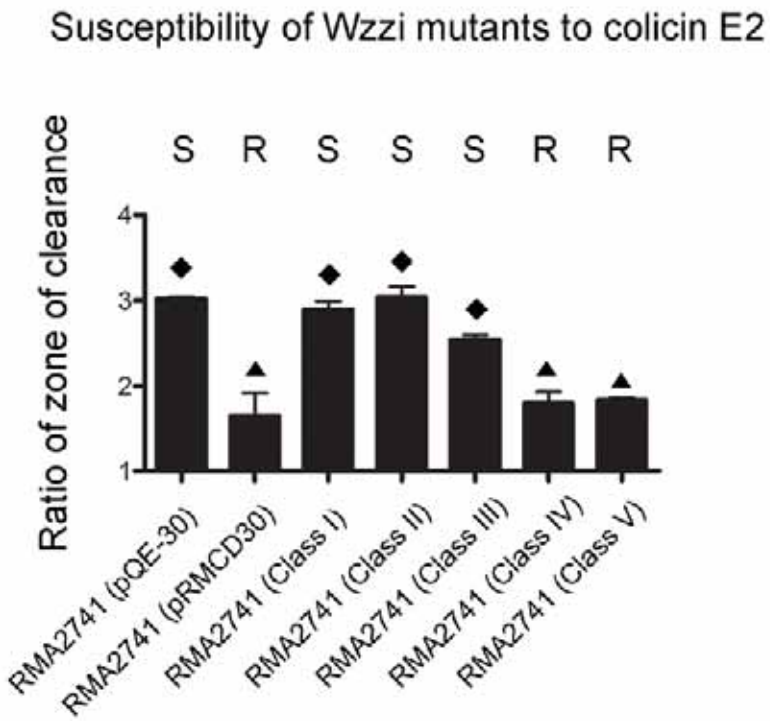

pQE-30
pRMCD30
i80
i199
i66
pQE-30
pRMCD30
i92
i128
i163
i231
i290
i191
i161
i32
i279
i255
i131
i81
i219
i247
i138
37 kDa
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22

Figure 3.4 Wzz_i susceptibility to the effects of colicin E2

S. flexneri strain RMA2741 carrying pQE-30, pRMCD30 or Wzz_i plasmids were grown in LB + Amp, induced with 0.5 mM IPTG for 1.5 h and streaked on an upper layer of agar in LB plates, while the base agar layer had a streak of chloroform-killed RMA2782 (a colicin E2-producing strain), as described in section 2.12. The cleared regions of the streaked Wzz_i strains proximal to the RMA2782 streak were measured, and the zone of clearance ratio was acquired for each Wzz_i mutant by dividing the size of the proximal clearance zone by the size of the RMA2782 streak. The average clearance zone ratio for each Wzz_i mutant class was graphed (Y axis) against the classes (X axis). The mean and standard errors are indicated in the graph. Experiment performed 3 times (n=3). 'S' indicates susceptibility to colicin E2 and 'R' indicates resistance. ◆ values are not significantly different compared to each other, however are significant compared to, ▲ as determined by one way ANOVA ($p<0.05$) (section 2.12).


Susceptibility of Wzzi mutants to colicin E2
Ratio of zone of clearance
S R S S S R R
4
3
2
1
RMA2741 (pQE-30)
RMA2741 (pRMCD30)
RMA2741 (Class I)
RMA2741 (Class II)
RMA2741 (Class III)
RMA2741 (Class IV)
RMA2741 (Class V)

IV mutants which conferred a near wild-type LPS Oag modal chain length exhibited a similarly small average clearance zone ratio (1.8) by RMA2741 (pRMCD30) (Figure 3.4). The Class V mutants had the same average zone of clearance (1.8) compared to Class IV, indicating low susceptibility to colicin E2 (Figure 3.4). The data obtained from these experiments indicate that strains producing Wzz_i mutants which confer wild-type or longer Oag modal chain lengths are more resistant to the effects of colicin E2 than stains conferring random or shorter Oag modal chain length. These data show that there is a strong correlation between LPS Oag modal chain length and sensitivity to colicin E2. A one way AVOVA (section 2.12) was performed on these values using GraphPad Prism version 5.03, and showed that differences displayed by RMA2741 (pQE-30), Class I, Class II, and Class III were not significant, however were statistically significant compared to RMA2741 (pRMCD30), Class IV and Class V ($p<0.05$).

3.6 CHEMICAL CROSS-LINKING OF WZZ_i MUTANT PROTEINS

Previous chemical cross-linking studies to assess the oligomeric properties of Wzz_{SF} have demonstrated that Wzz_{SF} is able to oligomerise (Daniels and Morona, 1999), and has been confirmed by 3D structural data (Tocilj *et al.*, 2008). Chemical cross-linking using formaldehyde was conducted on Wzz_{SF} wild-type and representative Wzz_i proteins (section 2.10.6) to determine if the mutants were capable of forming oligomers comparable to wild-type Wzz_{SF} or if the 5-aa insertions had influenced inter-monomeric interactions. The mutants selected for cross-linking analyses were i290 (Class I), i219 (Class II), i92 (Class III), and i128 and i131 (Class V). *S. flexneri* strains carrying plasmids encoding the Wzz_i proteins were grown, induced as described in section 2.10.3, and treated with 0.5% formaldehyde and subjected to SDS-PAGE and Western immunoblotting using an affinity purified polyclonal anti-Wzz_{SF} antibody (section 2.10.6, Table 2.5). The chemical cross-linking experiments were conducted on 3 occasions and yielded reproducible results; representative data from one of

these experiments is presented in Figure 3.5. The sizes of the detected cross-linked bands from all mutants analysed are summarised in Table 3.2. The data show that in non-treated samples, Wzz_{SF} wild-type monomeric (36 kDa) and dimeric forms (72 kDa) were readily detected, and an apparent tetrameric form (144 kDa) was also detected (Figure 3.5, lane 3). Wzz_{SF} wild-type samples cross-linked with formaldehyde revealed bands at ~30 kDa, 36 kDa, a doublet band at ~72 kDa and a smeared high molecular weight (HMW) band of >180 kDa, suggesting the presence of higher order oligomers (Figure 3.5, lane 4). The Class I mutant i290 had both monomeric and dimeric bands in both non cross-linked and cross-linked samples, and a higher MW band around the >180 kDa region (Figure 3.5, lanes 9 and 10). The presence of an extra band ~30 kDa was also detected in the non cross-linked sample of i290; this band was only ever detected after formaldehyde cross-linking for Wzz_{SF} wild-type and the other Wzz_i mutants (Figure 3.5, lanes 4, 6, 8, 12 and 14). There was also an additional i290 band at approximately 89 kDa which is not detected in other Wzz_i cross-linked profiles (Figure 3.5, lane 9). The Class II mutant i219 also displayed the presence of monomeric (36 kDa) and dimeric forms (72 kDa) of Wzz_i protein in cross-linked and non cross-linked samples, and the 30 kDa protein was also detected in the cross-linked sample (Figure 3.5, lanes 13 and 14; Table 3.2). HMW bands were decreased in intensity in the cross-linked sample as compared to wild-type. The Class III mutant i92 included in this experiment displayed a cross-linked profile comparable to wild-type Wzz_{SF}, having detectable bands of sizes 36 kDa, 72 kDa, and 144 kDa in the non-treated sample, and 30 kDa, 36 kDa and 72 kDa in the cross-linked sample (Figure 3.5, lanes 5 and 6, respectively). However, no high MW bands around the 180 kDa region were detected in the cross-linked sample, similar to the Class II mutant i219.

The two Class V mutants, i128 and i131 exhibited bands of sizes 36 kDa, 72 kDa and weak bands at ~144 kDa in the non-treated samples (Figure 3.5, lanes 7 and 11, respectively). However, cross-linking profiles of these two mutant proteins varied dramatically to the other Wzz_i proteins; the monomeric forms of Wzz_i (36 kDa) were

Table 3.2 Approximate sizes of bands detected in samples of cross-linked and non-cross-linked Wzz_{SF} and Wzz_i protein*

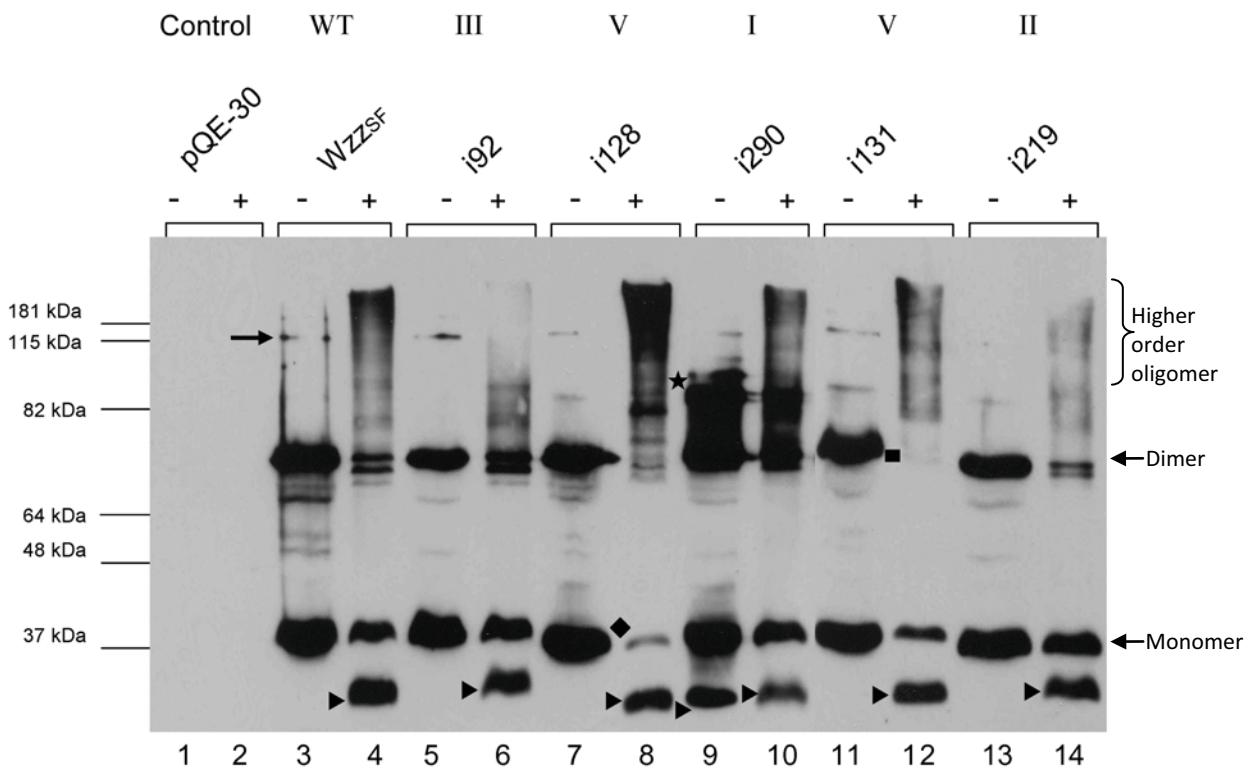
Wild-type		I		II		III		V		V		
Wzz_{SF}		$_i290$		$_i219$		$_i92$		$_i128$		$_i131$		
-	+	-	+	-	+	-	+	-	+	-	+	
	30	30	30		30		30		30		30	
36	36	36	36	36	36	36	36	36	↓36	36	↓36	Monomer
72	72	72	72	72	72	72	72	72	↓72	72		Dimer
		89	89									
144						144		↓144		↓144		Higher order oligomer
	180		↓180						180		180	

All values are relative MW in kDa, '+' denotes the addition of formaldehyde cross-linking, and '-' denotes the absence of cross-linking agent, ↓indicates lower intensity bands

* data from Figure 3.5

Figure 3.5 Analysis on Wzz_i mutants by cross-linking with formaldehyde

S. flexneri RMA2742 strains carrying plasmid encoded Wzz_i proteins were grown in LB + Amp, induced with IPTG for 1 h, harvested and washed in 10 mM KPO_4 and exposed to 0.5% (v/v) formaldehyde at 25ºC (+), controls were incubated without formaldehyde (-) as described in section 2.10.6. Samples were heated at 60ºC and electrophoresed on a SDS 12% polyacrylamide gel. Western immunoblotting was performed with affinity purified Wzz_{SF} (Daniels and Morona, 1999) polyclonal antisera at a dilution of 1:1000 (section 2.10.2). Sizes were determined with the Benchmark prestained molecular weight marker (Invitrogen). ► indicates the ~30-kDa form of Wzz_{SF}, ★ indicates the extra band in i290 (lane 9), ◆ indicates the depleted monomeric form of i128 (lane 8), and ■ indicates the lack of the dimeric form of i131 (lane 12). Each lane contained approximately 4 x 10^8 bacterial cells. Cross-linking data were reproduced in a minimum of 3 experimental repeats. The lanes are as follows: 1) RMA2741 (pQE-30) (-), 2) RMA2741 (pQE-30) (+), 3) RMA2741 (pRMCD30) (-), 4) RMA2741 (pRMCD30) (+), 5) RMA2741 (Wzz_{i92}) (-), 6) RMA2741 (Wzz_{i92}) (+), 7) RMA2741 (Wzz_{i128}) (-), 8) RMA2741 (Wzz_{i128}) (+), 9) RMA2741 (Wzz_{i290}) (-), 10) RMA2741 (Wzz_{i290}) (+), 11) RMA2741 (Wzz_{i131}) (-), 12) RMA2741 (Wzz_{i131}) (+), 13) RMA2741 (Wzz_{i219}) (-) and 14) RMA2741 (Wzz_{i219}) (+).


Control
WT
III
V
I
V
II
pQE-30
WZZSF
i92
i128
i290
i131
i219
- + - + - + - + - + - + - +
181 kDa
115 kDa
82 kDa
64 kDa
48 kDa
37 kDa
Higher order oligomer
Dimer
Monomer
1 2 3 4 5 6 7 8 9 10 11 12 13 14

significantly reduced in the cross-linked samples of both i128 and i131 (Figure 3.5, lanes 8 and 12, respectively), and the dimeric forms (72 kDa) were diminished in i128 and not detected at all in i131 cross-linked samples (Figure 3.5, lanes 8 and 12, respectively). Larger HMW bands were detected in i131 and easily detected in the cross-linked sample of i128, and were comparable to the HMW oligomers detected in the Wzz_{SF} wild-type cross-linking profile (Table 3.2).

3.7 STABILITY OF WZZ_{SF} WILD-TYPE AND WZZ_i DIMERS

In a previous study, it was established that Wzz_{SF} dimers could be detected even after solubilisation and SDS-PAGE, and also that formaldehyde cross-linked Wzz_{SF} dimers could be detected after being heated to 100°C (Daniels and Morona, 1999). *S. flexneri* strains carrying plasmids encoding representative Wzz_i mutants were grown, induced and cross linked with formaldehyde. Samples were solubilised and heated at 100°C and subjected to Western immunoblotting with anti-Wzz_{SF} to determine whether the mutants exhibited altered dimer stability comparable to Wzz_{SF} wild-type. The data indicate that Wzz_{SF} wild-type dimers were detected in the cross-linked samples, as were dimers of Class V mutants i128 and i131 (Figure 3.6, lanes 4, 8 and 14, respectively). Mutants i290, i219 and i92 from Classes I, II and III respectively, did not exhibit detectable dimers under these conditions (Figure 3.6, lanes 10, 16 and 6 respectively). Soluble His_6-Wzz_{SF} purified with Ni-NTA resin (section 2.10.3) was also included for comparison (Figure 3.6, lanes 11 and 12).

3.8 MAPPING OF WZZ_i INSERTIONS ONTO PCP 3D STRUCTURES

As recent studies have solved the 3D structures of PCP proteins WzzE, FepE and $WzzB_{ST}$ (Tocilj *et al.*, 2008), the location of the Wzz_i 5-aa insertions could be mapped onto these structures to obtain structure-function relationship information. Figure 3.7 presents the

Figure 3.6 Comparison of Wzz_{SF} and Wzz_i dimer stability

S. flexneri RMA2741 strains harbouring plasmid-encoded Wzz_i proteins were grown in LB + Amp and induced with IPTG for 1 h, harvested, washed in 10 mM KPO_4 and exposed to 0.5% (v/v) formaldehyde at 25ºC (+), controls were incubated without formaldehyde (-) as described in section 2.10.6. Both cross-linked and non cross-linked samples of Wzz_{SF} and Wzz_i mutants were heated to 100°C for 5 minutes and electrophoresed on a SDS 12% polyacrylamide gel and subjected to Western immunoblotting using affinity purified polyclonal anti-Wzz_{SF} (Daniels and Morona, 1999) at a dilution of 1:1000 (section 2.10.2). Prestained Benchmark protein marker was used to determine protein sizes (Invitrogen). Each lane contained approximately 1 x 10^8 bacterial cells. The lanes are as follows: 1) and 2) RMA2741 (pQE-30), 3) and 4) RMA2741 (pRMCD30), 5) and 6) RMA2741 (Class III, i92), 7) and 8) RMA2741 (Class V, i128), 9) and 10) (Class I, i290), 11) and 12) ~20µL soluble His_6-Wzz_{SF}, 13) and 14) RMA2741 (Class V, i131) and 15) and 16) RMA2741 (Class II, i219).

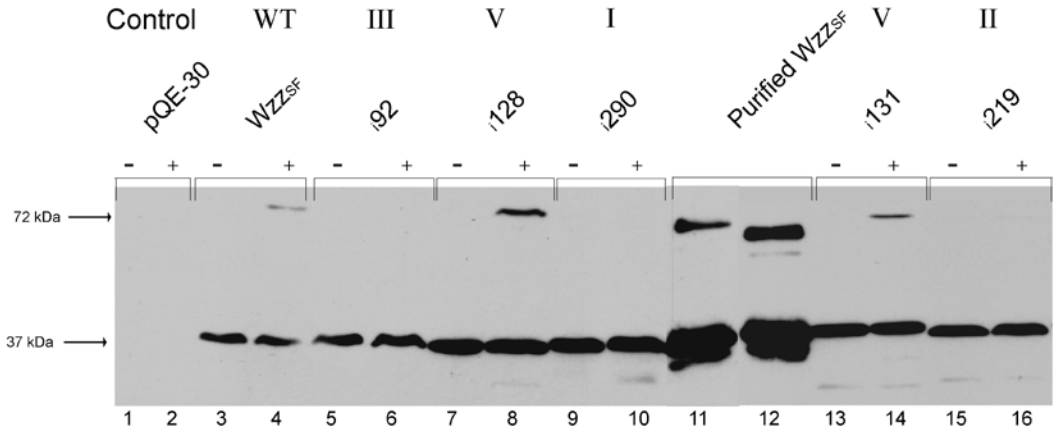

Control
WT
III
V
I
V
II
pQE-30
Wzz^{SF}
i92
i128
i290
Purified Wzz^{SF}
i131
i219
- + - + - + - + - + - + - +
72 kDa
37 kDa
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16

Figure 3.7 PCP1 protein alignments

The multiple sequence alignment of four PCP proteins, *S. typhimurium* Wzz ($WzzB_{ST}$ genbank NP_461024.1), *S. flexneri* Wzz (Wzz_{SF}, genbank CAA50783.1), *E. coli* WzzE (genbank NP_290416.1) and *E. coli* FepE (genbank NP_286314.1), using Clustal 2.0.12. Red residues indicate small hydrophobic amino acids, blue resides indicate acidic amino acids, magenta residues illustrate basic amino acids, and green residues indicate hydroxyl/basic/amine amino acids. '*' indicates positions which have a single, fully conserved residue. "." indicates where semi-conserved substitutions are observed, and ":"indicates where conserved substitutions have been observed, according to the colour specifications listed above (http://www.ebi.ac.uk/Tools/clustalw2/index.html). Numbers on the right represent the aa position.

```
WZZBST          MTVDSNTSSG---------RGNDPEQIDLIELLLQLWRGKMTIIVAVIIAILLAVGYLMI 51
WZZSF           MRVENNNVSG---------QNHDPEQIDLIDLLVQLWRGKMTIIISVIVAIALAIGYLAV 51
WZZE            MMTQPMPGKP---------AEDAENELDIRGLFRTLWAGKLWIIGMGLAFALIALAYTFF 51
FEPE            MSSLNIKQGSDAHFPDYPLASPSNNEIDLLNLISVLWRAKKTVMAVVFAFACAGLLISFI 60
                *                       :::*:  *:  ** .*  ::   :     .:    .

WZZBST          AKEKWTSTAIITQPDAAQVATYTN------ALNVLYGGNAPKISEVQANFISRFSSAFSA 105
WZZSF           AKEKWTSTAIITQPDVGQIAGYNN------AMNVIYGGQAAPKVSDLQETLIGRFSSAFSA 105
WZZE            ARQEWSSTAITDRPTVNMLGGYYSQQQFLRNLDVRSNMASADQPSVMDEAYKEFVMQLAS 111
FEPE            LPQKWTSAAVVTPPEPVQWQELEKSFTKLRVLDLDIKIDRTEAFNLFIKKFQSVSLLEEY 120
                  ::*:*:*:   *         .       :::      ..  .:       .

WZZBST          LS--------------EVLDNQKEREKLTIEQSVKGQALP----------LSVSYVSTT 140
WZZSF           LA--------------ETLDNQEEPEKLTIEPSVKNQQLP----------LTVSYVGQT 140
WZZE            WDTRREFWLQTDYYKQRMVGNSKADAALLDEMINNIQFIPGDFTRA---VNDSVKLIAET 168
FEPE            LRSSPYVMDQLKEAKIDELDLHRAIVALSEKMKAVDDNASKKKDEPSLYTSWTLSFTAPT 180
                                  :.  .    *  :    :  .           ::.  . *

WZZBST          AEGAQRRLAEYIQQVDEEVAKELEVDLKDNITLQTKTLQESLETQEVVAQEQKDLRIKQI 200
WZZSF           AEGAQMKLAQYIQQVDDKVNQELEKDLKDNIALGRKNLQDSLRTQEVVAQEQKDLRIRQI 200
WZZE            APDANNLLRQYVAFASQRAASHLNDELKGAWAARTIQMKAQVKRQEEVAKAIYDRRMNSI 228
FEPE            SEEAQTVLSGYIDYISTLVVKESLENVRNKLEIKTQFEKEKLAQDRIKTKNQLDANIQRL 240
                :  *:  *  *:   .  . ..   :::.         : .:  :.  ::   * .:. :

WZZBST          EEALRYADEAKITQPQIQQ-TQDVTQDT--MFLLGSDALKSMIQNEATRPLVFSPA--YY 255
WZZSF           QEALQYANQAQVTKPQVQQ-TEDVTQDT--LFLLGSEALESMIKHEATRPLVFSPN--YY 255
WZZE            EQALKIAEQHNISRSATDVPAEELPDSE--MFLLGRPMLQARLENLQAVGPAFDLD--YD 284
FEPE            NYSLDIANAAGIKKPVYSN-GQAVKDDPDFSISLGADGIERKLEIEKAVTDVAELNGELR 299
                : :*  *:   :.:.  .   : : :.    : **   ::  ::   :   . .

WZZBST          QTKQTLLDIKNLKVTADTVHVYRYVMKPTLPVRRDSPKTAITLVLAVLLGGMIGAGIVLG 315
WZZSF           QTRQNLLDIEKLKFDDLDIHAYRYVMKPTLPIRRDSPKKAITLILAVLLGGMVGAGIVLG 315
WZZE            QNRAMLNTLNVGPTLDPRFQTYRYLRTPEEPVKRDSPRRAFLMIMWGIVGGLIGAGVALT 344
FEPE            NRQYLVEQLTKAHVNDVNFTPFKYQLSPSLPVKKDGPGKAIIVILSALIGGMVACGGVLL 359
                : :  :  :         .  ::*  .*  *:::*.*  *: :::  ::**::.* .*

WZZBST          RNALRSYKPKAL------ 327
WZZSF           RNALRNYNAK-------- 325
WZZE            RRCSK------------- 349
FEPE            RYAMASRKQDAMMADHLV 377
                * .
```

structural alignment of $WzzB_{ST}$, Wzz_{SF}, WzzE and FepE. The last uninterrupted amino acid of each Wzz_i mutant was located on the monomeric structures of FepE (PDB number 3b8n), WzzE (PDB number 3b8o) and $WzzB_{ST}$ (PDB number 3b8p) (Figure 3.8) and also on the respective oligomeric structures (Figures 3.9, 3.10, 3.11 and 3.12). Analysis was primarily conducted on $WzzB_{ST}$, as it exhibits the highest level of amino acid sequence identity to Wzz_{SF} (Morona *et al.*, 2000).

3.8.1 LOCATION OF WZZ_i INSERTIONS MAPPED ONTO MONOMERIC PCP STRUCTURES

The six knock out Class I mutants (i66, i161, i163, i199, i279 and i290) were able to be mapped onto the monomeric structures of $WzzB_{ST}$, WzzE and FepE (Figure 3.8, A, B and C). Mutant i66 is predicted to be located on the very first turn of the first α helix, α1. Mutants i161 and i163 were predicted to be located approximately one third of the way into the long α helix extending from the α/β base domain (α6), while i199 was predicted to be located towards the uppermost region of α6. Mutations i279 and i290 were predicted to be located within the fourth β sheet of the α/β base domain (β4) and the loop closest to TM2 of the determined 3D structure, respectively.

Only three of the five Class II mutants, conferring very short type LPS Oag modal chain length (i191, i247 and i255) could be mapped onto the 3D structure, as i219 and i231 were in regions where structural data is unavailable for all three PCP proteins. The insertion mutation in i191 was predicted to be located on the uppermost quarter of α6 (Figure 3.8, A, B and C). The insertion mutation in i247 was predicted to be located on the loop between α7 and α8, and the insertion in mutant i255 was mapped onto the second turn in α8.

The two Class III mutants i92 and i138, which confer an Oag modal length slightly shorter than wild-type, contain insertions predicted to be located in the second α-helix (α2) and at the

base of the monomer, respectively (Figure 3.8, A, B and C). The insertion of mutant i92 was predicted to be mapped onto the uppermost region of α2; on the second turn for $WzzB_{ST}$, and on the first turn in FepE. This region of the structure is absent in WzzE. The insertion of mutant i138 was predicted to be mapped either on the very top of β3 ($WzzB_{ST}$ and WzzE) or on the loop between β3 and α6 (FepE), at the α/β base domain (Figure 3.8, A, B and C).

The Class IV insertion mutants i80 and i81, conferring wild-type LPS Oag modal length, have insertions that are predicted to be located within the last turn of α1 (Figure 3.8, A, B and C).

Class V insertions mutations i128 and i131 conferring longer LPS Oag modal lengths were mapped on the loop between two β strands, β2 and β3, towards the base of the structure (Figure 3.8, A, B and C).With the exception of Class I insertion mutations (which were predicted to be located throughout Wzz_{SF}), there appeared to be a correlation between location of the insertions and the observed phenotypes of the mutants in Classes II, III, IV and V.

3.8.2 LOCATION OF INSERTIONS MAPPED TO OLIGOMERIC STRUCTURES

In the monomeric structure, Class I mutant insertions are predicted to be located throughout Wzz_{SF}. Figure 3.9 illustrates the locations of the Class I mutants in $WzzB_{ST}$ (Figure 3.9A and B), WzzE (Figure 3.9C) and FepE (Figure 3.9D). Class I mutant i66 is predicted to be on the periphery of the monomer, with close proximity to α2 and the long extended α6 helix of neighbouring monomers (Figure 3.9B). Class I mutants i161 and i163, with insertions predicted to be located on the bottom half of α6, appear to be embedded within the monomer, close to α2 (Figure 3.9B). Class I mutant i199, was predicted to have the insertion sequence located at the uppermost region of α6, isolated from other structural features (Figure 3.9A and B). Mutant i279 (Class I) was predicted to contain the insertion in the central region of the β4 sheet, on the periphery of the monomeric structure, very close to

Figure 3.8 Wzz_i insertions mapped to 3D structures of $WzzB_{ST}$, FepE and WzzE monomers

The location of the Wzz_i insertions mapped to the 3D structures of $WzzB_{ST}$ (PDB 3b8p), WzzE (PDB 3b8o) and FepE (PDB 3b8n). The last uninterrupted amino acid of Wzz_{SF} was mapped onto each monomeric structure. A) $WzzB_{ST}$ monomer, B) WzzE monomer, C) FepE monomer. The mutant classes are illustrated in different colours; Class I mutants are orange, Class II mutants are yellow, Class III are green, Class IV are magenta and Class V are blue. The α helices and β sheets are labelled where appropriate. Images were generated using Pymol ™ software (DeLano Scientific LLC 2008).

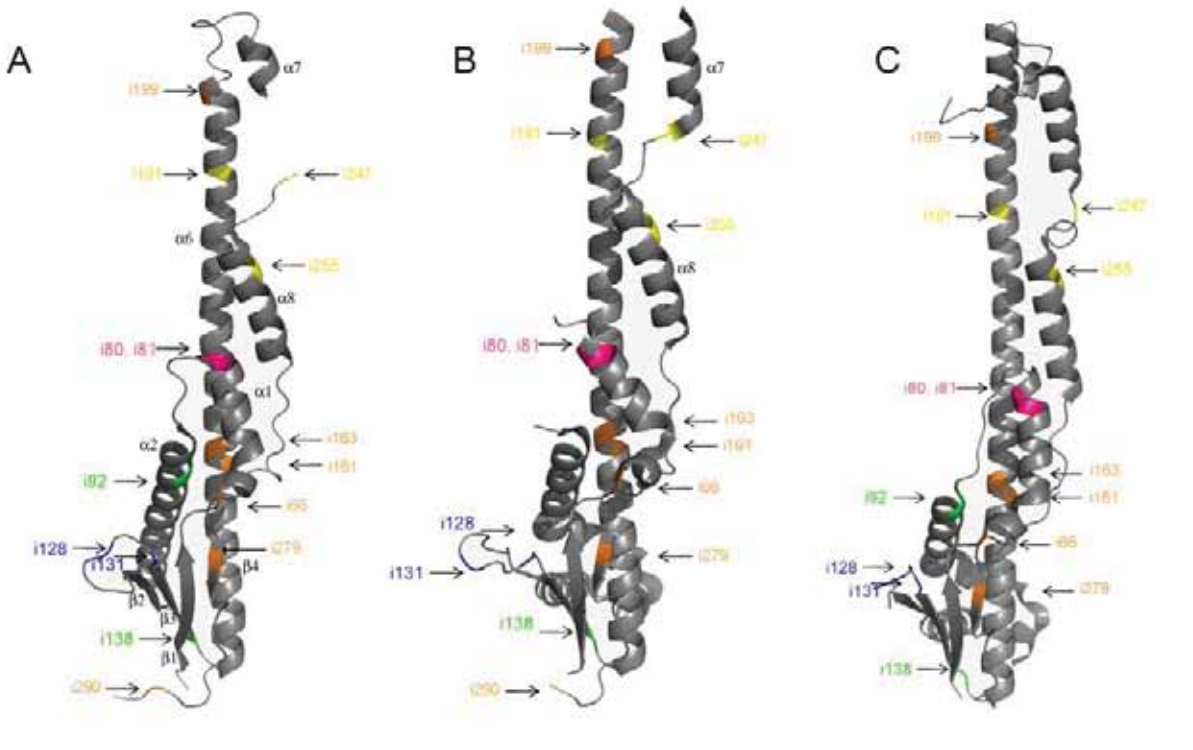
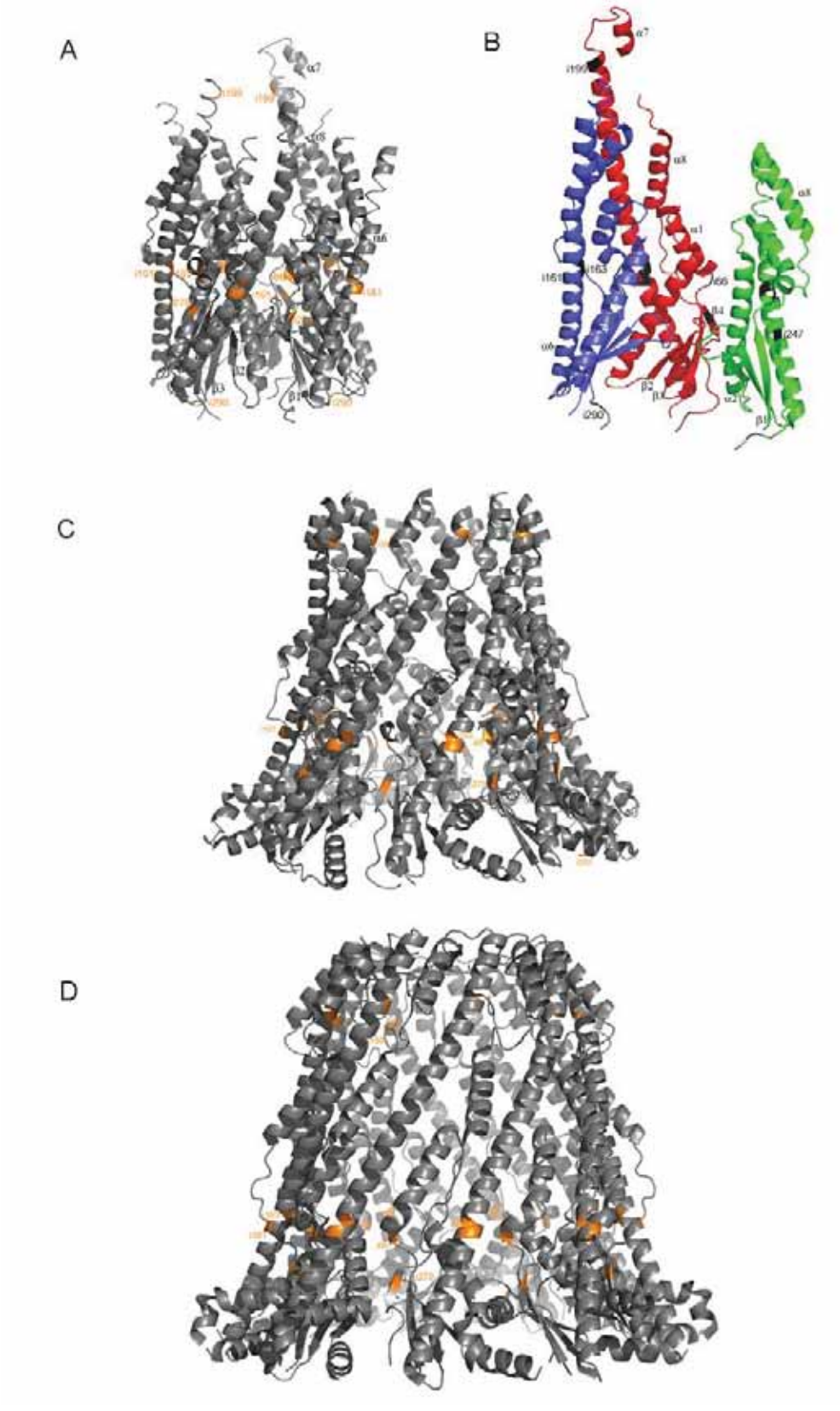

A
α7
i199
i191
i247
α6
i255
α8
i80, i81
α1
i163
α2
i161
i92
i95
i128
i279
i131
β4
β2
β3
i138
β1
i290
B
i199
α7
i191
i247
i255
α8
i80, i81
i163
i191
i95
i128
i279
i131
i138
i290
C
i199
i247
i191
i255
i80, i81
i163
i92
i161
i95
i128
i279
i131
i138

Figure 3.9 Wzz_i Class I insertions mapped to 3D oligomeric structures of $WzzB_{ST}$, FepE and WzzE

The location of the Wzz_i insertions mapped to the 3D structures of Wzz_{ST} (PDB 3b8p), WzzE (PDB 3b8o) and FepE (PDB 3b8n). The last uninterrupted amino acid of Wzz_{SF} was mapped to each oligomeric structure. A) $WzzB_{ST}$ pentamer, B) three coloured monomers of $WzzB_{ST}$ (for greater clarity), C) WzzE octamer, and D) FepE nonamer. The Class I mutants are illustrated and labelled in orange in A, C and D, and illustrated in black for panel B. The α helices and β strands are indicated also. Images were generated using Pymol ™ software (DeLano Scientific LLC 2008).


A
B
C
D

the neighbouring α2 (Figure 3.9B). Mutant i290 (Class I) is predicted to be located on the lowest points in the oligomer and would most likely have close proximity to the transmembrane regions (Figure 3.9A, 3.9B and 3.9C). The insertions from this class are mapped to both internal regions of the oligomers, i.e., embedded within helices and on the lower regions of the oligomer, or also on the exterior surface of the oligomeric structures.

Class II mutation locations were also mapped onto the oligomeric structures; Figure 3.10 illustrates the predicted locations of these insertions in $WzzB_{ST}$ (Figure 3.10A and B), WzzE (Figure 3.10C and D) and FepE (Figure 3.10E). The data indicate that i191, i247 and i255 are all located to the upper half of the oligomer (Figure 3.10A, B, C and D). Mutation i191, predicted to be mapped to the top of α6, appeared to be quite remote from other helices or structural components of Wzz (Figure 3.10A and B), and i247 was predicted to be positioned on the cusp of α7 and α8, close to α6 of neighbouring monomers (Figure 3.10D). Mutation i255, located on the second turn of α8, appeared to be situated closer to the inner surface of the cavity (Figure 3.10A and B). These insertions are all mapped to the surface of the oligomeric structures.

When the Class III insertions were mapped onto the oligomeric structures (Figure 3.11), i92 which is positioned within α2, appears to be situated close to the lining of the inner cavity (Figure 3.11A, B, C and D). Mutation i138 was predicted to be situated at the base of the oligomer, embedded between α6 and β3, close to the membrane surface (Figure 3.11A, B, C, D and E). These insertions appear to be primarily located to internal regions of the oligomeric structures (Table 3.3). The Class IV mutants i80 and i81 were predicted to be embedded internally between α2 and α8 (Figure 3.11B).

Interestingly, the two Class V mutants, i128 and i131, when mapped to the oligomeric structures all PCP proteins revealed that they were located to the residues in the inner cavity of the oligomers (Figure 3.12A, B, C, D and E).

Figure 3.10 Wzz_i Class II insertions mapped to 3D structures of $WzzB_{ST}$, FepE and WzzE

The location of the Wzz_i Class II insertions mapped on the 3D structures of Wzz_{ST} (PDB 3b8p), WzzE (PDB 3b8o) and FepE (PDB 3b8n). The last uninterrupted amino acid of Wzz_{SF} was mapped on each oligomeric structure. A) $WzzB_{ST}$ pentamer, B) three coloured monomers of $WzzB_{ST}$ (for greater clarity), C) WzzE octamer, and D) FepE nonamer. The Class II mutants are illustrated and labelled in yellow (A, C and D) and illustrated in black in B. The α helices and β strands are indicated. Images were generated using Pymol ™ software (DeLano Scientific LLC 2008).

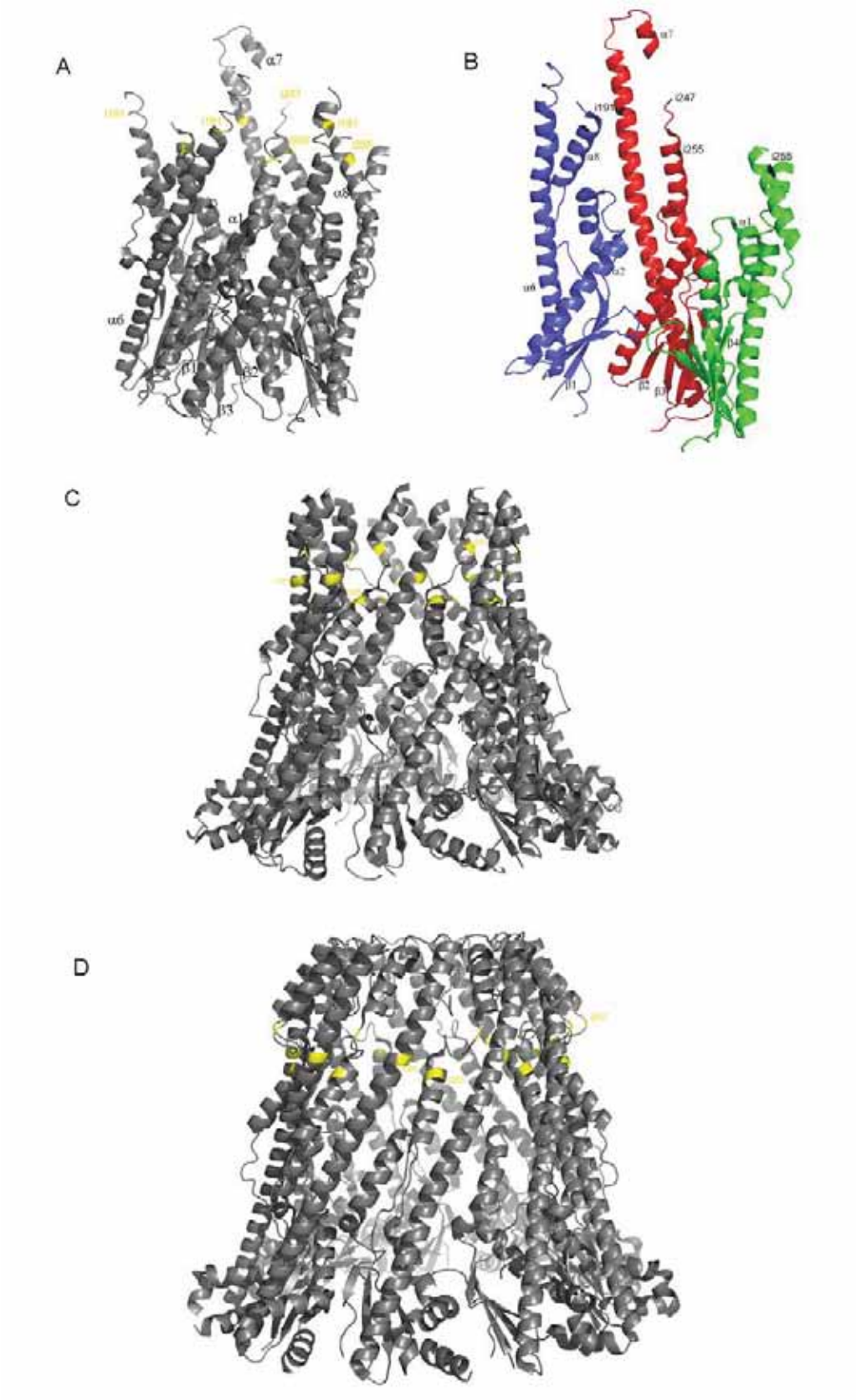

A
α7
α8
α1
α6
β1
β2
β3
B
α7
α8
α6
α2
β1
β2
β3
β4
α1
C
D

Figure 3.11 Wzz_i Class III and Class IV insertions mapped to 3D structures of $WzzB_{ST}$, FepE and WzzE

The location of the Wzz_i Class III and Class IV insertions mapped on the 3D structures of Wzz_{ST} (PDB 3b8p), WzzE (PDB 3b8o) and FepE (PDB 3b8n). The last uninterrupted amino acid of Wzz_{SF} was mapped on each oligomeric structure. A) $WzzB_{ST}$ pentamer, B) three coloured monomers of $WzzB_{ST}$ (for greater clarity), C) a top view of $WzzB_{ST}$, D) WzzE octamer, and E) FepE nonamer. The Class III and IV mutants are illustrated and labelled in green and magenta respectively (A, C, D and E), and illustrated in black in B. The α helices and β strands are indicated. Images were generated using Pymol ™ software (DeLano Scientific LLC 2008).

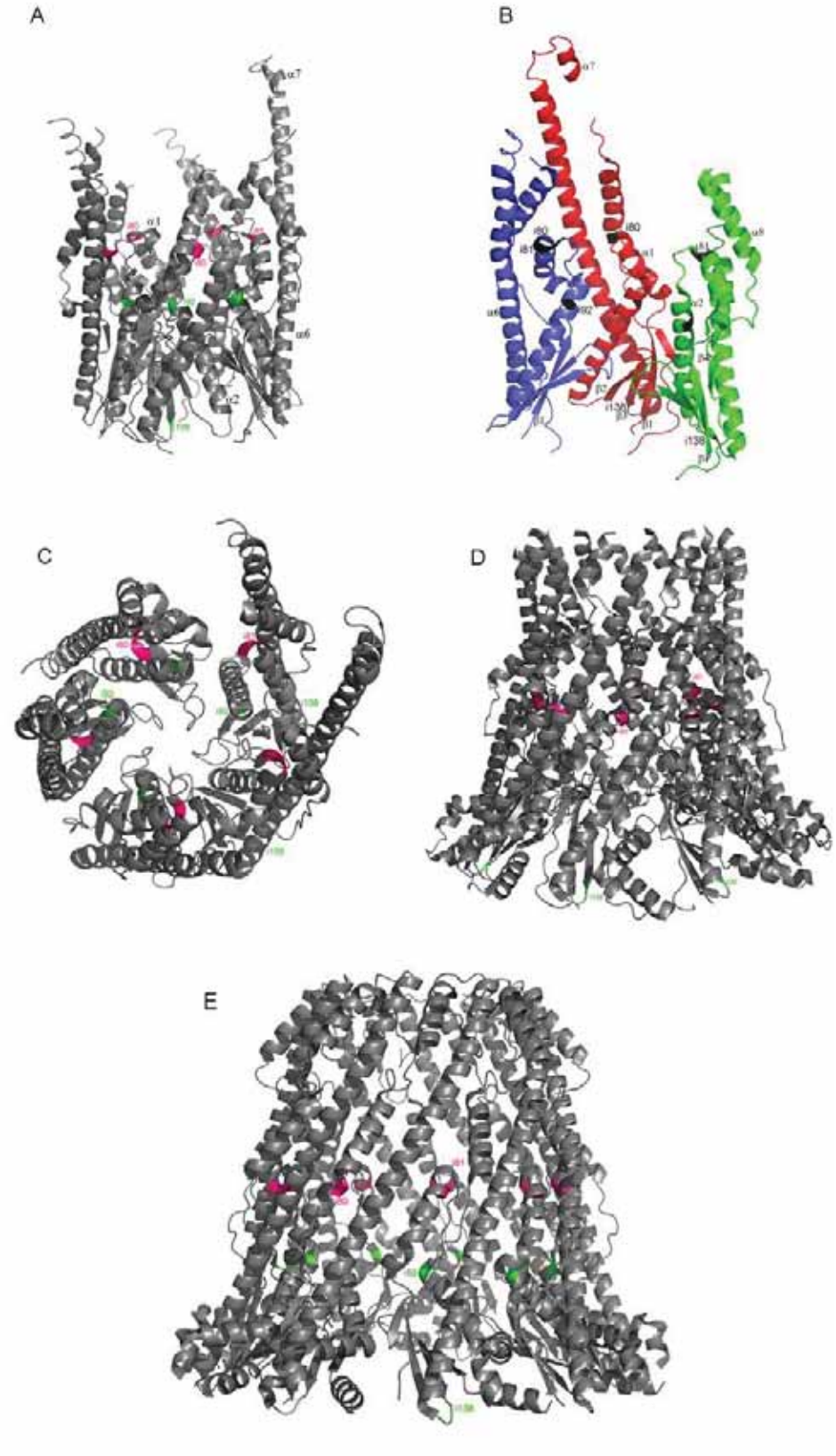

A
B
C
D
E

Table 3.3 The locations of each Wzz_i insertion mapped to structures $WzzB_{ST}$, WzzE and FepE

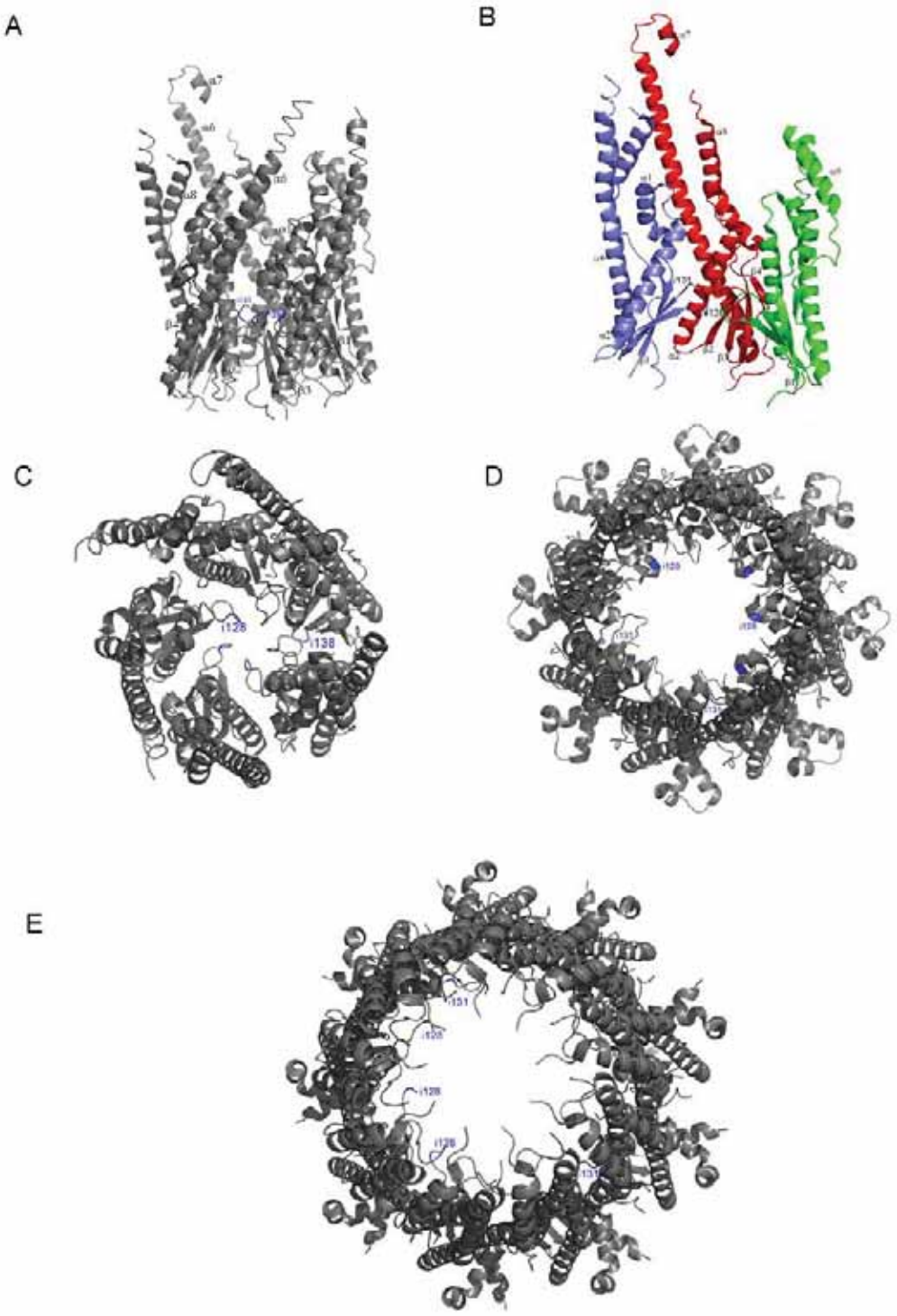
Wzz_i mutants (SF)(ST)(WzzE) (FepE)	Monomeric location (ST)	Monomeric location (WzzE)	Monomeric location (FepE)
D66 D66 T66 E75	First turn of α1	First turn of α1	First turn of α1
I80 L80 L80 L89	Last turn of α1	Last turn of α1	Last turn of α1
Y81 Y81 R81 R90	Last turn of α1	Last turn of α1	Last turn of α1
Q92 Q92 N/A R100	Second turn of α2	N/A	First turn of α2
Q128 Q128 R156 L168	Loop between β2 and β3	Loop between β2 and β3	Loop between β2 and β3
P131 P131 N159 S171	Loop between β2 and β3	Loop between β2 and β3	Loop between β2 and β3
G138 S138 A166 A178	Top of β3	Top of β3	Loop between β3 and α6
Q161 K161 S189 K201	A third into α6	A third into α6	A third into α6

L163 L163 L191 S203	A third into α6	A third into α6	A third into α6
E191 E191 A219 N231	Uppermost α6	Uppermost α6	Uppermost α6
Q199 Q199 S227 R239	Upper region of α6	Upper region of α6	Upper region of α6
Q219 N/A N/A N/A	N/A	N/A	N/A
L231 N/A N/A N/A	N/A	N/A	N/A
P247 247 G276 T289	Loop between α7 and α8	Last turn of α7	Loop between α7 and α8
Y255 Y255 D284 R299	Second turn of α8	Second turn of α8	Second turn of α8
Y279 Y279 Y308 Y323	β4	β4	β4
D290 D290 D319 N/A	Loop closest to TM2	Loop closest to TM2	N/A

Insertions were mapped to 3D structures $WzzB_{ST}$ (PDB number 3b8p), WzzE (PDB number 3b8o) and FepE (PDB number 3b8n). Analysis of structures and predicted location of insertions was performed using Pymol ™ software (DeLano Scientific LLC 2008). Wzz_i mutants in *S. flexneri* Wzz_{SF} are listed in red, and respective mutants in *S. typhimurium* $WzzB_{ST}$ are listed in blue, whilst the corresponding mutants mapped in *E. coli* WzzE are listed in green, and magenta for *E. coli* FepE.

Figure 3.12 Wzz_i Class V insertions mapped to 3D structures of $WzzB_{ST}$, FepE and WzzE

The location of the Wzz_i Class V insertions mapped on the 3D structures of Wzz_{ST} (PDB 3b8p), WzzE (PDB 3b8o) and FepE (PDB 3b8n). The last uninterrupted amino acid of Wzz_{SF} was mapped to each oligomeric structure. A) $WzzB_{ST}$ pentamer, B) three coloured monomers of $WzzB_{ST}$ (for greater clarity), C) a top view (looking from the periplasm towards the inner membrane) of the $WzzB_{ST}$ pentamer, D) a top view of WzzE octamer, and E) a top view of the FepE nonamer. The Class V mutants are illustrated and labelled in blue (A, C, D and E), and illustrated in black in B. The α helices and β strands are indicated. Images were generated using Pymol ™ software (DeLano Scientific LLC 2008).


A
B
C
D
E

3.9 SUMMARY

In this chapter, a collection of in-frame linker insertion mutants of Wzz_{SF} were generated and characterised according to the differences observed in their phenotypic effects on the LPS Oag modal chain. Five different classes of mutants were characterised; Class I mutants did not restore non-random LPS Oag modal chain length and exhibited no Wzz function, Class II mutants conferred very short LPS Oag modal chain length to 2-10 RUs, Class III mutants conferred shorter LPS Oag modal chain length to 8-14 RUs, whilst Class IV mutants resulted in near wild-type LPS Oag modal chain length (11-19 RUs) and Class V mutants conferred longer LPS Oag modal chain length (16-25 RUs). Cross-linking analyses indicated that higher order oligomers were easily detected in Class V mutants and wild-type Wzz_{SF}, and that stable dimers were also easily detected in Class V mutants and wild-type, but not detected in representative mutants from Classes I, II and III. The mutants were assessed on their sensitivity to colicin E2, and showed that there is a strong correlation of LPS Oag modal chain length and susceptibility to colicin, as Classes IV and V had greater resistance to the lethal action of colicin E2. Assessing the location of the mutations in the 3D structures of $WzzB_{ST}$, WzzE and FepE, illustrated that Class V mutants are predicted to have insertions located to residues in the inner oligomeric cavity.

CHAPTER FOUR

WZZ:WZZ PROTEIN INTERACTION STUDY

4.1 INTRODUCTION

In a previous study, wzz_{SF} located on pET-17b was mutated such that a variety of substitution mutations were generated (Daniels, 1999) (Table 2.2). The resulting mutant proteins were: Wzz_{K31A}, encoded by pRMCD119, Wzz_{K267N}, encoded by pRMCD108, Wzz_{P286A} (pRMCD117), and Wzz_{P292A} (pRMCD116) (Figure 4.1) (Table 2.2). Plasmids incorporating double substitutions in Wzz_{SF} were also created; pRMCD113 had substitutions at glycine residues (G305A and G311A within TM2), encoding protein $Wzz_{G305A/G311A}$ and pRMCD122 had the substitutions M32T and I35C encoding protein $Wzz_{M32T/I35C}$ (Table 2.2). The sequences from these altered Wzz_{SF} proteins are aligned with Wzz_{SF} and Wzz_{ST} in Figure 4.1. These studies also explored the LPS Oag modal chain length conferred by *S. typhimurium* Wzz_{ST} (pRMCD80), and Wzz_{ST} and Wzz_{SF} hybrid proteins (Daniels and Morona, 1999). *S. typhimurium* Wzz_{ST} was incorporated in those experiments as *S. flexneri* Wzz_{SF} and *S. typhimurium* Wzz_{ST} share 77% amino acid sequence identity (Morona *et al.*, 2000), yet exhibits very dissimilar LPS Oag modal chain lengths (Wzz_{SF} results in modal chain length of 11-17 RUs, whilst Wzz_{ST} confers 19-30 RUs). The hybrid Wzz proteins were constructed by utilizing a unique *Bgl*II site present in the open reading frames of the two genes in order to assemble the hybrids $Wzz_{ST/SF}$, which contained either the first half of Wzz_{ST} and the second of Wzz_{SF} (pRMCD104), or vice versa to produce the hybrid $Wzz_{SF/ST}$ (encoded by pRMCD106) (Figure 4.2) (Daniels and Morona, 1999).

These studies demonstrated that Wzz_{K31A} (pRMCD119) resulted in a loss of activity, as judged by the inability to complement the *wzz* defect in the strain (Daniels and Morona, 1999). Expression of Wzz_{K267N} (pRMCD108) resulted in an increase in LPS Oag modal chain length, from wild-type 11-17 RUs to the intermediate type 13-20 RUs (Daniels and Morona,

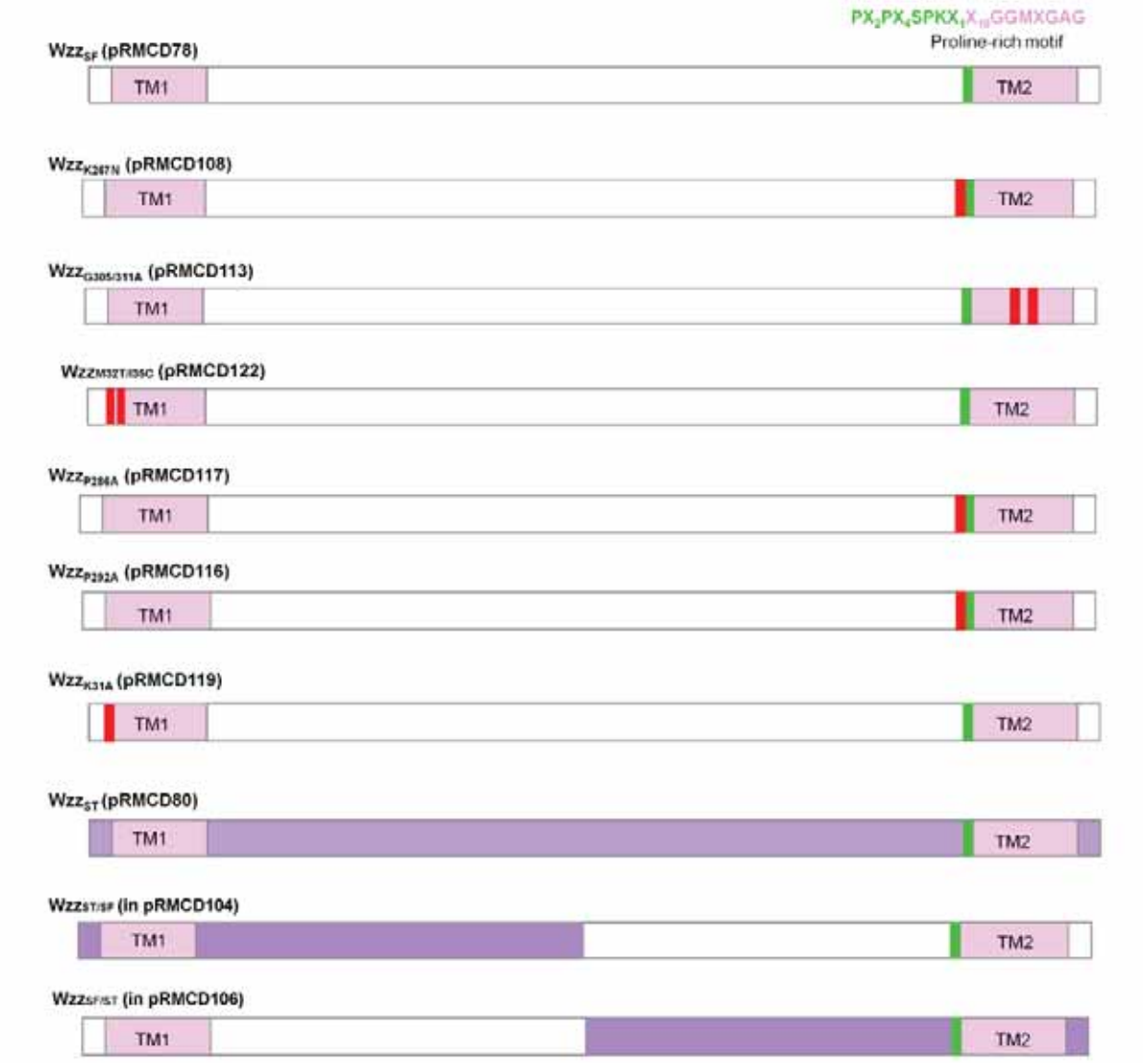
Figure 4.1 Wzz protein alignments

The multiple sequence alignment of Wzz_{SF} (genbank accession number CAA50783.1/Swissprot P37792) proteins $Wzz_{G305A/G311A}$, Wzz_{P292A}, Wzz_{P286A}, Wzz_{K31A}, Wzz_{K267N} and *S. typhimurium* Wzz (Wzz_{ST}, genbank accession number NP_461024.1/Swisprot Q04866) using Clustal 2.0.12. Red residues indicate small hydrophobic amino acids, blue resides indicate acidic amino acids, magenta residues illustrate basic amino acids, and green residues indicate hydroxyl/basic/amine amino acids. '*' indicates positions which have a single, fully conserved residue. "." indicates where semi-conserved substitutions are observed, and ":"indicates where conserved substitutions have been observed, according to the colour specifications mentioned above (http://www.ebi.ac.uk/Tools/clustalw2/index.html). Numbers on the right represent the aa position. TM1 and TM2 regions are underlined, and amino acid substitutions are illustrated in black.

```
WZZG305A/G311A      MRVENNNVSGQNHDPEQIDLIDLLVQLWRGKMTIIISVIVAIALAIGYLAVAKEKWTSTA 60
WZZP292A            MRVENNNVSGQNHDPEQIDLIDLLVQLWRGKMTIIISVIVAIALAIGYLAVAKEKWTSTA 60
WZZP286A            MRVENNNVSGQNHDPEQIDLIDLLVQLWRGKMTIIISVIVAIALAIGYLAVAKEKWTSTA 60
WZZK31A             MRVENNNVSGQNHDPEQIDLIDLLVQLWRGAMTIIISVIVAIALAIGYLAVAKEKWTSTA 60
WZZSF               MRVENNNVSGQNHDPEQIDLIDLLVQLWRGKMTIIISVIVAIALAIGYLAVAKEKWTSTA 60
WZZK267N            MRVENNNVSGQNHDPEQIDLIDLLVQLWRGKMTIIISVIVAIALAIGYLAVAKEKWTSTA 60
WZZST               MTVDSNTSSGRGNDPEQIDLIELLLQLWRGKMTIIVAVIIAILLAVGYLMIAKEKWTSTA 60
                    * *:.*. **:.:********:**:***** ****::**:** **:*** :*********

WZZG305A/G311A      IITQPDVGQIAGYNNAMNVIYGQAAPKVSDLQETLIGRFSSAFSALAETLDNQEEPEKLT 120
WZZP292A            IITQPDVGQIAGYNNAMNVIYGQAAPKVSDLQETLIGRFSSAFSALAETLDNQEEPEKLT 120
WZZP286A            IITQPDVGQIAGYNNAMNVIYGQAAPKVSDLQETLIGRFSSAFSALAETLDNQEEPEKLT 120
WZZK31A             IITQPDVGQIAGYNNAMNVIYGQAAPKVSDLQETLIGRFSSAFSALAETLDNQEEPEKLT 120
WZZSF               IITQPDVGQIAGYNNAMNVIYGQAAPKVSDLQETLIGRFSSAFSALAETLDNQEEPEKLT 120
WZZK267N            IITQPDVGQIAGYNNAMNVIYGQAAPKVSDLQETLIGRFSSAFSALAETLDNQEEPEKLT 120
WZZST               IITQPDAAQVATYTNALNVLYGGNAPKISEVQANFISRFSSAFSALSEVLDNQKEREKLT 120
                    ******..*:* *.**:**:**  ***:*::* .:*.*********:*.****:* ****

WZZG305A/G311A      IEPSVKNQQLPLTVSYVGQTAEGAQMKLAQYIQQVDDKVNQELEKDLKDNIALGRKNLQD 180
WZZP292A            IEPSVKNQQLPLTVSYVGQTAEGAQMKLAQYIQQVDDKVNQELEKDLKDNIALGRKNLQD 180
WZZP286A            IEPSVKNQQLPLTVSYVGQTAEGAQMKLAQYIQQVDDKVNQELEKDLKDNIALGRKNLQD 180
WZZK31A             IEPSVKNQQLPLTVSYVGQTAEGAQMKLAQYIQQVDDKVNQELEKDLKDNIALGRKNLQD 180
WZZSF               IEPSVKNQQLPLTVSYVGQTAEGAQMKLAQYIQQVDDKVNQELEKDLKDNIALGRKNLQD 180
WZZK267N            IEPSVKNQQLPLTVSYVGQTAEGAQMKLAQYIQQVDDKVNQELEKDLKDNIALGRKNLQD 180
WZZST               IEQSVKGQALPLSVSYVSTTAEGAQRRLAEYIQQVDEEVAKELEVDLKDNITLQTKTLQE 180
                    ** ***.* ***:****. ****** :**:******::* :*** ******:*  *.**:

WZZG305A/G311A      SLRTQEVVAQEQKDLRIRQIQEALQYANQAQVTKPQVQQTEDVTQDTLFLLGSEALESMI 240
WZZP292A            SLRTQEVVAQEQKDLRIRQIQEALQYANQAQVTKPQVQQTEDVTQDTLFLLGSEALESMI 240
WZZP286A            SLRTQEVVAQEQKDLRIRQIQEALQYANQAQVTKPQVQQTEDVTQDTLFLLGSEALESMI 240
WZZK31A             SLRTQEVVAQEQKDLRIRQIQEALQYANQAQVTKPQVQQTEDVTQDTLFLLGSEALESMI 240
WZZSF               SLRTQEVVAQEQKDLRIRQIQEALQYANQAQVTKPQVQQTEDVTQDTLFLLGSEALESMI 240
WZZK267N            SLRTQEVVAQEQKDLRIRQIQEALQYANQAQVTKPQVQQTEDVTQDTLFLLGSEALESMI 240
WZZST               SLETQEVVAQEQKDLRIKQIEEALRYADEAKITQPQIQQTQDVTQDTMFLLGSDALKSMI 240
                    **.**************:**:***:**::*::*:**:***:******:*****:**:***

WZZG305A/G311A      KHEATRPLVFSPNYYQTRQNLLDIEKLKFDDLDIHAYRYVMKPTLPIRRDSPKKAITLIL 300
WZZP292A            KHEATRPLVFSPNYYQTRQNLLDIEKLKFDDLDIHAYRYVMKPTLPIRRDSAKKAITLIL 300
WZZP286A            KHEATRPLVFSPNYYQTRQNLLDIEKLKFDDLDIHAYRYVMKPTLAIRRDSPKKAITLIL 300
WZZK31A             KHEATRPLVFSPNYYQTRQNLLDIEKLKFDDLDIHAYRYVMKPTLPIRRDSPKKAITLIL 300
WZZSF               KHEATRPLVFSPNYYQTRQNLLDIEKLKFDDLDIHAYRYVMKPTLPIRRDSPKKAITLIL 300
WZZK267N            KHEATRPLVFSPNYYQTRQNLLDIENLKFDDLDIHAYRYVMKPTLPIRRDSPKKAITLIL 300
WZZST               QNEATRPLVFSPAYYQTKQTLLDIKNLKVTADTVHVYRYVMKPTLPVRRDSPKTAITLVL 300
                    ::********** ****:*.****::**.    :*.*********.:****.*.****:*

WZZG305A/G311A      AVLLAGMVGAAIVLGRNALRNYNAK-- 325
WZZP292A            AVLLGGMVGAGIVLGRNALRNYNAK-- 325
WZZP286A            AVLLGGMVGAGIVLGRNALRNYNAK-- 325
WZZK31A             AVLLGGMVGAGIVLGRNALRNYNAK-- 325
WZZSF               AVLLGGMVGAGIVLGRNALRNYNAK-- 325
WZZK267N            AVLLGGMVGAGIVLGRNALRNYNAK-- 325
WZZST               AVLLGGMIGAGIVLGRNALRSYKPKAL 327
                    ****.**:**.*********.*:.*
```

Figure 4.2 pET-17b-based Wzz_{ST} and Wzz_{SF} constructs used in co-expression assays

A schematic diagram illustrating Wzz_{SF}, (pRMCD78), mutant proteins Wzz_{K267N} (pRMCD108), $Wzz_{G305A/G311A}$ (pRMCD113), $Wzz_{M32T/I35C}$ (pRMCD122), Wzz_{P286A} (pRMCD117), Wzz_{P292A} (pRMCD116), Wzz_{K31A} (pRMCD119), Wzz_{ST} (pRMCD80), and hybrid proteins $Wzz_{ST/SF}$ (pRMCD104) and $Wzz_{SF/ST}$ (pRMCD106). All proteins are 325 amino acids, excluding Wzz_{ST} and $Wzz_{SF/ST}$, which are 327 amino acids. Indicated in this figure are the transmembrane regions (TM1 and TM2) in pink, and the proline-rich motif; in the text above Wzz_{SF}, green indicates the residues of the motif present in the periplasm, whilst pink indicates the residues present in TM2. Approximate locations of the substitution mutations are indicated in red. Purple indicates Wzz_{ST} or the hybrid region of Wzz_{ST}.


PX2PX4SPKX1X10GGMXGAG
Proline-rich motif
Wzz$_{SF}$ (pRMCD78)
TM1
TM2
Wzz$_{K267N}$ (pRMCD108)
TM1
TM2
Wzz$_{G305/311A}$ (pRMCD113)
TM1
Wzz$_{M32T/I95C}$ (pRMCD122)
TM1
TM2
Wzz$_{P286A}$ (pRMCD117)
TM1
TM2
Wzz$_{P292A}$ (pRMCD116)
TM1
TM2
Wzz$_{K31A}$ (pRMCD119)
TM1
TM2
Wzz$_{ST}$ (pRMCD80)
TM1
TM2
Wzz$_{ST/SF}$ (in pRMCD104)
TM1
TM2
Wzz$_{SF/ST}$ (in pRMCD106)
TM1
TM2

1999), while Wzz_{P286A} (pRMCD117) had a reduced level of wild-type Oag modal chain length conferred, and Wzz_{P292A} (pRMCD116), resulted in non-regulated LPS Oag modal chain length, indicating a loss of activity (Daniels and Morona, 1999).

The resulting LPS Oag modal chain length conferred by pRMCD113 ($Wzz_{G305A/G311A}$) reduced the wild-type 11-17 RUs to 3-8 RUs (Daniels and Morona, 1999). A similar phenotype was observed with $Wzz_{M32T/I35C}$ (pRMCD122). When assessing the LPS phenotype conferred by pRMCD80 (Wzz_{ST}), the resulting LPS Oag profiles indicated that even though expressed in *S. flexneri*, pRMCD80 conferred an *S. typhimurium*-like LPS Oag modal chain length of L-type 19-30 RUs (Daniels and Morona, 1999). Analysis of the LPS profile of the two Wzz hybrids pRMCD104 ($Wzz_{ST/SF}$) and pRMCD106 ($Wzz_{SF/ST}$) demonstrated that $Wzz_{ST/SF}$ conferred an Oag modal chain length of 14-19 RUs, similar to wild-type Wzz_{SF} modal chain length, and $Wzz_{SF/ST}$ resulted in an LPS Oag chain length of 17-26 RUs, closer to the L-type Oag modal chain length conferred by Wzz_{ST} (Daniels and Morona, 1999).

These previous studies indicated that the Wzz TM regions play a critical role in establishing and influencing the resulting LPS Oag modal chain length (Daniels and Morona, 1999). The aim of the work in this chapter was to investigate the basis for Oag modal length determination by investigating potential protein interactions occurring between Wzz proteins utilising both genetic and biochemical approaches. Based on the observations of Daniels and Morona (1999), it was hypothesised that wild-type Wzz_{SF} co-expression with mutant Wzz proteins, would prove indicative of whether the two Wzz proteins were capable of interacting, and whether this may affect the resulting function as judged by the resulting LPS Oag modal chain length profiles.

CHAPTER FOUR

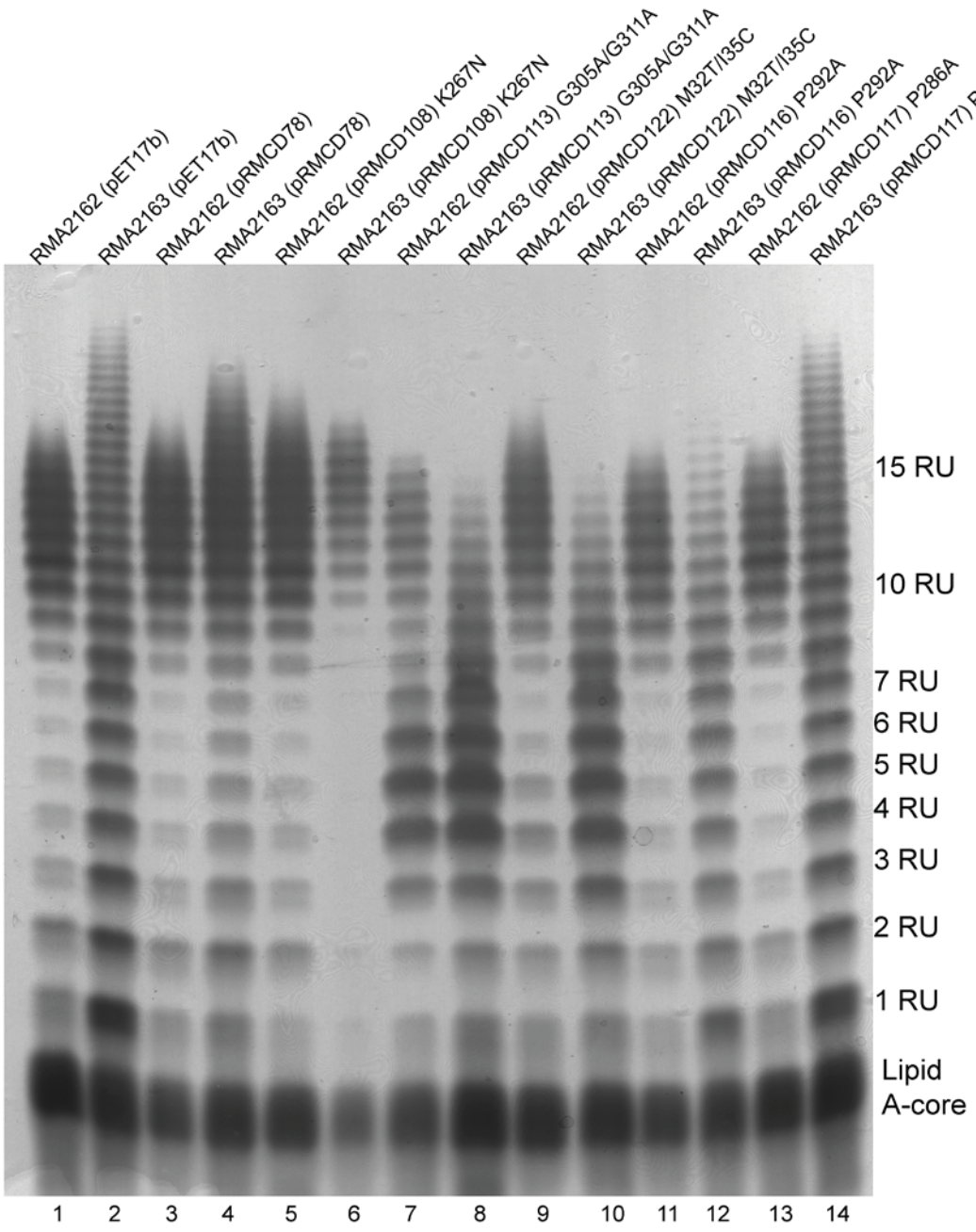
4.2 THE ASSESSMENT OF MUTANT WZZ GENETIC DOMINANCE

4.2.1 MUTATIONS K267N, M32T/I35C AND MUTATIONS IN TM2

As Wzz oligomerisation has previously been established (Daniels and Morona, 1999), the functional significance of this interaction was explored by co-expressing wild-type and mutant forms of Wzz. Plasmids encoding Wzz_{SF}, Wzz_{K267N}, $Wzz_{G305A/G311A}$, $Wzz_{M32T/I35C}$, Wzz_{P286A} and Wzz_{P292A} (Figure 4.1) were transformed into two strains; RMA2163, a *S. flexneri* serotype Y SFL1 *wzz*::Km^R strain, cured of both the large virulence plasmid and pHS-2 plasmid (encoding Wzz_{pHS-2}), and the corresponding parent strain RMA2162, possessing wild-type wzz_{SF} (Table 2.2, Figure 4.1). LPS samples were prepared from these strains and subjected to SDS-PAGE and silver staining as described in section 2.11. The resulting LPS profiles showed that control strains carrying the pET-17b vector control and pRMCD78 (encoding wzz_{SF}) in the *wzz* deficient strain RMA2163, had random LPS Oag chain length and wild-type Oag modal chain length, respectively, as expected (Figure 4.3, lanes 2 and 4). When expressed in the strain RMA2162, the resulting modal length conferred by both constructs was the same as the wild-type, 11-17 RUs (Figure 4.3, lanes 1 and 3). Plasmid pRMCD108, encoding Wzz_{K267N}, did not have a dramatic effect on the resulting LPS Oag modal chain length, as the modal length was only slightly increased to an average length of 12-18 RUs when expressed in RMA2163 (Figure 4.3, lane 6, Table 4.1), similar to the result previously reported (Daniels and Morona, 1999). When expressed in RMA2162, a wild-type Oag modal chain length was observed (Figure 4.3, lane 5). The plasmids pRMCD113 ($Wzz_{G305A/G311A}$) and pRMCD122 ($Wzz_{M32T/I35C}$) in RMA2163 conferred similar very short (VS) Oag modal chain lengths of 3-8 RUs (Figure 4.3, lanes 8 and 10 respectively, Table 4.1). However, when expressed in RMA2162, the two Wzz mutant proteins conferred novel LPS Oag profiles. Strain RMA2162 (pRMCD113) exhibited bimodality, i.e. two distinct and different Oag chain modal length were evident; the wild-type average of 11-17 RUs, and the VS Oag modal length of 3-8 RUs (Figure 4.3, lane 7).

Figure 4.3 Analysis of LPS produced by Wzz_{SF} mutants expressed in *S. flexneri* RMA2162 and RMA2163

Strains harbouring plasmids were grown in LB + Amp for 3 h, and LPS samples were prepared, electrophoresed on a SDS 15% polyacrylamide gel and silver stained as described in section 2.11. Strains in each lane are as follows: 1) RMA2162 (pET-17b), 2) RMA2163 (pET-17b), 3) RMA2162 (pRMCD78), 4) RMA2163 (pRMCD78), 5) RMA2162 (pRMCD108), 6) RMA2163 (pRMCD108), 7) RMA2162 (pRMCD113), 8) RMA2163 (pRMCD113), 9) RMA2162 (pRMCD122), 10) RMA2163 (pRMCD122), 11) RMA2162 (pRMCD116), 12) RMA2163 (pRMCD116), 13) RMA2162 (pRMCD117) and 14) RMA2163 (pRMCD117). The lipid A-core and number of repeat units (RUs) are indicated on the right. Each lane contains approximately 1.3 x 10^8 cells.


RMA2162 (pET17b)
RMA2163 (pET17b)
RMA2162 (pRMCD78)
RMA2163 (pRMCD78)
RMA2162 (pRMCD108) K267N
RMA2163 (pRMCD108) K267N
RMA2162 (pRMCD113) G305A/G311A
RMA2163 (pRMCD113) G305A/G311A
RMA2162 (pRMCD122) M32T/I35C
RMA2163 (pRMCD122) M32T/I35C
RMA2162 (pRMCD116) P292A
RMA2163 (pRMCD116) P292A
RMA2162 (pRMCD117) P286A
RMA2163 (pRMCD117) P286A
15 RU
10 RU
7 RU
6 RU
5 RU
4 RU
3 RU
2 RU
1 RU
Lipid A-core
1 2 3 4 5 6 7 8 9 10 11 12 13 14

However, when pRMCD122 was expressed in RMA2162, only the wild-type Oag modal chain length was observed (Figure 4.3, lane 9). Plasmids pRMCD116 and pRMCD117 encoding Wzz_{P292A} and Wzz_{P286A}, respectively, did not complement the *wzz* mutation in RMA2163 (Figure 4.3, lanes 12 and 14). While the previous study reported that Wzz_{P286A} exhibited a diminished level of wild-type LPS Oag modal chain length control (Daniels and Morona, 1999), this was not observed in the current study. Proteins Wzz_{P292A} and Wzz_{P286A} only exhibited wild-type LPS Oag modal chain length when expressed in RMA2162, thereby illustrating in this study that these mutations that completely knock out Wzz function, were recessive to the wild-type, and did not exhibit negative dominance over Wzz wild-type function (Figure 4.3, lanes 11 and 13, Table 4.1).

Based on the observations made and hypothesizing that Wzz-Wzz interaction is important for determination of modal length, the data suggest that $Wzz_{G505A/G311A}$ is unable to interact with Wzz_{SF}, as two modal lengths indicate two separate functions, and that $Wzz_{M32T/I35C}$ is most likely capable of interacting with wild-type Wzz_{SF} as wild-type LPS Oag modal chain length is observed and hence the $Wzz_{M32T/I35C}$ defect is corrected. Conversely, Wzz_{P286A} and Wzz_{P292A} exhibit no functionality, as random modal chain length is unchanged when either protein is expressed in the *wzz* deficient strain, and no effect was observed on wild-type Wzz_{SF} function. Both proline residues are clearly needed for wild-type Wzz_{SF} function.

4.2.2 COMPLEMENTATION WITH WZZ_{ST} AND WZZ_{ST}/WZZ_{SF} HYBRID PROTEINS, AND WZZ_{K31A}

Plasmids pET-17b, pRMCD78 (Wzz_{SF}), pRMCD80 (Wzz_{ST}), pRMCD119 (Wzz_{K31A}), pRMCD104 ($Wzz_{ST/SF}$) and pRMCD106 ($Wzz_{SF/ST}$) were transformed into the strains RMA2162 and RMA2163. LPS samples were prepared from these strains and samples were analysed by SDS-PAGE electrophoresis and silver staining as described in section 2.11.

Table 4.1

Plasmids	RMA2163 (RUs)	RMA2162 (RUs)
*pRMCD78 (Wzz_{SF})	10-17	10-17
pRMCD108 (K267N)	12-18	10-17
pRMCD113 (G305/G311A)	3-8	3-8 10-17
pRMCD122 (M32T/I35C)	3-8	10-17
pRMCD117 (P286A)	Random	10-17
pRMCD116 (P292A)	Random	10-17
pRMCD119 (K31A)	Random	10-17
pRMCD80 (Wzz_{ST})	17-26	12-20
pRMCD104 ($Wzz_{ST/SF}$)	11-23	10-18
pRMCD106 ($Wzz_{SF/ST}$)	14-25	10-21
[a]pRMCD30 (His_6-Wzz_{SF})	10-17	10-17
pBAD-Wzz_{SF} (FLAG-Wzz_{SF})	10-17	10-17
pBAD-Wzz_{ST} (FLAG-Wzz_{ST})	17-26	12-21
pBAD-$Wzz_{G305A/G311A}$ (FLAG-$Wzz_{G305A/G311A}$)	3-8	3-7 10-17
[b] Wzz_{i191} Class II Wzz_i	3-8	2-7 10-17
Wzz_{i219} Class II Wzz_i	2-10	2-7 10-16
Wzz_{i231} Class II Wzz_i	3-7	3-7 10-16
Wzz_{i247} Class II Wzz_i	2-8	2-6 10-17
Wzz_{i255} Class II Wzz_i	3-8	2-6 10-17
[c] Wzz_{i128} Class V Wzz_i	16-21	12-20
Wzz_{i131} Class V Wzz_i	16-25	12-22

The lengths conferred by the pET-17b, pQE-30 and pBAD33-based plasmids expressed in strains RMA2163 and RMA2162
*based on data from Figure 4.3 and 4.4, [a] based on data from Figure 4.8, [b] based on data from Figure 4.13 and [c] based on data from Figure 4.14

RMA2162 and RMA2163 harbouring pRMCD78 (Wzz_{SF}) conferred wild-type modal length of 11-17 RUs, as expected (Figure 4.4, lane 4, Table 4.1) and wild-type length when expressed in RMA2162 (Figure 4.4, lane 3, Table 4.1). RMA2163 carrying pRMCD80 (Wzz_{ST}) exhibited L-type LPS Oag modal chain length of 17-28 RUs (Figure 4.4, lane 6), as similarly observed in previous studies (Daniels and Morona, 1999). However, when pRMCD80 was expressed in RMA2162, the resulting LPS had an Oag modal chain length of 12-20 RUs, decreased from L-type, and was closer to the wild-type S-type modal length of 11-17 RUs (Figure 4.4, lane 5, Table 4.1). This was as unexpected result, as two Oag modal lengths were expected.

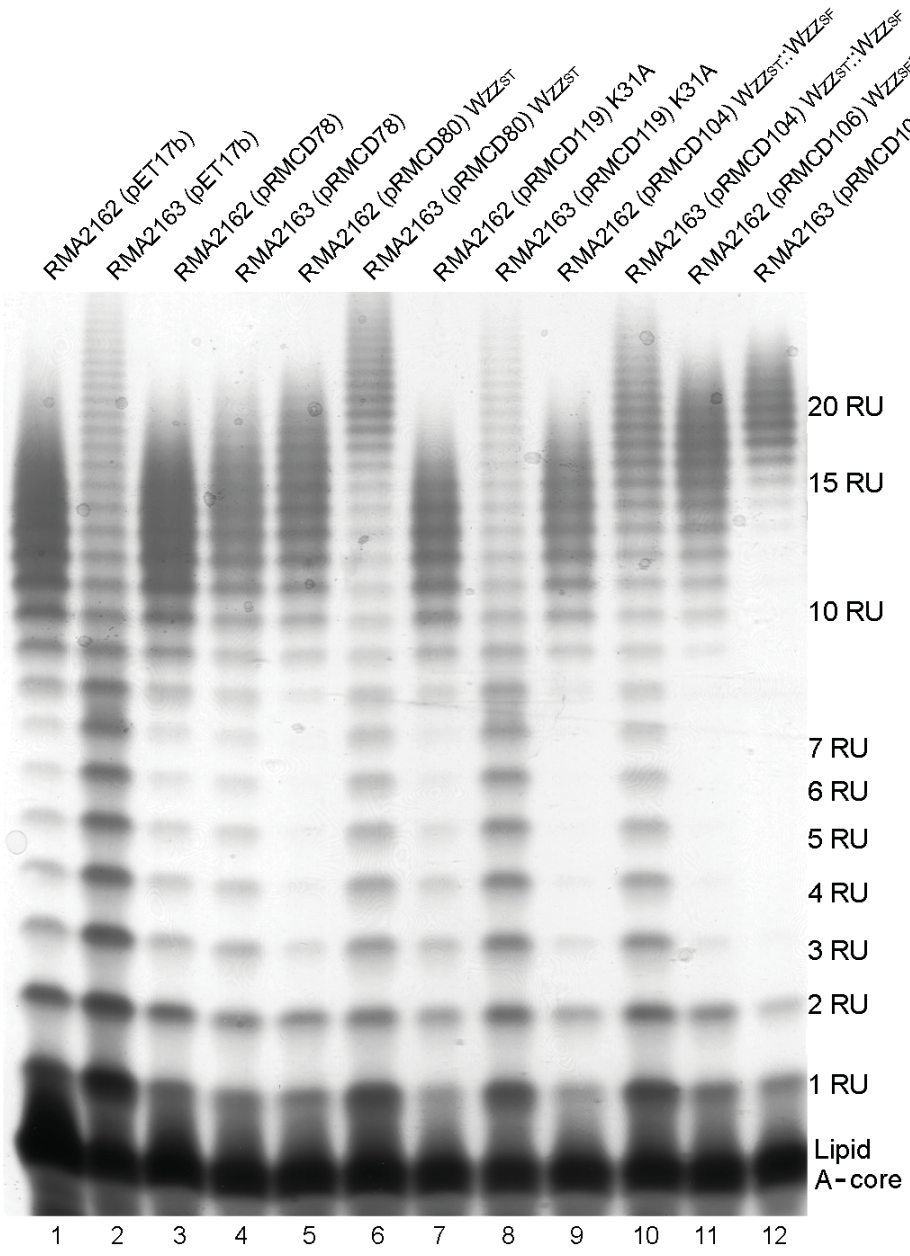
RMA2163 harbouring plasmid pRMCD119 (encoding Wzz_{K31A}), resulted in a complete loss of Wzz function, as detected by the failure of this Wzz protein to complement LPS Oag modal chain length control (Figure 4.4, lane 8), and exhibited a wild-type LPS phenotype in RMA2162 (Figure 4.4, lane 7). This was expected and confirmed previously reported observations (Daniels and Morona, 1999).

RMA2163 carrying pRMCD104 ($Wzz_{ST/SF}$) had an LPS Oag modal chain length of 11-19 RUs (Figure 4.4 lane 10, Table 4.1). Previously reported observations showed that pRMCD104 expressed in a *wzz* deficient strain resulted in LPS Oag modal chain length of 14-19 RUs (Daniels and Morona, 1999). $Wzz_{ST/SF}$ appears to have less ability to regulate LPS Oag modal chain length, as the average distribution is not as distinct as is normally produced when wild-type Wzz imparts modal length control. RMA2162 (pRMCD104), which possesses a functional wild-type Wzz_{SF}, had LPS with Oag modal chain length of 10-18 RUs which is closer to the wild-type modal chain length (Figure 4.4, lane 9, Table 4.1).

RMA2163 harbouring plasmid pRMCD106 (encoding for $Wzz_{SF/ST}$) conferred an Oag modal chain length of 14-25 RUs (Figure 4.4, lane 12, Table 4.1) , consistent with previously reported Oag chain length of 17-26 RUs (Daniels and Morona, 1999). When expressed in RMA2162, pRMCD106 ($Wzz_{SF/ST}$) resulted in Oag modal chain length of 10-21 RUs which is slightly longer than that conferred by pRMCD78 (Wzz_{SF}) (Figure 4.4, lane 3

Figure 4.4 Analysis of LPS of *S. flexneri* strains expressing Wzz_{K31A}, Wzz_{ST} and Wzz_{ST}/Wzz_{SF} hybrid proteins

RMA2162 and RMA2163 strains were grown in LB + Amp for 3 h, and LPS samples were prepared, electrophoresed on a SDS 15% polyacrylamide gel and silver stained as described in section 2.11. Strains in each lane are as follows: 1) RMA2162 (pET-17b), 2) RMA2163 (pET-17b), 3) RMA2162 (pRMCD78), 4) RMA2163 (pRMCD78), 5) RMA2162 (pRMCD80), 6) RMA2163 (pRMCD80), 7) RMA2162 (pRMCD119), 8) RMA2163 (pRMCD119), 9) RMA2162 (pRMCD104), 10) RMA2163 (pRMCD104), 11) RMA2162 (pRMCD106) and 12) RMA2163 (pRMCD106). The lipid A-core and number of repeat units (RUs) are indicated on the right. Each lane contains approximately 1.3×10^8 cells.


RMA2162 (pET17b)
RMA2163 (pET17b)
RMA2162 (pRMCD78)
RMA2163 (pRMCD78)
RMA2162 (pRMCD80) WZZST
RMA2163 (pRMCD80) WZZST
RMA2162 (pRMCD119) K31A
RMA2163 (pRMCD119) K31A
RMA2162 (pRMCD104) WZZST::WZZSF
RMA2163 (pRMCD104) WZZST::WZZSF
RMA2162 (pRMCD106) WZZSF::WZZST
RMA2163 (pRMCD106) WZZSF::WZZST
20 RU
15 RU
10 RU
7 RU
6 RU
5 RU
4 RU
3 RU
2 RU
1 RU
Lipid A-core
1 2 3 4 5 6 7 8 9 10 11 12

and 11, Table 4.1). Whilst previous data and current data exhibit slight differences (see section 7.8), both hybrid proteins result in slightly longer Oag modal chain length than Wzz_{SF} and shorter length than Wzz_{ST}, and as the LPS Oag modal chain length observed when the hybrid proteins are co-expressed with Wzz_{SF} appears to be closer to wild-type length than the hybrid length, Wzz_{SF} appears to be dominant over both $Wzz_{ST/SF}$ and $Wzz_{SF/ST}$.

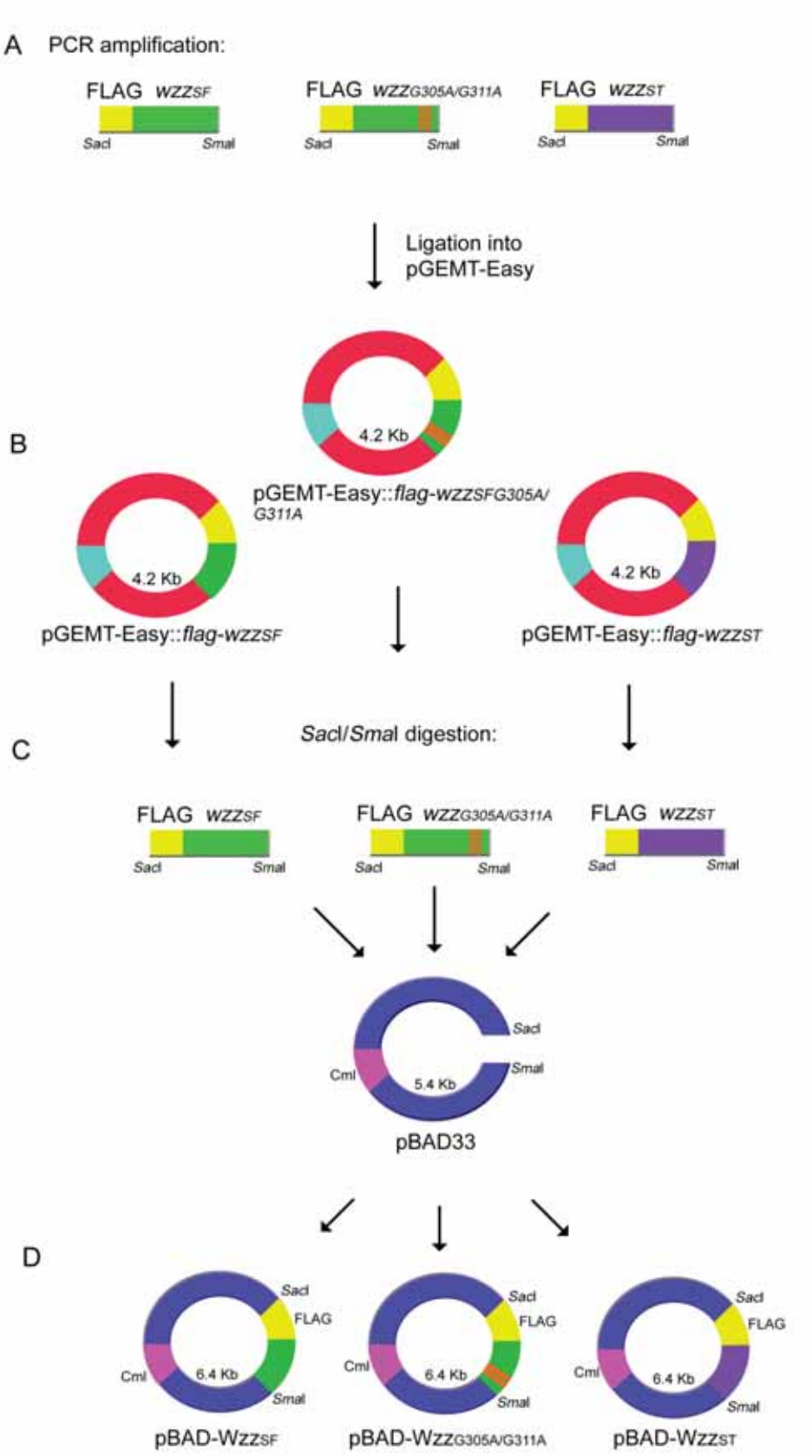
The results from these experiments suggested that Wzz_{ST} and Wzz_{ST}/Wzz_{SF} hybrids were capable of interacting with Wzz_{SF} even though they share significantly less sequence identity with Wzz_{SF} than Wzz_{SF} does with Wzz_{SF} mutants. To explain the bimodal LPS produced by RMA2162 (pRMCD113), I hypothesised that glycine residues at aa 305 and 311 (substituted to alanines in pRMCD113) are necessary for establishing Wzz:Wzz interactions, and despite sharing over 99% sequence identity, these altered amino acids in TM2 prevented $Wzz_{G305A/G311A}$ protein from interacting with wild-type Wzz_{SF}. I also hypothesised that the conserved aa sequence identity at these positions in the TM2 region shared by Wzz_{ST} and Wzz_{SF} (Figure 4.2) may explain why these two proteins are able to interact, unlike Wzz_{SF} and $Wzz_{G305A/G311A}$. To investigate these putative protein interactions, constructs were made in order to epitope tag these proteins and to perform co-purification assays.

4.3 CONSTRUCTION PLASMIDS ENCODING FLAG-TAGGED WZZ

To directly investigate the ability of Wzz_{ST} and $Wzz_{G305A/G311A}$ to interact with wild-type Wzz_{SF}, constructs were designed and produced to explore the putative interactions of these proteins via physical means, i.e. a pull-down type assay. The vector pBAD33 (Guzman *et al.*, 1995) was utilised for the construction of these plasmids (Figure 4.5). Oligonucleotide primers were designed that incorporated a 1 x FLAG tag, restriction sites for the enzymes *Sac*I and *Sma*I to amplify wild-type Wzz_{SF}, Wzz_{ST} and $Wzz_{G305A/G311A}$ (Table 2.4). Templates to amplify wzz_{SF}, wzz_{ST} and $wzz_{G305A/G311A}$ were derived from plasmids pRMCD30, pRMCD80 and pRMCD113, respectively (Figure 4.5). The amplified fragments were cloned

Figure 4.5 Construction of pBAD33-based plasmids encoding FLAG-tagged Wzz proteins

Plasmids were constructed as described in section 4.3. Briefly, forward and reverse primers were designed to incorporate *Sac*I and *Sma*I restriction sites respectively, in order to amplify wzz_{SF}, $wzz_{G305A/G311A}$, and wzz_{ST} (A) from templates pRMCD30 (wzz_{SF}), pRMCD113 ($wzz_{G305A/G311A}$) and pRMCD80 (wzz_{ST}) respectively, and ligated into pGEMT-Easy (B). The vector pBAD33 was used (incorporating an arabinose inducible promoter, pACYC184 origin of replication, pBAD promoter, and chloramphenicol resistance), digested with *Sac*I and *Sma*I (C), and ligated with the similarly digested pGEMT-Easy plasmids carrying wzz_{SF}, $wzz_{G305A/G311A}$, and wzz_{ST}, to produce pBAD-Wzz_{SF}, pBAD-$Wzz_{G305A/G311A}$, and pBAD-Wzz_{ST} (D). Indicated in blue is pBAD33, chloramphenicol resistance in magenta, the FLAG tag in yellow, pGEMT-Easy in red, Amp resistance in cyan, wzz_{SF}/$wzz_{G305A/G311A}$ in green (and the double glycine mutation in orange), and wzz_{ST} in purple.


A
PCR amplification:
FLAG wzzSF
SacI
SmaI
FLAG wzzG305A/G311A
SacI
SmaI
FLAG wzzST
SacI
SmaI
Ligation into
pGEMT-Easy
B
4.2 Kb
pGEMT-Easy::flag-wzzSFG305A/
G311A
4.2 Kb
pGEMT-Easy::flag-wzzSF
4.2 Kb
pGEMT-Easy::flag-wzzST
C
SacI/SmaI digestion:
FLAG wzzSF
SacI
SmaI
FLAG wzzG305A/G311A
SacI
SmaI
FLAG wzzST
SacI
SmaI
SacI
SmaI
CmI
5.4 Kb
pBAD33
D
SacI
FLAG
CmI
6.4 Kb
SmaI
pBAD-WzzSF
SacI
FLAG
CmI
6.4 Kb
SmaI
pBAD-WzzG305A/G311A
SacI
FLAG
CmI
6.4 Kb
SmaI
pBAD-WzzST

into pGEM-T Easy, restricted with the aforementioned enzymes, and ligated into similarly digested pBAD33 (Figure 4.5). The integrity of the sequences cloned into pBAD33 was confirmed by DNA sequencing (Figure 4.6).

4.4 COMPLEMENTATION OF LPS OAG MODAL LENGTH BY FLAG-WZZ$_{SF}$

To assure the functionality of pBAD-Wzz$_{SF}$ and its ability to express detectable levels of FLAG-Wzz$_{SF}$ protein, the construct and pBAD33 vector control were transformed into strain RMA2741 (*S. flexneri* serotype Y *wzz*::Km^{R} F'*lacI*^{q}) (Table 2.1). Strains were grown and induced with arabinose, and samples for LPS analyses were prepared (section 2.11). LPS analyses indicated that RMA2741 (pBAD33) had LPS with unregulated Oag chain length (Figure 4.7A, lane 3). However, FLAG-Wzz$_{SF}$ completely restored the wild-type LPS Oag modal chain length of 11-17 RUs and was identical to pRMCD30 in this regard (Figure 4.7A, lanes 4 and 2 respectively). This indicated that the FLAG-tagged Wzz$_{SF}$ protein expressed by pBAD-Wzz$_{SF}$ was fully functional. Plasmids pBAD-Wzz$_{ST}$ and pBAD-Wzz$_{G305A/G311A}$ were also transformed into strain RMA2741, and samples for LPS analyses were likewise prepared and analysed to ensure that the modal chain lengths conferred by the new constructs were those that were observed above (section 4.2). The resulting LPS Oag modal chain lengths confirmed the 3-10 RU VS type modal length conferred by FLAG-Wzz$_{G305A/G311A}$ (Figure 4.7B, lane 2) and the longer L- type LPS Oag modal chain length (17-28 RUs) conferred by FLAG-Wzz$_{ST}$ (Figure 4.7B, lane 3). These LPS profiles are similar to those obtained for the pET-17b based plasmids (section 4.2, Figures 4.3 and 4.4).

Figure 4.6 DNA sequence of wzz_{SF}, $wzz_{G305A/G311A}$ and wzz_{ST} cloned fragments in pBAD33 via *Sac*I and *Sma*I sites

Listed here are the sequences of the inserts cloned into pBAD33 to generate the three pBAD33-based constructs, A) pBAD-Wzz_{SF}, B) pBAD-$Wzz_{G305A/G311A}$, and C) pBAD-Wzz_{ST}. Magenta indicates the *Sac*I sites, yellow indicates the start codon, green shows the FLAG tag, blue indicates the second codon of the gene and the stop codon, and red illustrates the *Sma*I sites.

A) pBAD-Wzz$_{SF}$ insert

```
1       GAGCTCAGGA GATATCTTAT GGACTACAAG GACGACGACG ACAAGAGAGT AGAAAATAAT
61      AATGTTTCTG GGCAAAACCA TGACCCGGAA CAGATTGATT TGATTGATTT ACTAGTGCAG
121     TTGTGGCGTG GCAAGATGAC AATTATCATT TCCGTCATTG TGGCTATTGC CCTGGCTATT
181     GGTTATTTGG CAGTAGCGAA GGAGAAATGG ACGTCAACAG CAATTATCAC TCAGCCCGAC
241     GTGGGGCAAA TTGCTGGCTA TAACAATGCC ATGAATGTTA TCTATGGTCA GGCTGCACCG
301     AAAGTATCGG ATTTGCAGGA GACGTTAATT GGTCGCTTCA GTTCTGCCTT CTCTGCATTA
361     GCAGAAACGC TGGATAATCA GGAAGAGCCA GAAAAACTTA CCATCGAACC TTCTGTTAAG
421     AACCAGCAAT TACCATTGAC TGTTTCTTAT GTTGGGCAAA CTGCAGAGGG CGCACAAATG
481     AAGTTGGCCC AATACATTCA GCAAGTTGAT GATAAAGTGA ATCAAGAGCT AGAAAAGGAT
541     CTCAAGGACA ACATTGCTCT GGGACGGAAA AACTTGCAGG ACTCTTTAAG AACCCAGGAA
601     GTGGTCGCGC AGGAGCAGAA AGATCTGCGT ATCCGTCAGA TTCAGGAAGC GTTGCAGTAT
661     GCGAATCAGG CGCAGGTGAC AAAGCCACAG GTTCAGCAGA CTGAAGATGT GACGCAAGAT
721     ACGTTGTTCC TTCTAGGGAG CGAAGCGCTG GAGTCGATGA TTAAGCATGA AGCGACTCGT
781     CCGTTGGTGT TCTCACCAAA CTACTATCAG ACACGTCAAA ACCTGTTGGA TATTGAAAAA
841     TTAAAGTTTG ATGATCTTGA TATTCATGCT TACCGCTATG TGATGAAACC GACGTTACCT
901     ATTCGTCGCG ATAGTCCGAA AAAGGCAATC ACCTTGATTC TGGCAGTGCT TCTGGGCGGC
961     ATGGTTGGCG CGGGGATTGT GTTGGGGCGT AACGCTCTGC GTAATTACAA CGCGAAGTAA
1021    TATTATTGTG CATTTAAGAG AAACGGGCAG GGTGGTGACA CCATGCCCGT TTTTTTTGCC
1081    GGATGCGATG CTGGCGCATC TTATCCGGCC TACGTGTGTT GAGATAATGT GTAGGCACGA
1141    TAAGTTTGCG CATCGGGCAA TGGCTCCGGG TGTGACAACA ACATCACACC TGCTCCCGGG
1201    G
```

B) pBAD-Wzz$_{G305A/G311A}$ insert

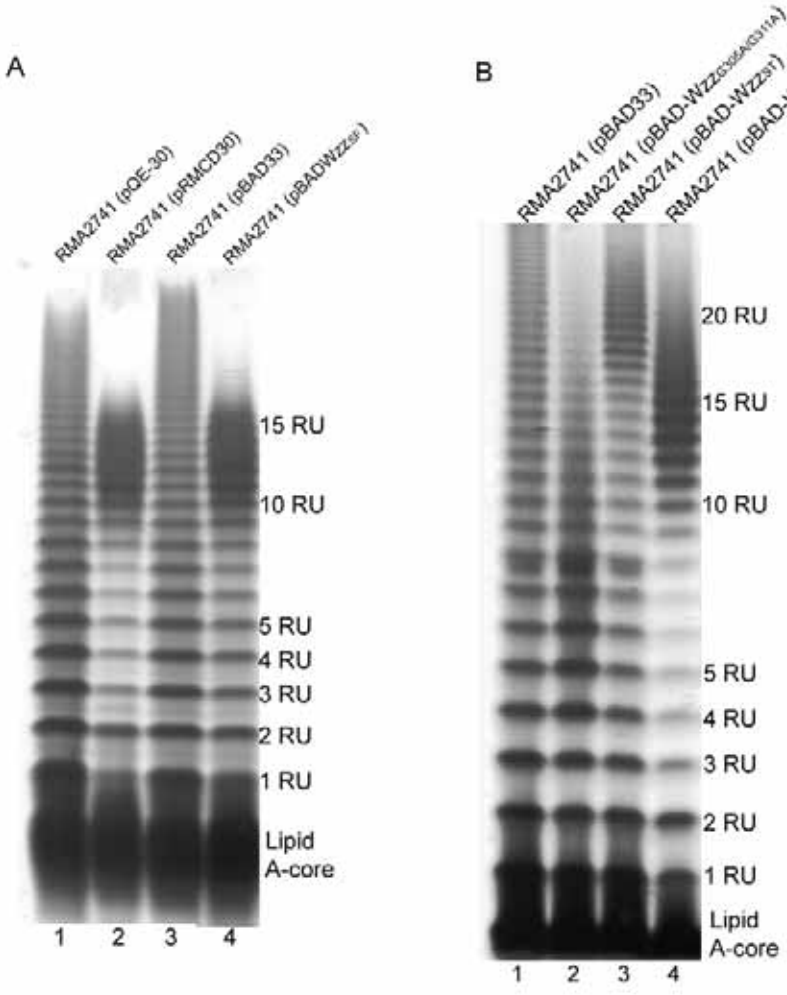
```
1       GAGCTCAGGA GATATCTTAT GGACTACAAG GACGACGACG ACAAGAGAGT AGAAAATAAT
61      AATGTTTCTG GGCAAAACCA TGACCCGGAA CAGATTGATT TGATTGATTT ACTAGTGCAG
121     TTGTGGCGTG GCAAGATGAC AATTATCATT TCCGTCATTG TGGCTATTGC CCTGGCTATT
181     GGTTATTTGG CAGTAGCGAA GGAGAAATGG ACGTCAACAG CAATTATCAC TCAGCCCGAC
241     GTGGGGCAAA TTGCTGGCTA TAACAATGCC ATGAATGTTA TCTATGGTCA GGCTGCACCG
301     AAAGTATCGG ATTTGCAGGA GACGTTAATT GGTCGCTTCA GTTCTGCCTT CTCTGCATTA
361     GCAGAAACGC TGGATAATCA GGAAGAGCCA GAAAAACTTA CCATCGAACC TTCTGTTAAG
421     AACCAGCAAT TACCATTGAC TGTTTCTTAT GTTGGGCAAA CTGCAGAGGG CGCACAAATG
481     AAGTTGGCCC AATACATTCA GCAAGTTGAT GATAAAGTGA ATCAAGAGCT AGAAAAGGAT
541     CTCAAGGACA ACATTGCTCT GGGACGGAAA AACTTGCAGG ACTCTTTAAG AACCCAGGAA
601     GTGGTCGCGC AGGAGCAGAA AGATCTGCGT ATCCGTCAGA TTCAGGAAGC GTTGCAGTAT
661     GCGAATCAGG CGCAGGTGAC AAAGCCACAG GTTCAGCAGA CTGAAGATGT GACGCAAGAT
721     ACGTTGTTCC TTCTAGGGAG CGAAGCGCTG GAGTCGATGA TTAAGCATGA AGCGACTCGT
781     CCGTTGGTGT TCTCACCAAA CTACTATCAG ACACGTCAAA ACCTGTTGGA TATTGAAAAA
841     TTAAAGTTTG ATGATCTTGA TATTCATGCT TACCGCTATG TGATGAAACC GACGTTACCT
901     ATTCGTCGCG ATAGTCCGAA AAAGGCAATC ACCTTGATTC TGGCAGTGCT TCTGGCCGGC
961     ATGGTTGGCG CGGCGATTGT GTTGGGGCGT AACGCTCTGC GTAATTACAA CGCGAAGTAA
1021    TATTATTGTG CATTTAAGAG AAACGGGCAG GGTGGTGACA CCATGCCCGT TTTTTTTGCC
1081    GGATGCGATG CTGGCGCATC TTATCCGGCC TACGTGTGTT GAGATAATGT GTAGGCACGA
1141    TAAGTTTGCG CATCGGGCAA TGGCTCCGGG TGTGACAACA ACATCACACC TGCTCCCGGG
1201    G
```

C) pBAD-Wzz$_{ST}$ insert

```
1       GAGCTCAGGA GATATCTTAT GGACTACAAG GACGACGACG ACAAGACAGT GGATAGTAAT
61      ACGTCTTCCG GGCGTGGGAA CGATCCGGAA CAGATTGATT TGATTGAGTT ATTGCTACAG
121     TTATGGCGTG GGAAGATGAC CATTATTGTA GCCGTTATTA TCGCCATTTT GCTGGCTGTA
181     GGCTACCTGA TGATAGCCAA AGAAAAATGG ACATCCACGG CGATTATTAC CCAACCTGAT
241     GCCGCGCAGG TTGCCACCTA TACCAACGCG CTCAACGTCT TGTATGGTGG GAATGCCCCC
301     AAAATCTCTG AAGTGCAGGC GAATTTTATA AGCCGCTTTA GCTCTGCGTT TTCCGCATTA
361     TCGGAAGTGC TGGATAATCA GAAAGAGCGG GAAAAGCTTA CCATTGAACA GTCGGTAAAA
421     GGGCAGGCGC TGCCACTCTC GGTTTCTTAT GTGAGTACTA CCGCTGAAGG GGCGCAGCGT
481     CGGCTGGCGG AATATATCCA ACAGGTGGAT GAAGAGGTCG CTAAGGAACT GGAAGTTGAC
541     CTGAAAGATA ACATCACGCT GCAAACCAAA ACGTTGCAGG AGTCCCTTGA GACGCAGGAA
601     GTTGTGGCGC AGGAGCAAAA AGATCTGCGT ATTAAGCAAA TCGAAGAAGC GTTGCGCTAT
661     GCGGATGAGG CCAAAATCAC GCAGCCGCAG ATTCAGCAAA CCCAGGATGT TACCCAGGAC
721     ACGATGTTCC TGTTGGGGAG CGATGCGCTA AAATCGATGA TACAGAACGA AGCGACGCGT
781     CCACTGGTCT TTTCTCCGGC CTATTACCAG ACGAAGCAGA CACTGCTGGA CATTAAAAAT
841     CTGAAAGTGA CTGCCGATAC GGTGCACGTC TATCGTTATG TGATGAAGCC GACGCTGCCG
901     GTCCGTCGCG ATAGCCCGAA AACAGCCATT ACCCTTGTGC TGGCTGTATT GCTGGGTGGG
961     ATGATCGGTG CCGGGATTGT GCTGGGACGC AATGCGCTAC GTAGCTATAA GCCAAAAGCC
1021    TTGTAACCCG GG
```

Figure 4.7 LPS Oag modal chain length conferred by pBAD-Wzz_{SF}, pBAD-$Wzz_{G305A/G311A}$ and pBAD-Wzz_{ST}

S. flexneri RMA2741 strains harbouring plasmids were grown in LB + Amp (for pQE-30 based plasmids) and LB + Cml (for pBAD33-based plasmids) and induced with IPTG (pQE-30-based plasmids) or arabinose (pBAD33-based plasmids) for 1.5 h as described in section 2.10.3 and 2.10.4. LPS samples were prepared, electrophoresed on a SDS 15% polyacrylamide gel and silver stained (section 2.11). Strains in each lane are as follows: A) 1) RMA2741 (pQE-30), 2) RMA2741 (pRMCD30), 3) RMA2741 (pBAD33), and 4) RMA2741 (pBAD-Wzz_{SF}). B) 1) RMA2741 (pBAD33), 2) RMA2741 (pBAD-$Wzz_{G305A/G311A}$), 3) RMA2741 (pBAD-Wzz_{ST}), and 4) RMA2741 (pBAD-Wzz_{SF}). The lipid A-core and number of repeat units (RUs) are indicated on the right. Each lane contains approximately 1.3 x 10^8 cells.


A
RMA2741 (pQE-30)
RMA2741 (pRMCD30)
RMA2741 (pBAD33)
RMA2741 (pBADWzz^ST)
15 RU
10 RU
5 RU
4 RU
3 RU
2 RU
1 RU
Lipid
A-core
1 2 3 4
B
RMA2741 (pBAD33)
RMA2741 (pBAD-Wzz^G305A/G311A)
RMA2741 (pBAD-Wzz^ST)
RMA2741 (pBAD-Wzz^SF)
20 RU
15 RU
10 RU
5 RU
4 RU
3 RU
2 RU
1 RU
Lipid
A-core
1 2 3 4

CHAPTER FOUR

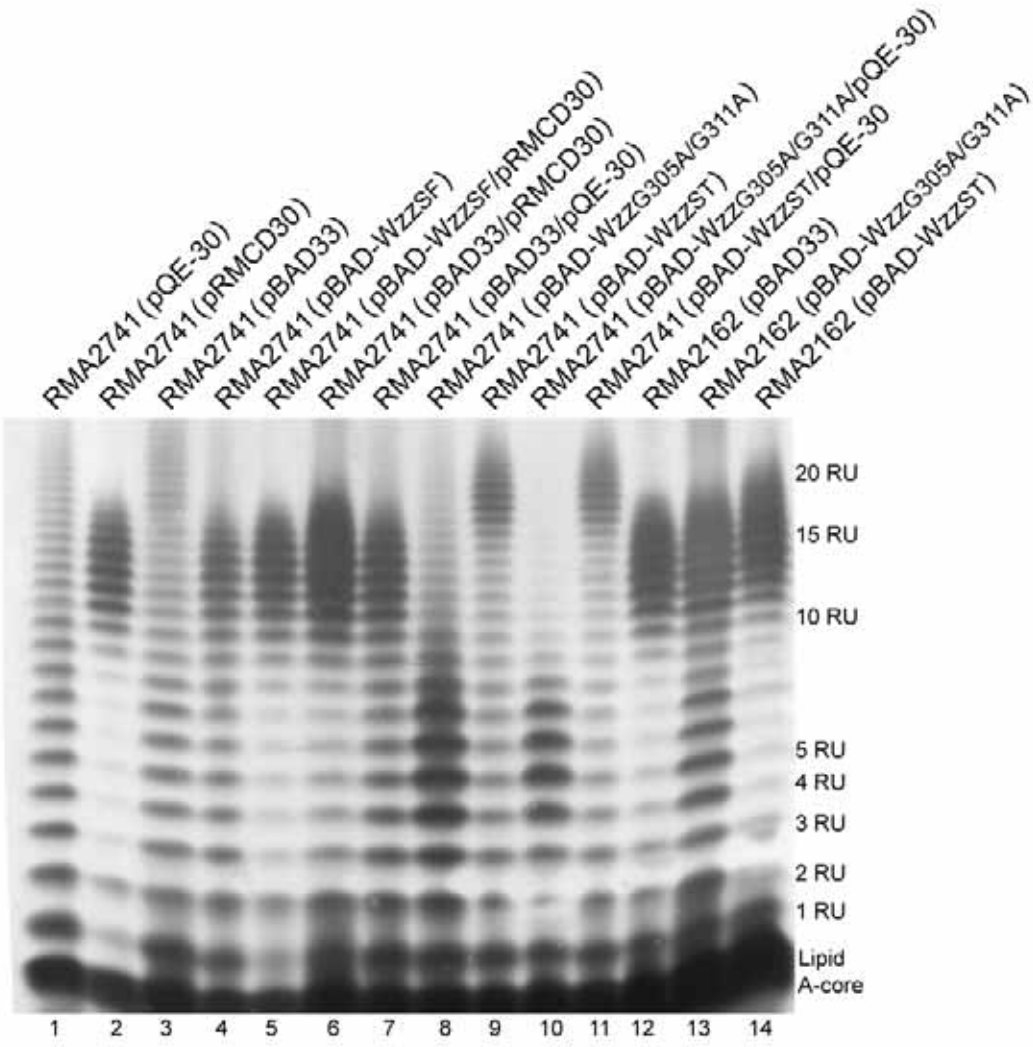
4.5 CO-EXPRESSION OF pBAD-BASED WZZ PROTEINS WITH WILD-TYPE WZZ_{SF}

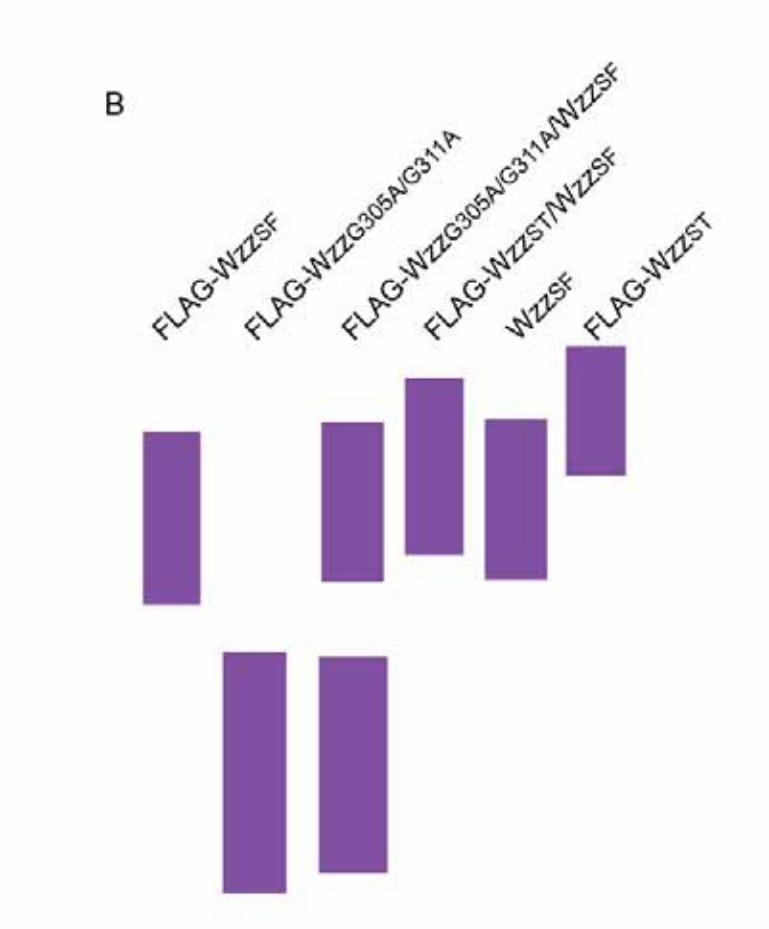
In order to confirm the original phenomenon of $Wzz_{G305A/G311A}$ and Wzz_{ST} conferring novel LPS profiles when co-expressed with Wzz_{SF} wild-type (section 4.2), the pBAD-based constructs were transformed into strain RMA2162. These strains were grown, induced with arabinose and LPS samples were prepared and subjected to SDS-PAGE and silver staining. The resulting LPS profile demonstrated that RMA2162 harbouring pBAD-Wzz_{ST} had LPS of a single Oag modal chain length of approximately 12-22 RUs (Figure 4.8A, lane 14), longer in total than wild-type Wzz_{SF} (Figure 4.8A, lane 13), but shorter than the L-type 17-28 RUs Oag conferred in RMA2741 (Figure 4.8, lane 9), and that observed when pBAD-Wzz_{ST} is co-expressed with pQE-30 (Figure 4.8, lane 11). Conversely, RMA2162 harbouring pBAD-$Wzz_{G305A/G311A}$ had an identical bimodal LPS Oag modal chain length profile as that shown by the previous co-expression assay with pET-17b based plasmids (section 4.2); two LPS Oag modal chain lengths exist within the same strain, the VS type conferred by pBAD-$Wzz_{G05A/G311A}$, and that of wild-type conferred by Wzz_{SF} in RMA2162 (Figure 4.8A, lane 13). Figure 4.8B illustrates a schematic summary of the LPS Oag chain modal lengths conferred by FLAG-$Wzz_{G305A/G311A}$ and FLAG-Wzz_{ST} with or without Wzz_{SF} (extrapolated from the gel in Figure 4.8A). This figure also indicates the average modal length of the strain co-expressing FLAG-Wzz_{ST} with wild-type Wzz_{SF}. A similar LPS Oag chain modal length profile was observed in the RMA2741 strain harbouring both pRMCD30 (His_6-Wzz_{SF}) and pBAD-Wzz_{ST} (FLAG-Wzz_{ST}), with only arabinose induction (Figure 4.8C, lane 2). LPS Oag chain bimodality was also observed for RMA2741 harbouring both pRMCD30 and pBAD-$Wzz_{G305A/G311A}$ induced only with arabinose, although the phenotype was not seen as distinctly on this gel (Figure 4.8C, lane 1). However, RMA2741 harbouring pRMCD30 and pBAD-$Wzz_{G305A/G311A}$ induced with IPTG appeared to

Figure 4.8 LPS Oag chain modal length conferred by Wzz proteins encoded by pBAD33-based plasmids, and co-expression of Wzz proteins encoded by pBAD plasmids with wild-type Wzz_{SF} and Wzz proteins encoded by pQE-30-based plasmids

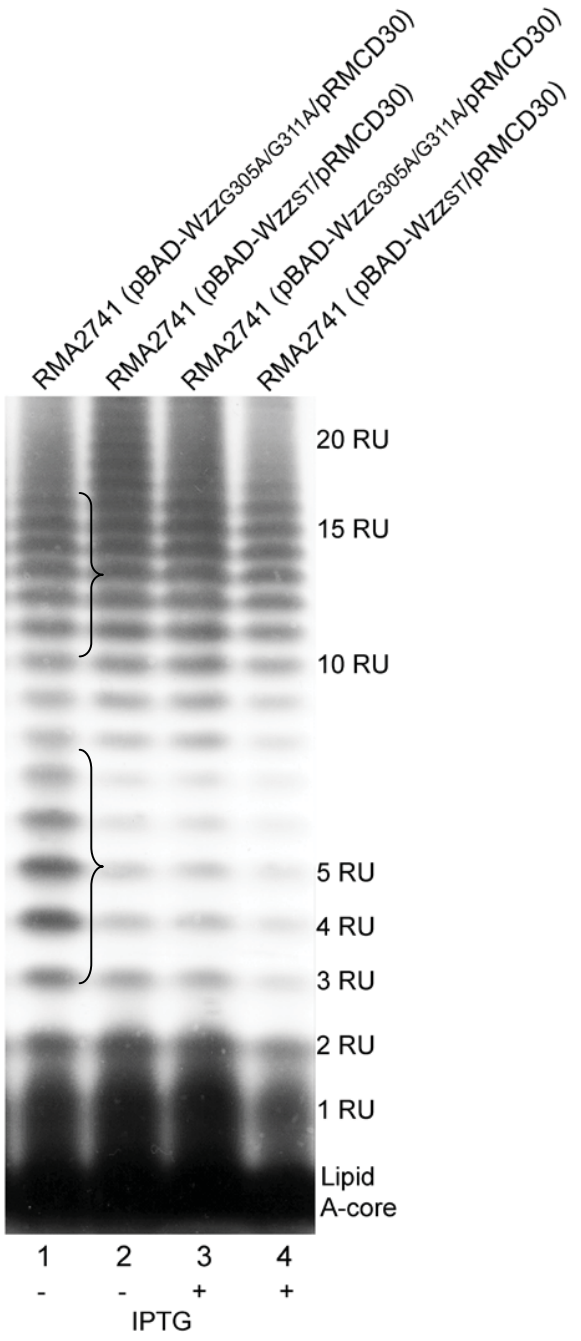
RMA2741 and RMA2162 strains harbouring plasmids were grown in LB + Amp (pQE-30 based plasmids) and LB + Cml (pBAD33-based plasmids), and LB +Amp/Cml for strains harbouring both plasmids. Strains (excluding C, lanes 3 and 4) were induced with 0.2% arabinose for 2 h as described in section 2.12. LPS samples were prepared, electrophoresed on a SDS 15% polyacrylamide gel and silver stained. The lipid A-core and number of repeat units (RUs) are indicated. Each lane contains approximately 1.3 x 10^8 cells. A) Strains in each lane are as follows: 1) RMA2741 (pBAD33), 2) RMA2741 (pBAD-Wzz_{SF}), 3) RMA2741 (pQE-30), and 4) RMA2741 (pRMCD30), 5) RMA2741 (pRMCD30/pBAD-Wzz_{SF}), 6) RMA2741 (pRMCD30/pBAD33), 7) RMA2741 (pQE-30/pBAD-Wzz_{SF}), 8) RMA2741 (pBAD-$Wzz_{G305A/G311A}$), 9) RMA2741 (pBAD-Wzz_{ST}), 10) RMA2741 (pBAD-$Wzz_{G305A/G311A}$/pQE-30), 11) RMA2741 (pBAD-Wzz_{ST}/pQE-30), 12) RMA2162 (pBAD-Wzz_{SF}), 13) RMA2162 (pBAD-$Wzz_{G305A/G311A}$) and 14) RMA2162 (pBAD-Wzz_{ST}). B) A schematic diagram illustrating the distribution of the LPS Oag chain lengths of strains expressing FLAG-Wzz_{SF}, FLAG-Wzz_{ST}, FLAG-$Wzz_{G305A/G311A}$, wild-type Wzz_{SF} (RMA2162 (pBAD33)), and also of FLAG-Wzz_{ST} or FLAG-$Wzz_{G305A/G311A}$ co-expressed with wild-type Wzz_{SF}; extrapolated from the LPS profiles in A). C) Lanes are as follows: 1) RMA2741 (pBAD-$Wzz_{G305A/G311A}$/pRMCD30), 2) RMA2741 (pBAD-Wzz_{ST}/pRMCD30), 3) RMA2741 (pBAD-$Wzz_{G305A/G311A}$/pRMCD30) and 4) RMA2741 (pBAD-Wzz_{ST}/pRMCD30). IPTG induction (or absence of) is indicated by (+) and (-).

A


RMA2741 (pQE-30)
RMA2741 (pRMCD30)
RMA2741 (pBAD33)
RMA2741 (pBAD-WzzSF)
RMA2741 (pBAD-WzzSF/pRMCD30)
RMA2741 (pBAD33/pRMCD30)
RMA2741 (pBAD33/pQE-30)
RMA2741 (pBAD-WzzG305A/G311A)
RMA2741 (pBAD-WzzST)
RMA2741 (pBAD-WzzG305A/G311A/pQE-30)
RMA2741 (pBAD-WzzST/pQE-30
RMA2162 (pBAD33)
RMA2162 (pBAD-WzzG305A/G311A)
RMA2162 (pBAD-WzzST)
20 RU
15 RU
10 RU
5 RU
4 RU
3 RU
2 RU
1 RU
Lipid
A-core
1 2 3 4 5 6 7 8 9 10 11 12 13 14


B
FLAG-WZZSF
FLAG-WZZG305A/G311A
FLAG-WZZG305A/G311A/WZZSF
FLAG-WZZST/WZZSF
WZZSF
FLAG-WZZST

C


RMA2741 (pBAD-WzzG305A/G311A/pRMCD30)
RMA2741 (pBAD-WzzST/pRMCD30)
RMA2741 (pBAD-WzzG305A/G311A/pRMCD30)
RMA2741 (pBAD-WzzST/pRMCD30)
20 RU
15 RU
10 RU
5 RU
4 RU
3 RU
2 RU
1 RU
Lipid
A-core
1 2 3 4
- - + +
IPTG

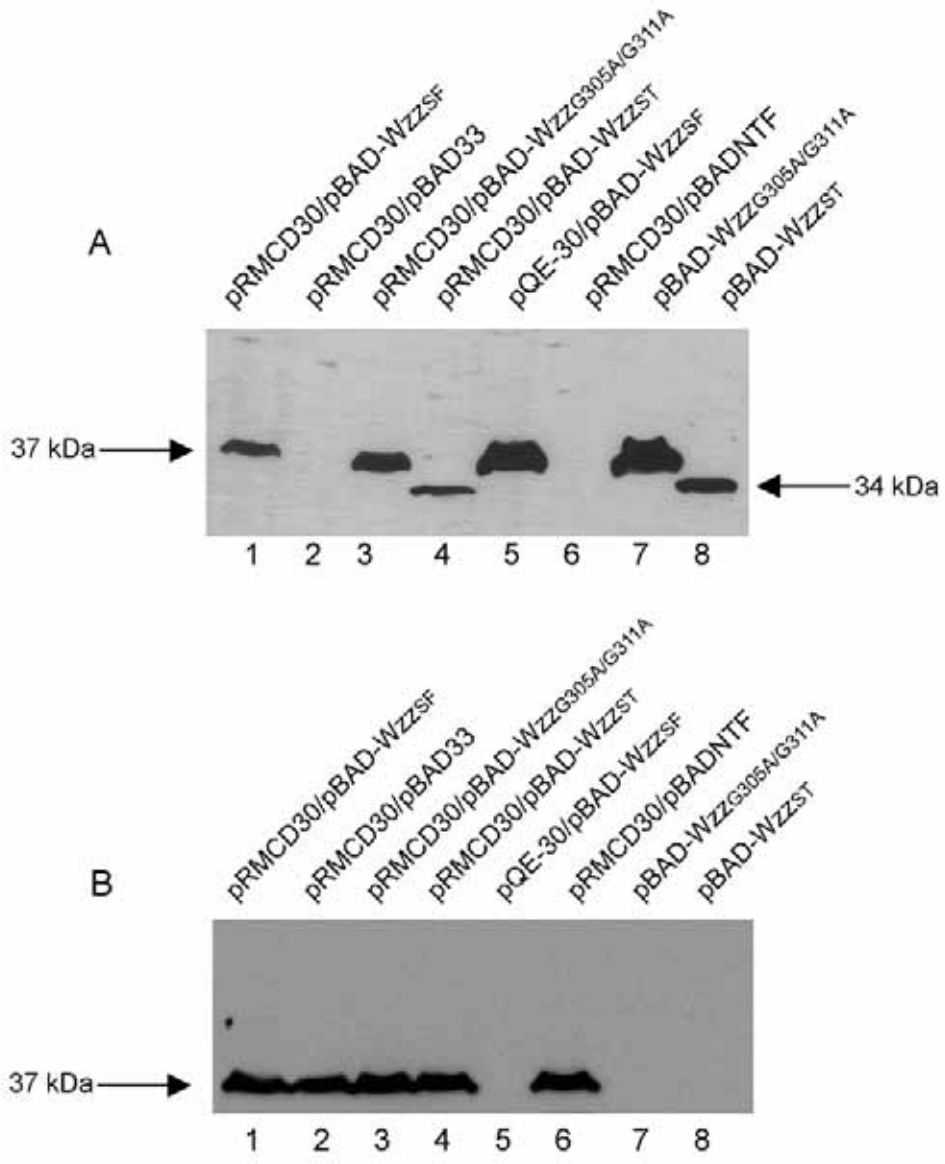
not exhibit bimodality (Figure 4.8C, lane 4), suggesting that higher levels of His_6-Wzz_{SF} can outcompete $Wzz_{G305A/G311A}$, and the bimodal phenotype is not observed (see section 7.7).

4.6 DETECTION OF FLAG-TAGGED WZZ PROTEINS

The pBAD33-based plasmids encoding FLAG-Wzz proteins were to be utilised in co-purification assays along with pRMCD30 (encoding His_6-Wzz_{SF}), hence Western immunoblotting was performed to detect FLAG-Wzz proteins with FLAG M2 mAb, and also to detect His_6-Wzz_{SF} with an anti-His mAb in strains co-expressing both plasmids. RMA2741 harbouring pBAD-Wzz_{ST}, pBAD-Wzz_{SF}, pBAD-$Wzz_{G305A/G311A}$, (singularly or with pQE-30 or pRMCD30) were grown, induced with arabinose and IPTG (section 2.10) and whole cell lysate samples were subjected to SDS-PAGE and Western immunoblotting with either anti-FLAG or anti-His. Figure 4.9A illustrates the detection of FLAG-Wzz proteins with anti-FLAG, and shows that the ~34 kDa FLAG-Wzz_{ST} protein was detected in RMA2741 (pBAD-Wzz_{ST}) (Figure 4.9A, lane 8), and also detected in RMA2741 (pBAD-Wzz_{ST} and pRMCD30) (Figure 4.9A, lane 4). The ~37 kDa FLAG-$Wzz_{G305A/G311A}$ protein was also detected, either when expressed alone or with His_6-Wzz_{SF} (Figure 4.9A lane 7 and 3, respectively). Under these conditions, FLAG-Wzz_{SF} was also detected in RMA2741 (pBAD-$Wzz_{SF,}$ pRMCD30) (Figure 4.9A, lane 1) and in RMA2741 (pBAD-Wzz_{SF}, pQE-30) (Figure 4.9A, lane 5). No bands were detected for RMA2741 (pRMCD30, pBAD33) (Figure 4.9A, lane 2), and RMA2741 (pRMCD30, pBADNTF (a FLAG-encoding vector control, Marolda *et al.* (2004)) (Figure 4.9A, lane 6). The same strains were subjected to Western immunoblotting with an anti-His mAb. His_6-Wzz_{SF} was detected in strains RMA2741 (pRMCD30/pBAD-Wzz_{SF}) (Figure 4.9B, lane 1), RMA2741 (pRMCD30/pBAD33) (Figure 4.9B, lane 2), RMA2741 (pRMCD30/pBAD-$Wzz_{G305A/G311A}$) (Figure 4.9B, lane 3), RMA2741 (pRMCD30/pBAD-Wzz_{ST}) (Figure 4.9B, lane 4) and RMA2741 (pRMCD30/pBADNTF) (Figure 4.9B, lane 6). No detectable protein was present in samples from RMA2741 (pQE-30/pBAD-Wzz_{SF})

Figure 4.9 Analysis of His_6-Wzz_{SF} and FLAG-Wzz_{SF} protein expression in *S. flexneri* strain RMA2741

Strains were grown in LB + Cml (and Amp for pQE-30 based plasmids), induced with 0.5 mM IPTG (pQE-30 based plasmids) and 0.2% arabinose (pBAD33-based plasmids) for 1 h, and whole cell lysates were electrophoresed on SDS 15% polyacrylamide gel, and detected with A) anti-FLAG (Invitrogen) with a dilution of 1:2000, and B) anti-His at a dilution of 1:1000 (section 2.11). Prestained Benchmark protein marker was used to determine protein sizes (Invitrogen). Under these conditions, FLAG-Wzz_{SF} migrates as a 37 kDa band, and FLAG-Wzz_{ST} as a 34 kDa band. Each lane contains approximately 2 x 10^8 cells. A) The lanes are as follows: 1) RMA2741 (pRMCD30/pBAD-Wzz_{SF}), 2) RMA2741 (pRMCD30/pBAD33), 3) RMA2741 (pRMCD30/pBAD-$Wzz_{G305A/G311A}$), 4) RMA2741 (pRMCD30/pBAD-Wzz_{ST}), 5) RMA2741 (pQE-30/pBAD-Wzz_{SF}), 6) RMA2741 (pRMCD30/pBADNTF), 7) RMA2741 (pBAD-$Wzz_{G305A/G311A}$), 8) RMA2741 (pBAD-Wzz_{ST}). B) The lanes are as follows: 1) RMA2741 (pRMCD30/pBAD-Wzz_{SF}), 2) RMA2741 (pRMCD30/pBAD33), 3) RMA2741 (pRMCD30/pBAD-$Wzz_{G305A/G311A}$), 4) RMA2741 (pRMCD30/pBAD-Wzz_{ST}), 5) RMA2741 (pQE-30/pBAD-Wzz_{SF}), 6) RMA2741 (pRMCD30/pBADNTF), 7) RMA2741 (pBAD-$Wzz_{G305A/G311A}$), 8) RMA2741 (pBAD-Wzz_{ST}).

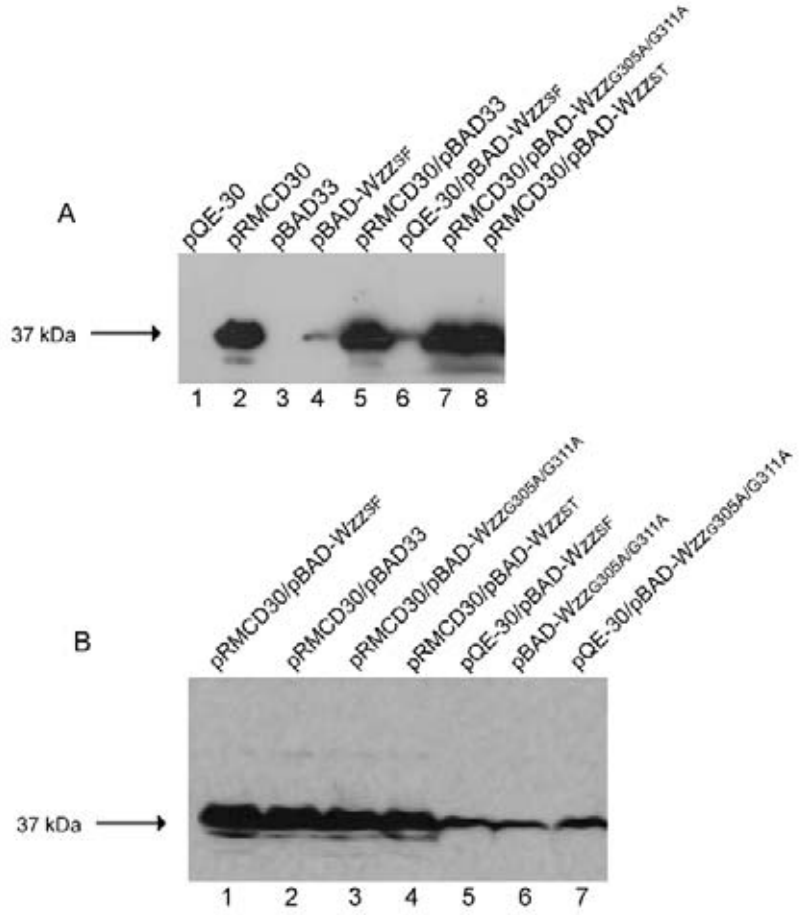

A
pRMCD30/pBAD-WzzSF
pRMCD30/pBAD33
pRMCD30/pBAD-WzzG305A/G311A
pRMCD30/pBAD-WzzST
pQE-30/pBAD-WzzSF
pRMCD30/pBADNTF
pBAD-WzzG305A/G311A
pBAD-WzzST
37 kDa
34 kDa
1 2 3 4 5 6 7 8
B
pRMCD30/pBAD-WzzSF
pRMCD30/pBAD33
pRMCD30/pBAD-WzzG305A/G311A
pRMCD30/pBAD-WzzST
pQE-30/pBAD-WzzSF
pRMCD30/pBADNTF
pBAD-WzzG305A/G311A
pBAD-WzzST
37 kDa
1 2 3 4 5 6 7 8

(Figure 4.9B, lane 5), RMA2741 (pBAD-$Wzz_{G305A/G311A}$) (Figure 4.9B, lane 7) or RMA2741 (pBAD-Wzz_{ST}) (Figure 4.9B, lane 8). Collectively, these data indicated that M2 FLAG did not detect His_6-Wzz_{SF} and anti-His did not detect FLAG-tagged Wzz proteins.

Wzz protein expression levels were compared by Western immunoblotting using affinity purified polyclonal anti-Wzz_{SF} on RMA2741 strains harbouring pRMCD30 or pBAD-Wzz_{SF} following induction with either IPTG or arabinose, respectively. The results illustrated that when expressed from pRMCD30, a relatively high level of His_6-Wzz_{SF} was detected (Figure 4.10A, lane 2), whilst when expressed from pBAD-Wzz_{SF}, the level of FLAG-Wzz_{SF} detected was considerably lower in comparison (Figure 4.10A, lane 4). This showed that although FLAG-tagged Wzz_{SF} proteins were readily detected by anti-FLAG (Figure 4.9A, lanes 1, 3-5 and 7-8), the actual levels of Wzz proteins expressed by pRMCD30 and pBAD-Wzz_{SF} were disproportionate under either IPTG or arabinose induction conditions, respectively. When RMA2741 (pQE-30, pBAD-Wzz_{SF}) was induced with both IPTG and arabinose, a very low amount of FLAG-Wzz_{SF} was detected (Figure 4.10A, lane 6), as compared to RMA2741 harbouring pRMCD30 with pBAD33, pBAD-$Wzz_{G305A/G311A}$ or pBAD-Wzz_{ST}, in which much higher levels of Wzz_{SF} were detected (Figure 4.10A, lanes 5, 7 and 8), due to the higher expression of His_6-Wzz_{SF} from pRMCD30. A relatively low level of expression of FLAG-$Wzz_{G305A/G311A}$ was also observed. RMA2741 (pBAD-$Wzz_{G305A/G311A}$) (Figure 4.10B, lane 6) and RMA2741 (pQE-30/pBAD-$Wzz_{G305A/G311A}$) (Figure 4.10B, lane 7) had a reduced level of Wzz_{SF} protein compared to Wzz produced by RMA2741 (pRMCD30/pBAD-Wzz_{SF}), RMA2741 (pRMCD30/pBAD33), RMA2741 (pRMCD30/pBAD-$Wzz_{G305A/G311A}$) and RMA2741 (pRMCD30/pBAD-Wzz_{ST}) (Figure 4.10B, lanes 1-4, respectively).

Figure 4.10 Comparison of Wzz protein expression by pRMCD30 and pBAD33-based constructs in *S. flexneri* strain RMA2741

Strains were grown in LB + Amp and or Cml, induced with 0.5 mM IPTG and 0.2% arabinose for 1 h (section 2.10.3 and 2.10.4), and whole cell lysates were subjected to 15% SDS-PAGE and detected with affinity purified polyclonal antibody anti-Wzz_{SF} at a dilution of 1:1000 (section 2.10.1 and 2.10.2). Prestained Benchmark protein marker was used to determine protein sizes (Invitrogen). The lanes are: A) 1) RMA2741 (pQE-30), 2) RMA2741 (pRMCD30), 3) RMA2741 (pBAD33), 4) RMA2741 (pBAD-Wzz_{SF}), 5) RMA2741 (pRMCD30/pBAD33), 6) RMA2741 (pQE-30/pBAD-Wzz_{SF}), 7) RMA2741 (pRMCD30/pBAD-$Wzz_{G305A/G311A}$) and 8) RMA2741 (pRMCD30/pBAD-Wzz_{ST}). B) The lanes are: 1) RMA2741 (pRMCD30/pBAD-Wzz_{SF}), 2) RMA2741 (pRMCD30/pBAD33), 3) RMA2741 (pRMCD30/pBAD-$Wzz_{G305A/G311A}$), 4) RMA2741 (pRMCD30/pBAD-Wzz_{ST}), 5) RMA2741 (pQE-30/pBAD-Wzz_{SF}), 6) RMA2741 (pBAD-$Wzz_{G305A/G311A}$), and 7) RMA2741 (pQE-30/pBAD-$Wzz_{G305A/G311A}$). Each lane contains approximately 2 x 10^8 cells.


A
pQE-30
pRMCD30
pBAD33
pBAD-WzzSF
pRMCD30/pBAD33
pQE-30/pBAD-WzzSF
pRMCD30/pBAD-WzzG305A/G311A
pRMCD30/pBAD-WzzST
37 kDa
1 2 3 4 5 6 7 8
B
pRMCD30/pBAD-WzzSF
pRMCD30/pBAD33
pRMCD30/pBAD-WzzG305A/G311A
pRMCD30/pBAD-WzzST
pQE-30/pBAD-WzzSF
pBAD-WzzG305A/G311A
pQE-30/pBAD-WzzG305A/G311A
37 kDa
1 2 3 4 5 6 7

CHAPTER FOUR

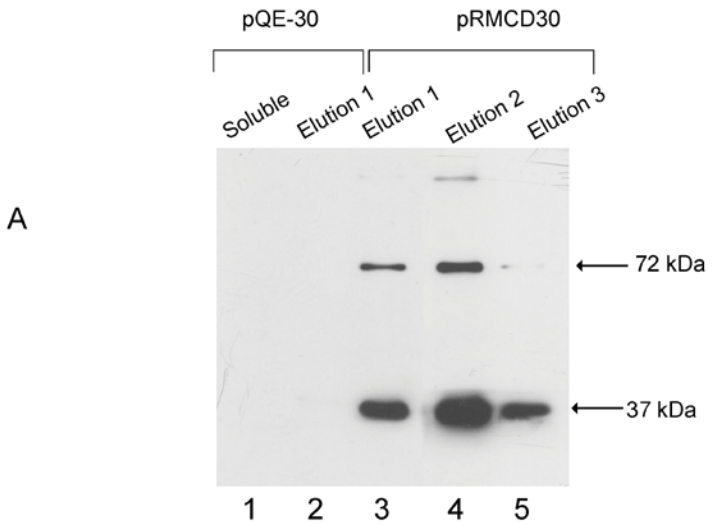
4.7 CO-PURIFICATION ASSAY

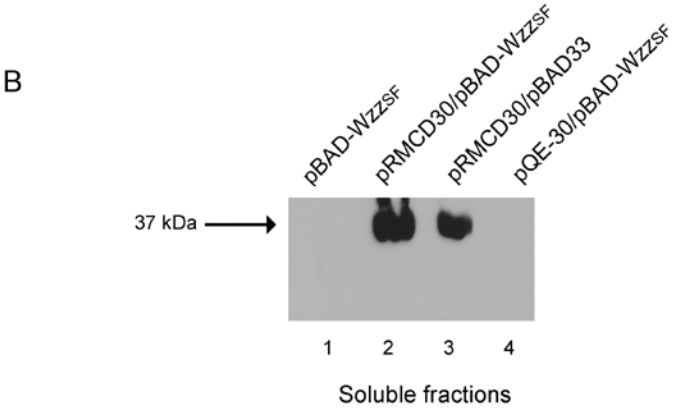
4.7.1 PURIFICATION OF HIS_6-WZZ_{SF}

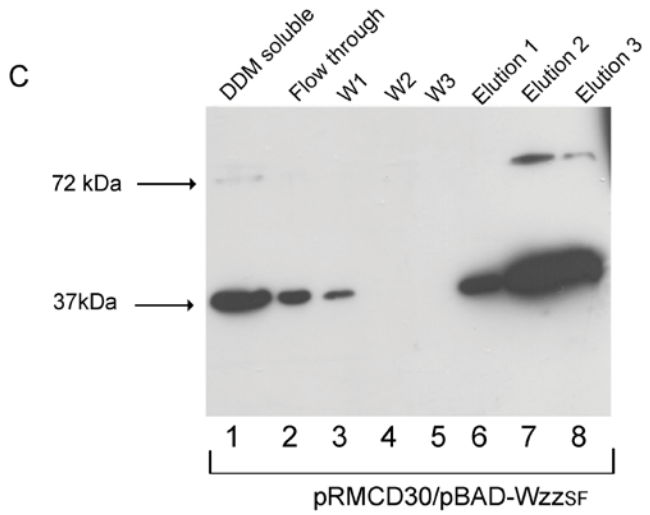
In order to purify His_6-Wzz_{SF}, an appropriate detergent was required to solubilise the protein. The detergents perfluoro-octanoic acid (PFO) (1% w/v), D-dodecyl β-D maltoside (DDM) (1% w/v), Triton X-100 (2% w/v), Sarkosyl (1% w/v), Nonidet P40 (1.5% w/v) and Zwittergent (1% w/v) (Table 2.3) were used to solubilise and purify His_6-Wzz_{SF}. RMA2741 (pRMCD30) was grown, induced with IPTG to express His_6-Wzz_{SF}, and cell lysis was achieved with a French press (section 2.10.3). Whole membranes were prepared and treated with the listed detergents, and the resulting soluble and insoluble fractions were subjected to SDS-PAGE and Western immunoblotting and probed with affinity purified anti-Wzz_{SF} (section 2.10.1 and 2.10.2). The results indicated that Wzz_{SF} protein was easily extracted using DDM as >90% was present in the soluble fraction (data not shown). Protein purification of the His_6-Wzz_{SF} using Qiagen Ni-NTA was then undertaken (section 2.10.3). RMA2741 carrying pRMCD30 was grown, induced with 0.5 mM IPTG as described in section 2.10.3, and whole membranes were solubilised with 1% (w/v) DDM. The DDM soluble fraction was incubated with the Ni-NTA, washed, and eluted with four fractions of increasing concentrations of imidazole (section 2.10.3). Figure 4.11A illustrates the detection of His_6-Wzz_{SF} with a high concentration of purified protein detected in the second elution fraction (Figure 4.11A, lane 4). Figure 4.11B illustrates the detection of His_6-Wzz_{SF} in soluble fractions of RMA2741 (pRMCD30/pBAD-Wzz_{SF}) and RMA2741 (pRMCD30/pBAD33) (Figure 4.11B, lanes 2 and 3, respectively) and as expected His_6-Wzz_{SF} could not be detect in DDM soluble samples from RMA2741 (pBAD-Wzz_{SF}) and RMA2741 (pQE-30/pBAD-Wzz_{SF}) (Figure 4.11B, lanes 1 and 4, respectively). Figure 4.11C shows the detection of Wzz_{SF} protein in fractions from strain RMA2741 harbouring both pRMCD30 (His_6-Wzz_{SF}) and pBAD-Wzz_{SF} (FLAG-Wzz_{SF}) (Figure 4.11C).

Figure 4.11 Analysis of Wzz_{SF} protein purification

S. flexneri RMA2741 strains harbouring pQE-30-based or pBAD-33 based constructs were grown in LB + Amp (for pQE-30 based plasmids) and/or Cml (pBAD33-based plasmids). Strains in A) were induced with 0.5 mM IPTG for 1.5 h, and strains in B) and C) were induced with 0.2% (w/v) arabinose for 1.5 h. Strains were treated with DDM, and purified with Ni-NTA (section 2.11). SDS-PAGE was performed on a 15% gel, and ~25 µL from each fraction was loaded in each well. Fractions were detected with affinity purified polyclonal antibody anti-Wzz_{SF} diluted 1:1000 (A and C), or anti-His mAb diluted 1:1000 (B). Prestained Benchmark protein marker was used to determine protein sizes (Invitrogen). The lanes are as follows: A) 1) RMA2741 (pQE-30) soluble fraction, 2) RMA2741 (pQE-30) elution fraction, 3) RMA2741 (pRMCD30) elution fraction 1, 4) RMA2741 (pRMCD30) elution fraction 2, and 5) RMA2741 (pRMCD30) elution fraction 3. B) The lanes are as follows: 1) RMA2741 (pBAD-Wzz_{SF}) soluble fraction, 2) RMA2741 (pRMCD30/pBAD-Wzz_{SF}) soluble fraction, 3) RMA2741 (pRMCD30/pBAD33) soluble fraction, and 4) RMA2741 (pQE-30/pBAD-Wzz_{SF}) soluble fraction. C) The soluble fractions, wash fractions and elution fractions of RMA2741 (pRMCD30/pBAD-Wzz_{SF}) are as follows: 1) DDM soluble fraction, 2) flow through, 3) wash 1, 4) wash 2, 5) wash 3, 6) elution fraction 1, 7) elution fraction 2, and 8) elution fraction 3.


A
pQE-30
pRMCD30
Soluble
Elution 1
Elution 1
Elution 2
Elution 3
72 kDa
37 kDa
1
2
3
4
5


B
pBAD-WzzSF
pRMCD30/pBAD-WzzSF
pRMCD30/pBAD33
pQE-30/pBAD-WzzSF
37 kDa
1
2
3
4
Soluble fractions

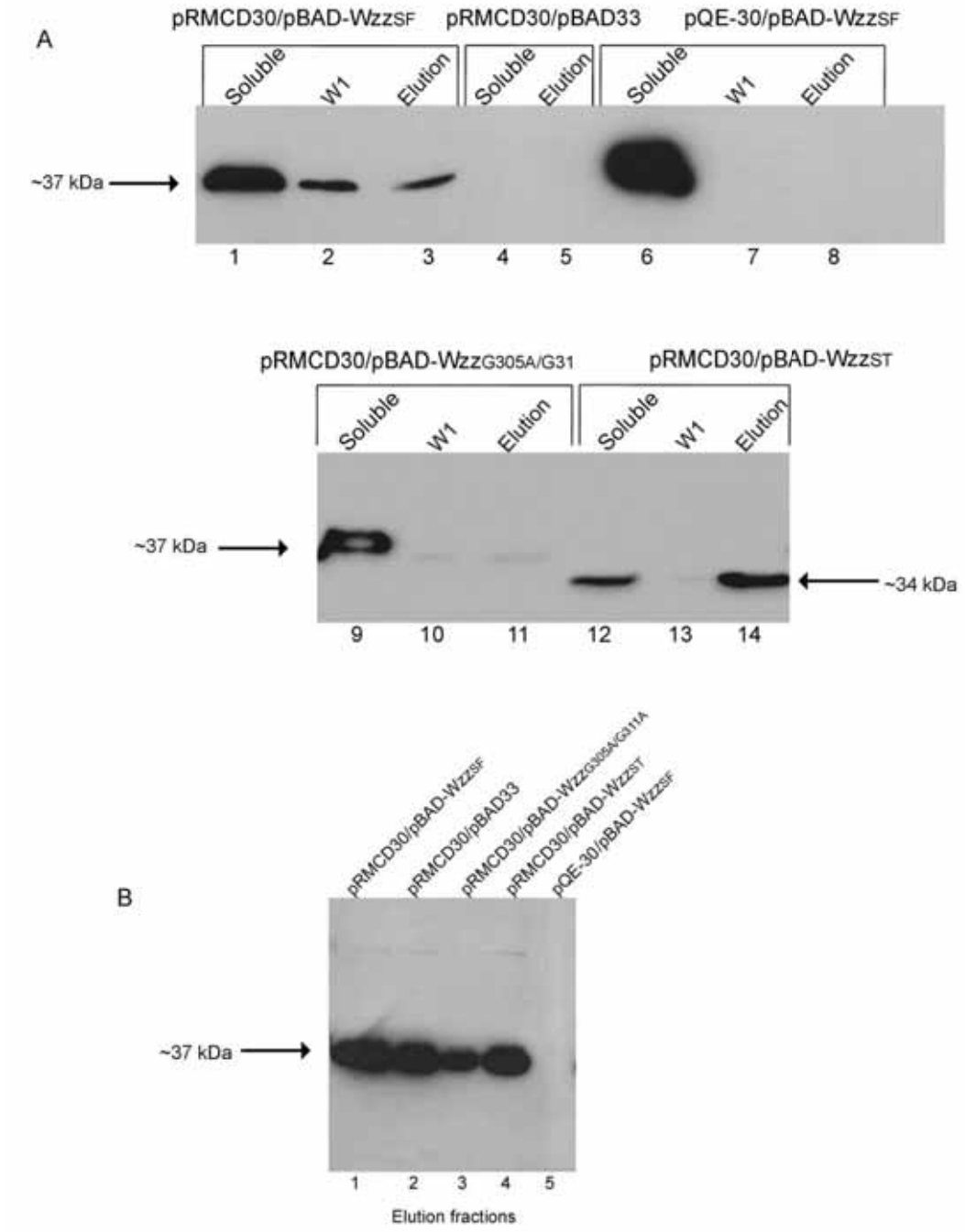

C
DDM soluble
Flow through
W1
W2
W3
Elution 1
Elution 2
Elution 3
72 kDa
37kDa
1
2
3
4
5
6
7
8
pRMCD30/pBAD-WzzSF

4.7.2 CO-PURIFICATION OF HIS$_6$-WZZ$_{SF}$ AND FLAG-WZZ$_{SF}$

LPS Oag modal chain length phenotypes observed in sections 4.2 and 4.4 were predicted to be the result of Wzz-Wzz interactions, such that bimodality was a result of reduced interactions between Wzz$_{G305A/G311A}$ and Wzz$_{SF}$, and that mono-modality was achieved via successful interactions between Wzz$_{ST}$ and Wzz$_{SF}$, a co-purification assay was undertaken. The rationale for this experiment was to determine if purification of His$_6$-Wzz$_{SF}$ would correspondingly co-purify FLAG-Wzz$_{ST}$ or either FLAG-Wzz$_{G305A/G311A}$. RMA2741 harbouring pRMCD30 and either pBAD-Wzz$_{SF}$, pBAD-Wzz$_{ST}$ or pBAD-Wzz$_{G305A/G311A}$, were grown overnight, subcultured in 80 mL of LB, and due to the higher expression levels of pRMCD30 compared to pBAD-Wzz$_{SF}$, strains were induced with 0.2% (w/v) arabinose only (section 2.10.3.2) and grown for a further 2 h, centrifuged and lysed with a French Press (section 2.10.3.2). The lysate was centrifuged once more to remove unlysed cells and any potential inclusion bodies (section 2.10.3.2), and whole membranes were prepared from the ultracentrifugation of the lysate. The resulting pellet was then treated with 1% (w/v) DDM. This fraction was subjected to further ultracentrifugation, and the supernatant was then incubated with equilibrated Qiagen Ni-NTA beads (section 2.10.3.2), washed, and bound His$_6$-Wzz$_{SF}$ was eluted off the beads using 250 mM of imidazole (section 2.10.3). Fractions were subjected to Western immunoblotting with α-FLAG (M2) to detect the presence of interacting FLAG-tagged protein in the elution fractions, and also anti-His to confirm the presence of His$_6$-Wzz$_{SF}$ in the samples. When the resulting RMA2741 (pRMCD30, pBAD-Wzz$_{SF}$) fractions were subjected to Western immunoblotting with anti-FLAG antibody, FLAG-Wzz$_{SF}$ was detected in the fractions containing His$_6$-Wzz$_{SF}$ (Figure 4.12A, lane 3), as expected. When His$_6$-Wzz$_{SF}$ was purified from RMA2741 (pBAD33), FLAG-tagged protein was not detected in the elution fraction (Figure 4.12A, lane 5). Likewise, in the sample prepared from strain RMA2741 (pQE-30, pBAD-Wzz$_{SF}$), FLAG-tagged protein was absent in

Figure 4.12 Purification of His_6-Wzz_{SF} and detection of co-purified FLAG-Wzz_{ST} protein

S. flexneri RMA2741 strains harbouring pQE-30 and pBAD33-based constructs were grown in LB + Amp (pQE-30 based plasmids) and Cml (pBAD33-based plasmids), induced with arabinose for 1.5 h, treated with 1% (w/v) DDM, and purified with Ni-NTA (section 2.10.3). SDS-PAGE was performed on a 15% gel (section 2.10.1), and ~25 µL from each fraction was loaded in each well. Fractions were detected with anti-FLAG mAb (Sigma) diluted 1:2000 (section 2.10.2). Prestained Benchmark protein marker was used to determine protein sizes (Invitrogen). A) The lanes are: 1) RMA2741 (pRMCD30/pBAD-Wzz_{SF}) soluble fraction, 2) RMA2741 (pRMCD30/pBAD-Wzz_{SF}) wash fraction, 3) RMA2741 (pRMCD30/pBAD-Wzz_{SF}) elution fraction, 4) RMA2741 (pRMCD30/pBAD33) soluble fraction, 5) RMA2741 (pRMCD30/pBAD33) elution fraction, 6) RMA2741 (pQE-30/pBAD-Wzz_{SF}) soluble fraction, 7) RMA2741 (pQE-30/pBAD-Wzz_{SF}) wash fraction, 8) (pQE-30/pBAD-Wzz_{SF}) elution fraction, 9) RMA2741 (pRMCD30/pBAD-$Wzz_{G305A/G311A}$) soluble fraction, 10) RMA2741 (pRMCD30/pBAD-$Wzz_{G305A/G311A}$) wash fraction, 11) RMA2741 (pRMCD30/pBAD-$Wzz_{G305A/G311A}$) elution fraction, 12) RMA2741 (pRMCD30/pBAD-Wzz_{ST}) soluble fraction, 13) RMA2741 (pRMCD30/pBAD-Wzz_{ST}) wash fraction and 14) RMA2741 (pRMCD30/pBAD-Wzz_{ST}) elution fraction. B) Elution fractions were detected with anti-His mAb diluted 1:1000. Lanes are as follows: 1) RMA2741 (pRMCD30/pBAD-Wzz_{SF}) elution fraction, 2) RMA2741 (pRMCD30/pBAD33) elution fraction, 3) RMA2741 (pRMCD30/pBAD-$Wzz_{G305A/G311A}$) elution fraction, 4) RMA2741 (pRMCD30/pBAD-Wzz_{ST}) elution fraction, and 5) RMA2741 (pQE-30/pBAD-Wzz_{SF}) elution fraction.


A
pRMCD30/pBAD-WzzSF
pRMCD30/pBAD33
pQE-30/pBAD-WzzSF
Soluble
W1
Elution
Soluble
Elution
Soluble
W1
Elution
~37 kDa
1
2
3
4
5
6
7
8
pRMCD30/pBAD-WzzG305A/G31
pRMCD30/pBAD-WzzST
Soluble
W1
Elution
Soluble
W1
Elution
~37 kDa
~34 kDa
9
10
11
12
13
14
B
pRMCD30/pBAD-WzzSF
pRMCD30/pBAD33
pRMCD30/pBAD-WzzG305A/G311A
pRMCD30/pBAD-WzzST
pQE-30/pBAD-WzzSF
~37 kDa
1
2
3
4
5
Elution fractions

the elution fraction (Figure 4.12A, lane 8). When His_6-Wzz_{SF} was purified from strain RMA2741 (pBAD-$Wzz_{G305A/G311A}$ and pRMCD30), there was very little FLAG-$Wzz_{G305A/G311A}$ protein detected in the elution fraction (Figure 4.12A, lane 11), indicating that His_6-Wzz_{SF} and FLAG-$Wzz_{G305A/G311A}$ interact poorly. In samples prepared from strain RMA2741 (pBAD-Wzz_{ST}, pRMCD30), FLAG-Wzz_{ST} was readily detected in the elution fraction, indicating that FLAG-Wzz_{ST} and His_6-Wzz_{SF} are capable of interacting (Figure 4.12A, lane 14). As expected, His_6-Wzz_{SF} was detected in the elution fractions of RMA2741 (pRMCD30/pBAD-Wzz_{SF}), RMA2741 (pRMCD30/pBAD33), and RMA2741 (pRMCD30/pBAD-$Wzz_{G305A/G311A}$) (Figure 4.12B, lanes 1-4, respectively). There was no detection of His_6-Wzz_{SF} in the elution fraction of RMA2741 (pQE-30/pBAD-Wzz_{SF}) (Figure 4.12B, lane 5), as expected.

These data illustrate that Wzz_{ST} interacts with wild-type Wzz_{SF}, however $Wzz_{G305A/G311A}$ does not interact very well with Wzz_{SF}, and provides direct evidence to support the hypothesis to explain the LPS profiles described in section 4.2 and 4.4.

4.8 ANALYSIS OF LPS PRODUCED BY STRAINS CO-EXPRESSING WZZi CLASS II MUTANTS AND WZZ_{SF} WILD-TYPE

The bimodal LPS Oag modal chain length observed when $Wzz_{G305A/G311A}$ and wild-type Wzz_{SF} were co-expressed, and conversely the single Oag modal chain length resulting from the co-expression of Wzz_{ST} and wild-type Wzz_{SF} (section 4.2 and 4.4), suggested that VS type LPS Oag modal chain length may be related to the inability of the corresponding mutant Wzz_{SF} to interact with wild-type Wzz_{SF}. This is supported by the data in section 4.6. We wanted to explore whether the observed lack of interaction between $Wzz_{G305A/G311A}$ and Wzz_{SF} was uniquely due to the conserved TM2 residue changes, or whether perhaps it may be a general property of other mutations that confer a VS LPS Oag modal chain length, e.g the Class II Wzz_i mutants characterised in section 3.3. Thus, a similar assay to that described in

section 4.4 was employed; however, due to the requirement of *lacI*^{q} in the cell to provide repression of the pQE-30 promoter, RMA2162 and RMA2163 strains were not suitable for expression of the Class II Wzz_i proteins. Therefore, pCDF-Duet1 (Novagen) was transformed into the strains, generating MPRMA143 (RMA2163 *wzz*::Km^R (pCDF-Duet1)), and MPRMA142 (RMA2162 (pCDF-Duet1) (Table 2.1). These strains were electroporated with plasmids encoding the Class II Wzz_i proteins i191, i219, i231, i247 and i255 (Table 2.2, Figure 3.3). MPMA142 and MPRMA143 harbouring Wzz_i were grown, induced with IPTG and subjected to SDS-PAGE and silver staining (section 2.11). The LPS Oag modal chain length conferred by all Class II proteins in the MPRMA143 background was the VS 2-10 RUs (Figure 4.13, lanes 1, 2, 3, 9 and 10), as expected (section 3.3). Conversely, the LPS Oag modal chain length exhibited by the Class II proteins when expressed in MPRMA142 showed bimodality (Figure 4.13, lanes 4, 5, 6, 11 and 12) similar to that seen by co-expression of $Wzz_{G305A/G311A}$ and Wzz_{SF}.

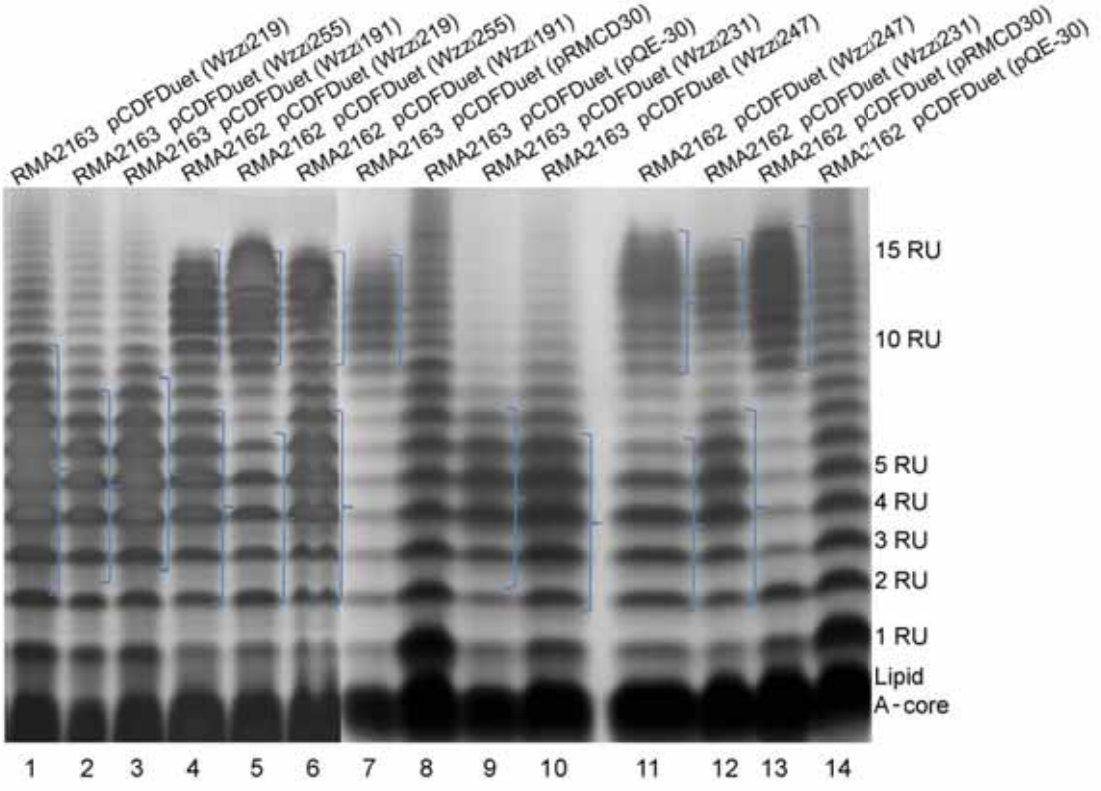
From these data, it can be concluded that although the type of mutations differ drastically between $Wzz_{G305A/G311A}$ and the Class II Wzz_i mutants, there exists a general correlation between co-dominance of the Wzz proteins exhibiting VS conferred Oag modal chain length with wild-type Wzz_{SF}. As the bimodal LPS profile exhibited by $Wzz_{G305A/G311A}$ lead to the discovery that interactions between $Wzz_{G305A/G311A}$ and wild-type Wzz_{SF} was poor (section 4.6.2), the similarly observed LPS profile exhibited by Class II co-expressed with wild-type Wzz_{SF} mutants suggests these Wzz_i proteins may also be unable to interact with Wzz_{SF}. Due to time constraints, this could not be assessed.

4.9 ANALYSIS OF LPS PRODUCED BY STRAINS CO-EXPRESSING WZZi CLASS V MUTANTS AND WZZ_{SF} WILD-TYPE

Since the Class II Wzz_i mutants displayed bimodal LPS Oag modal chain length profiles when co-expressed with wild-type Wzz_{SF}, the interaction of L-type (~16-25 RUs)

Figure 4.13 LPS Oag modal chain length conferred by Class II Wzz_i mutants expressed in RMA2162 (pCDF-Duet1) and RMA2163 (pCDF-Duet1)

MPRMA142 (RMA2162 (pCDF-Duet1)) and MPRMA143 (RMA2163 (pCDF-Duet1)) harbouring Wzz_i Class II plasmids Wzz_{i191}, Wzz_{i219}, Wzz_{i231}, Wzz_{i247} and Wzz_{i255} were grown in LB + Amp and induced with 0.5 mM IPTG for 1.5 h (section 2.10.3). LPS samples were prepared, electrophoresed on a SDS 15% polyacrylamide gel and silver stained (section 2.11). Strains in each lane are as follows: 1) RMA2163 (pCDF-Duet1/i219), 2) RMA2163 (pCDF-Duet1/i255), 3) RMA2163 (pCDF-Duet1/i191), 4) RMA2162 (pCDF-Duet1/i219), 5) RMA2162 (pCDF-Duet1/i255), 6) RMA2162 (pCDF-Duet1/i191), 7) RMA2163 (pCDF-Duet1/pRMCD30), 8) RMA2163 (pCDF-Duet1/pQE-30), 9) RMA2163 (pCDF-Duet1/i231), 10) RMA2163 (pCDF-Duet1/i247), 11) RMA2162 (pCDF-Duet1/i247), 12) RMA2162 (pCDF-Duet1/i231), 13) RMA2162 (pCDF-Duet1/pRMCD30) and 14) RMA2162 (pCDF-Duet1/pQE-30). The lipid A-core and number of repeat units (RUs) are indicated, and significant banding of Oag modal chain length is indicated with brackets. Each lane contains approximately 1.3 x 10^8 cells.


RMA2163 pCDFDuet (Wzzi219)
RMA2163 pCDFDuet (Wzzi255)
RMA2163 pCDFDuet (Wzzi191)
RMA2162 pCDFDuet (Wzzi219)
RMA2162 pCDFDuet (Wzzi255)
RMA2162 pCDFDuet (Wzzi191)
RMA2163 pCDFDuet (pRMCD30)
RMA2163 pCDFDuet (pQE-30)
RMA2163 pCDFDuet (Wzzi231)
RMA2163 pCDFDuet (Wzzi247)
RMA2162 pCDFDuet (Wzzi247)
RMA2162 pCDFDuet (Wzzi231)
RMA2162 pCDFDuet (pRMCD30)
RMA2162 pCDFDuet (pQE-30)
15 RU
10 RU
5 RU
4 RU
3 RU
2 RU
1 RU
Lipid
A-core
1 2 3 4 5 6 7 8 9 10 11 12 13 14

Oag modal length Class V Wzz_i mutants (section 3.3) with wild-type Wzz_{SF} was investigated. MPRMA142 and MPRMA143 (Table 2.1) harbouring the plasmids encoding Wzz_{i128} and Wzz_{i131}, were grown, induced with IPTG and LPS samples were subjected to SDS-PAGE and silver staining (section 2.11). The resulting LPS profiles revealed that when the Wzz_i proteins were expressed in MPRMA143, the LPS Oag modal chain length was longer (~16-25 RUs) than wild type, as previously determined in section 3.3. When expressed in the MPRMA142 background, the LPS had a single LPS Oag modal chain length (Figure 4.14, lanes 4 and 6), and was shorter than when expressed in MPRMA143 (Figure 4.14, lanes 5 and 7). The LPS Oag modal chain lengths conferred by Wzz_{i128} and Wzz_{i131} were ~12-20 and ~12-22 RUs, respectively, slightly longer than the wild-type LPS Oag modal chain length (Figure 4.14, lane 1 and 3). A bimodal LPS Oag chain length was not observed and this would have easily been observed in these gels. This phenomenon was previously observed when wild type Wzz_{SF} and either Wzz_{ST} or FLAG-Wzz_{ST} were co-expressed (sections 4.2 and 4.4).

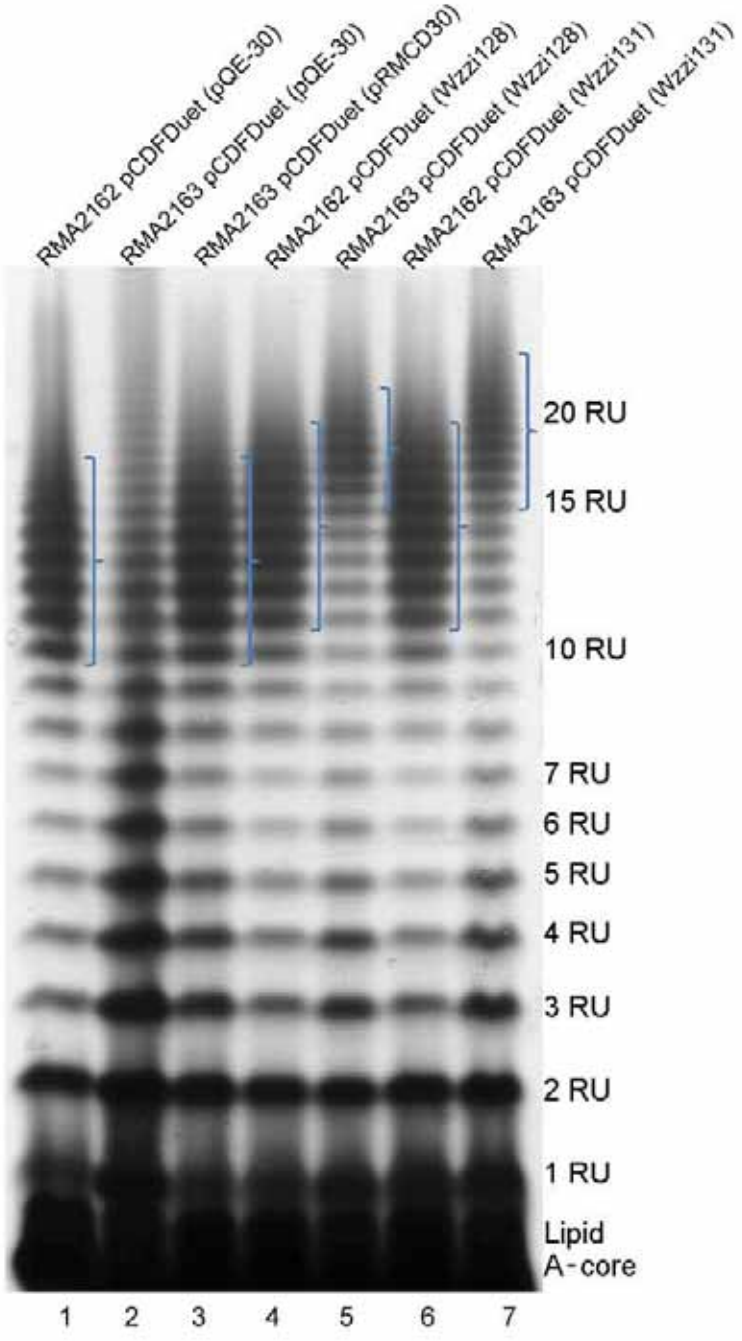
Collectively, the data suggest that Class II Wzz_i mutants are unable to interact with wild-type Wzz_{SF}, whereas Class V Wzz_i mutants are able to interact with wild-type Wzz_{SF}.

4.10 SUMMARY

In this chapter, the effect on LPS of Wzz mutant proteins co-expressed with wild-type Wzz_{SF} was investigated. Bimodal LPS Oag modal chain length was observed for strains expressing $Wzz_{G305A/G311A}$ and wild-type Wzz_{SF}, whilst Wzz_{ST} co-expressed with wild-type resulted in LPS which had mono-modal Oag chain modal length. Plasmids expressing FLAG-tagged $Wzz_{G305A/G311A}$, Wzz_{SF} and Wzz_{ST} were constructed, and strains expressing these plasmids emulated the previously observed LPS Oag modal chain lengths with pET-17b based plasmids, including the bimodality observed with $Wzz_{G305A/G311A}$. Co-purification assays were conducted to determine if the LPS Oag modal chain length phenotypes were a result of Wzz_{ST} interacting with wild-type Wzz_{SF}, and $Wzz_{G305A/G311A}$ failing to interact with

Figure 4.14 LPS Oag modal chain length conferred by Class V Wzz_i mutants expressed in RMA2162 (pCDF-Duet1) and RMA2163 (pCDF-Duet1)

MPRMA142 (*S. flexneri* RMA2162 (pCDF-Duet1)) and MPRMA143 (RMA2163 (pCDF-Duet1)) harbouring Wzz_i plasmids encoding the Class V Wzz_{i128} and Wzz_{i131} proteins were grown in LB + Amp, induced with 0.5 mM IPTG for 1.5 h (section 2.10.3) and LPS samples were prepared, electrophoresed on a SDS 15% polyacrylamide gel and silver stained (section 2.11). Strains in each lane are as follows: 1) RMA2162 (pCDF-Duet1/pQE-30), 2) RMA2163 (pCDF-Duet1/pQE-30), 3) RMA2163 (pCDF-Duet1/pRMCD30), 4) RMA2162 (pCDF-Duet1/Wzz_{i128}), 5) RMA2163 (pCDF-Duet1/Wzz_{i128}), 6) RMA2162 (pCDF-Duet1/Wzz_{i131}) and 7) RMA2163 (pCDF-Duet1/Wzz_{i131}). The lipid A-core and number of repeat units (RUs) are indicated, and significant banding of Oag modal chain length is indicated with brackets. Each lane contains approximately 1.3 x 10^8 cells.


RMA2162 pCDFDuet (pQE-30)
RMA2163 pCDFDuet (pQE-30)
RMA2163 pCDFDuet (pRMCD30)
RMA2162 pCDFDuet (Wzzi128)
RMA2163 pCDFDuet (Wzzi128)
RMA2162 pCDFDuet (Wzzi131)
RMA2163 pCDFDuet (Wzzi131)
20 RU
15 RU
10 RU
7 RU
6 RU
5 RU
4 RU
3 RU
2 RU
1 RU
Lipid
A-core
1
2
3
4
5
6
7

wild-type Wzz_{SF}. It was shown that FLAG-Wzz_{ST} co-purified with His_6-Wzz_{SF}, indicating that the two proteins interacted, whereas FLAG-$Wzz_{G305A/G311A}$ did not co-purify with His_6-Wzz_{SF}, indicating that the two proteins interacted poorly. The resulting LPS Oag modal chain length conferred by co-expression of wild-type Wzz_{SF} and the VS-conferring Class II Wzz_i mutants was also investigated, and yielded the similarly observed bimodal phenotype exhibited by co-expression of wild-type Wzz_{SF} and $Wzz_{G305A/G311A}$. This indicates that bimodality appears to be a general property of VS-conferring Wzz_i mutants and that these mutants may in turn be unable to interact with Wzz_{SF}. Class V Wzz_i mutants which conferred a L-type LPS Oag modal chain length similar to Wzz_{ST} were also co-expressed with wild-type Wzz_{SF}, with the resulting strains exhibiting single LPS Oag modal chain lengths, suggesting Class V Wzz_i mutants are capable of interacting with wild-type Wzz_{SF}.