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 FIVE

STUDIES ON THE WZZ FLUORESCENT FUSION PROTEIN

5.1 INTRODUCTION

Fluorescent protein (FP) tags have been used extensively to visualise the position and behaviour of proteins inside the cell under different experimental conditions. Recently, studies using fluorescent tags have revealed new insight into protein physiology and localisation. In a previous study conducted in our laboratory, Wzz_{SF} was tagged with GFP, however this proved to be uninformative in determining the sub-cellular localisation of Wzz_{SF} (Daniels, 1999). The resulting hybrid protein demonstrated a decreased level of function in complementation assays with respect to LPS Oag modal chain length control (~15% of wild type activity) (Daniels, 1999). In this initial attempt to create a Wzz FP fusion, the GFP tag was fused at the C-terminus of Wzz_{SF}. However, because complete functionality was exhibited by the albeit much smaller N-terminal tagged FLAG-Wzz_{SF} and His_6-Wzz_{SF} proteins described in Chapter 4, it was decided to investigate if Wzz_{SF} tagged with a fluorescent protein at the N-terminus would yield a functional protein. This chapter describes the construction of a fully functional Wzz_{SF} protein tagged with the red FP mCherry (Shaner *et al.*, 2004), for the purpose of identifying Wzz_{SF} localisation within *S. flexneri* cells.

5.2 CONSTRUCTION OF mCHERRY-WZZ_{SF}

Primers MCherryF and MCherryR (Table 2.4) were designed to incorporate *Bam*HI and *Sph*I sites in the 5' and 3' of the mCherry coding region and also a linker sequence, and were used to amplify the coding region from plasmid pRSET-B (Shaner *et al.*, 2004). The amplified *mcherry* was cloned into pGEM-T Easy, and digested with the restriction enzyme

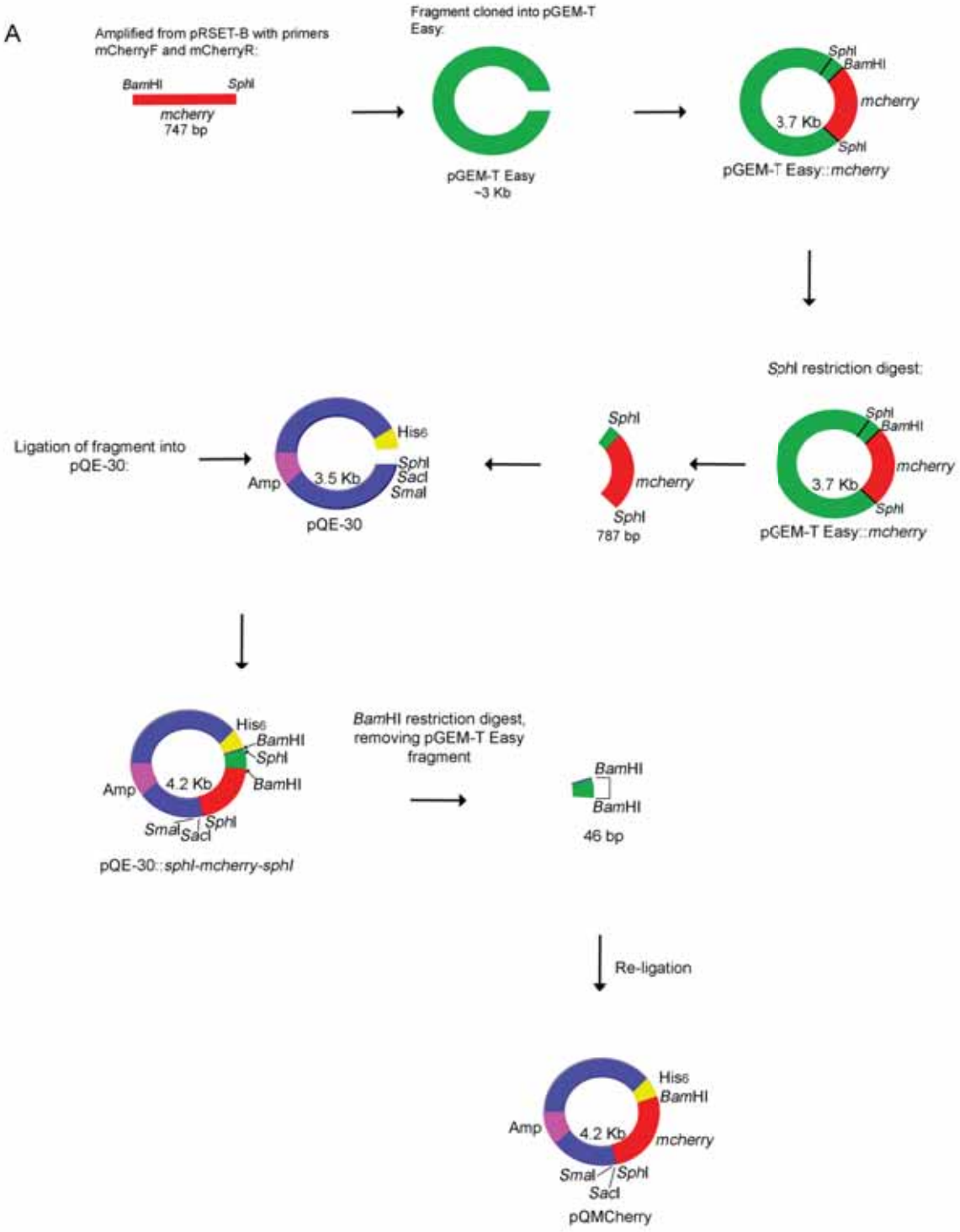
*Sph*I only, making use of a flanking *Sph*I restriction site in pGEM-T Easy in order to release the *Sph*I-*mCherry*-*Sph*I fragment, which was then cloned into pQE-30 (Figure 5.1A). The orientation of the fragment was determined with restriction digest analysis in order to ensure that *mcherry* was cloned in frame with the His_6-tag and the start codon (Figure 5.1A). Once verified, the construct was digested with *Bam*HI, in order to excise the 46 bp *Bam*HI-*Bam*HI fragment that remained (Figure 5.1A). This left only *mcherry* cloned into pQE-30 to create the construct pQMCherry (Table 2.2). To clone wzz_{SF} into pQMCherry, pRMCD30 was digested with *Sac*I and *Sma*I to release wzz_{SF}. This fragment was ligated into a similarly digested pQMCherry, in order to produce the construct pQMCherry-Wzz_{SF} (Figure 5.1B, Table 2.2). The sequences of *mcherry* cloned in pGEM-T Easy and pQMCherry-Wzz_{SF} were verified with DNA sequencing (Figure 5.2).

5.3 COMPLEMENTATION BY mCHERRY-WZZ_{SF}

The ability of the mCherry-tagged Wzz_{SF} protein ability to complement a *wzz* mutant was investigated. Strain RMA2741, (*S. flexneri* F'$lacI^q$) was electroporated with plasmids pQMCherry, and pQMCherry-Wzz_{SF}. These strains along with RMA2741 harbouring either pQE-30 or pRMCD30 (encoding His_6-Wzz_{SF}) were grown, induced with IPTG, and LPS samples were prepared and subjected to SDS-PAGE and silver staining (section 2.11). Analysis of the LPS profiles indicated that strain RMA2741 (pQE-30) (Figure 5.3A, lane 1) had LPS with random length Oag chains, and that wild-type Oag modal chain length is restored in RMA2741 (pRMCD30), as expected (Figure 5.3A, lane 2). RMA2741 (pQMCherry) had LPS with random length Oag chains (Figure 5.3A, lane 3). However, RMA2741 (pQMCherry-Wzz_{SF}) had wild-type Oag chain modal length of 11-17 RUs (Figure 5.3A, lane 4) and the level of modal length regulation was comparable to wild-type. These data indicated that as mCherry-Wzz_{SF} was fully functional, the mCherry tag did not detectably interfere with Wzz_{SF} function. Western immunoblotting was also conducted on

Figure 5.1 Construction of pQMCherry and pQMCherry-Wzz$_{SF}$

Plasmids pQMCherry and pQMCherry-Wzz$_{SF}$ were constructed as described in section 5.2. Briefly, primers incorporating *Bam*HI forward and *Sph*I reverse primers were used to amplify *mcherry* from pRSET-B (A) (Shaner *et al*., 2004), and cloned into pGEM-T Easy (Promega). This fragment was then excised from pGEM-T Easy with *Sph*I restriction digestion, utilising the pGEM-T Easy *Sph*I site located 46-bp prior to the *Bam*HI-*mcherry*-*Sph*I insertion. This *Sph*I-*mcherry*-*Sph*I fragment was cloned into pQE-30 (Qiagen), and the 46-bp fragment was excised from this construct with a *Bam*HI restriction digest, utilising the pQE-30 *Bam*HI site and the incorporated *Bam*HI site in the fragment, generating pQMCherry. B) The 1.1 Kb *wzz$_{SF}$* fragment from pRMCD30 was excised with a *Sac*I and *Sma*I restriction digest, and was ligated into a similarly digested pQMCherry to generate pQMCherry-Wzz$_{SF}$. The *mcherry* fragment is illustrated in red, pGEM-T Easy in green, pQE-30 in blue, and the His$_6$ tag is illustrated in yellow. Amp resistance conferred by pQE-30 is indicated in magenta and *wzz$_{SF}$* is illustrated in purple.


A
Amplified from pRSET-B with primers mCherryF and mCherryR:
BamHI
SphI
mcherry
747 bp
Fragment cloned into pGEM-T Easy:
pGEM-T Easy
~3 Kb
SphI
BamHI
mcherry
3.7 Kb
SphI
pGEM-T Easy::mcherry
SphI restriction digest:
SphI
787 bp
Ligation of fragment into pQE-30:
His6
SphI
SacI
SmaI
Amp
3.5 Kb
pQE-30
BamHI restriction digest, removing pGEM-T Easy fragment
BamHI
BamHI
46 bp
Amp
4.2 Kb
pQE-30::sphI-mcherry-sphI
Re-ligation
pQMCherry

B

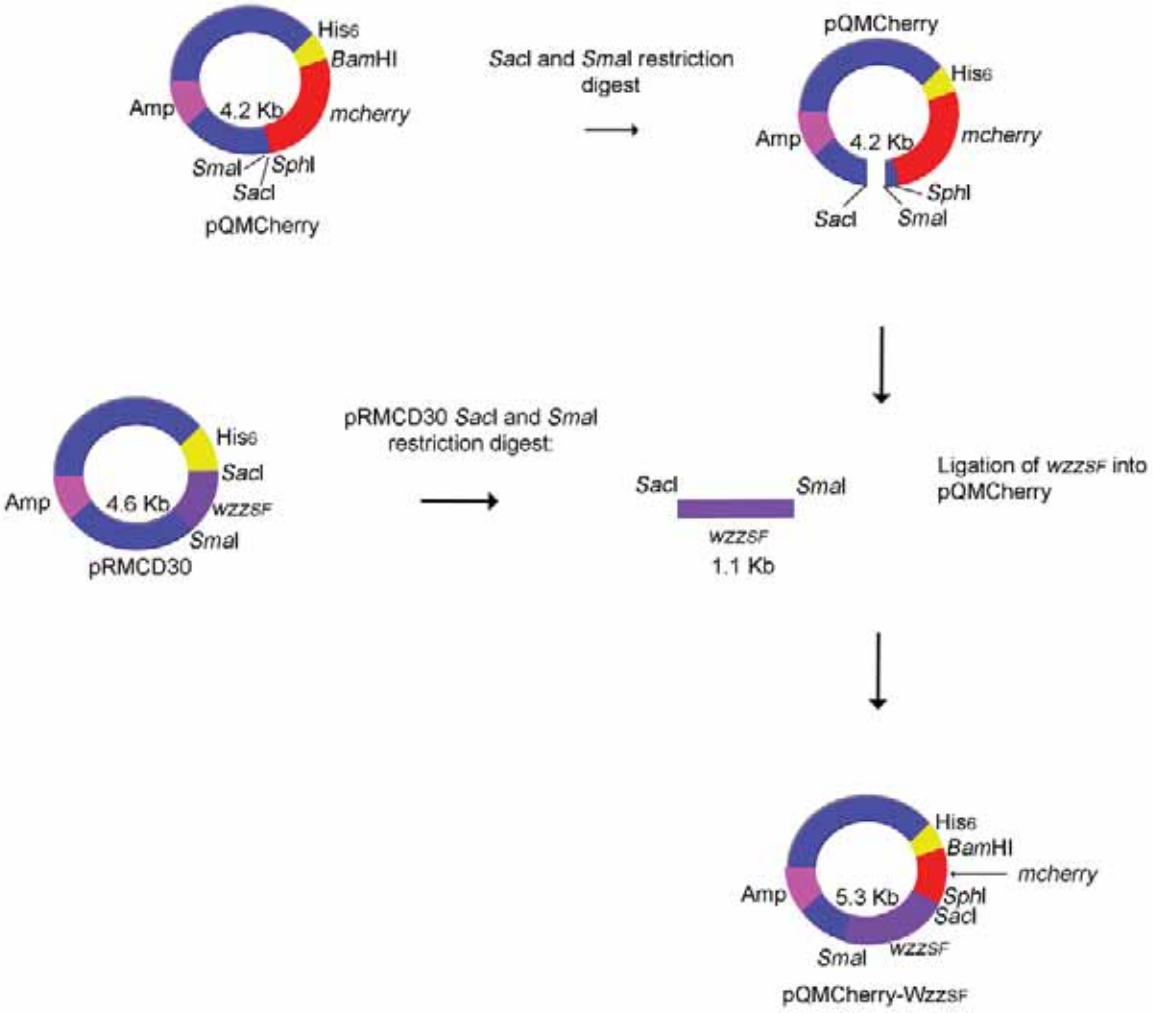
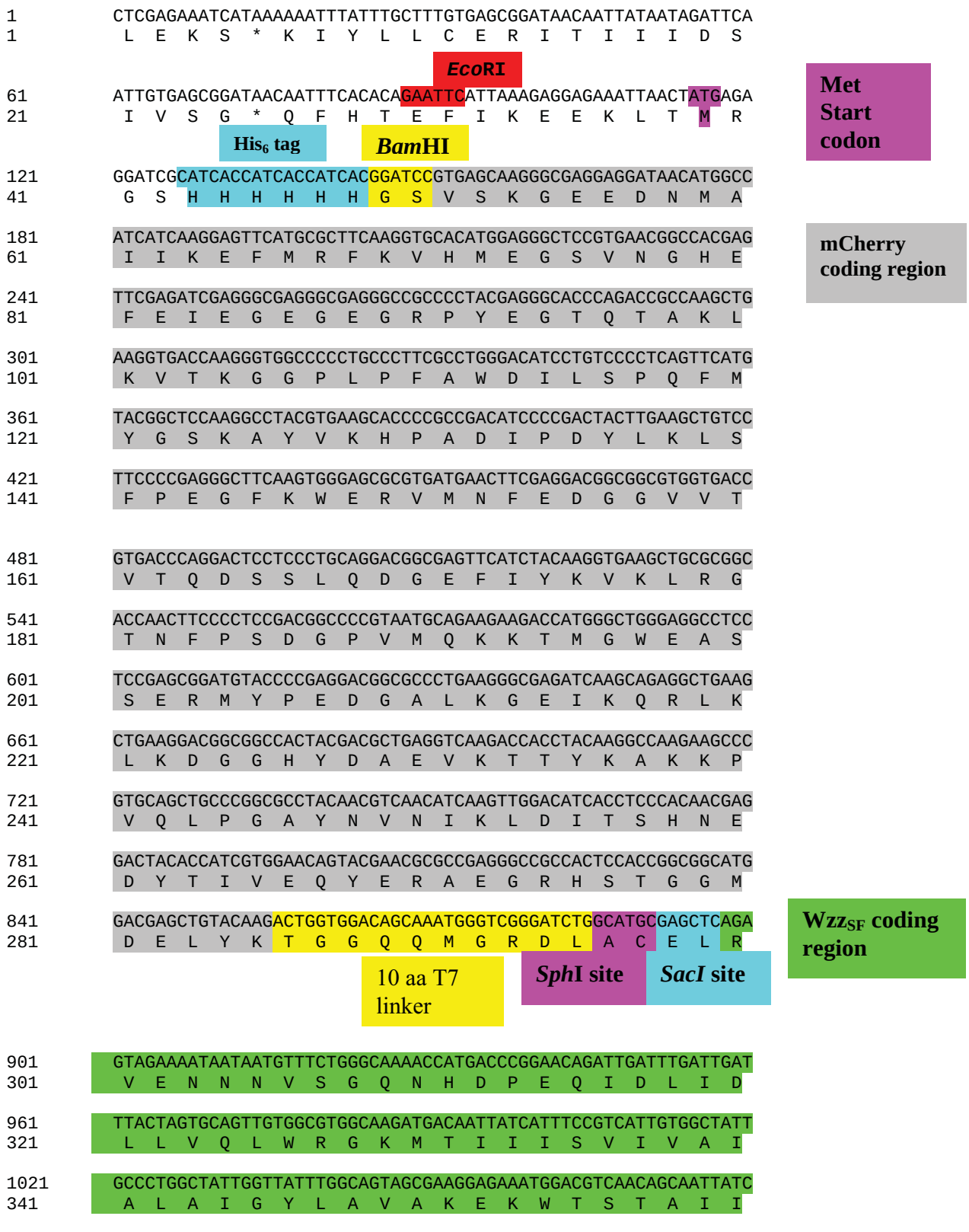

His6
BamHI
Amp
4.2 Kb
mcherry
SmaI
SphI
SacI
pQMCherry
SacI and SmaI restriction digest
pQMCherry
His6
Amp
4.2 Kb
mcherry
SphI
SacI
SmaI
His6
SacI
Amp
4.6 Kb
wzzSF
SmaI
pRMCD30
pRMCD30 SacI and SmaI restriction digest:
SacI
SmaI
wzzSF
1.1 Kb
Ligation of wzzSF into pQMCherry
His6
BamHI
mcherry
Amp
5.3 Kb
SphI
SacI
SmaI
wzzSF
pQMCherry-WzzSF

Figure 5.2 The DNA sequence and predicted amino acid sequence of the mCherry-Wzz_{SF} fragment in pQMCherry-Wzz_{SF}

The DNA sequence of the mCherry-Wzz_{SF} insert in pQMCherry-Wzz_{SF} (*mcherry* genbank accession number AY678264.1, *wzz*$_{SF}$ X71970.1), confirmed by DNA sequencing. The start codon is indicated in magenta, the *Bam*HI site is indicated in yellow, the His_6-tag is highlighted in teal, *mcherry* is highlighted in gray, the linker between mCherry and Wzz is highlighted in yellow, the *Sph*I site is indicated in magenta, the *Sac*I site is highlighted in teal, and *wzz*$_{SF}$ is highlighted in green, while the *Sma*I site is indicated in yellow.

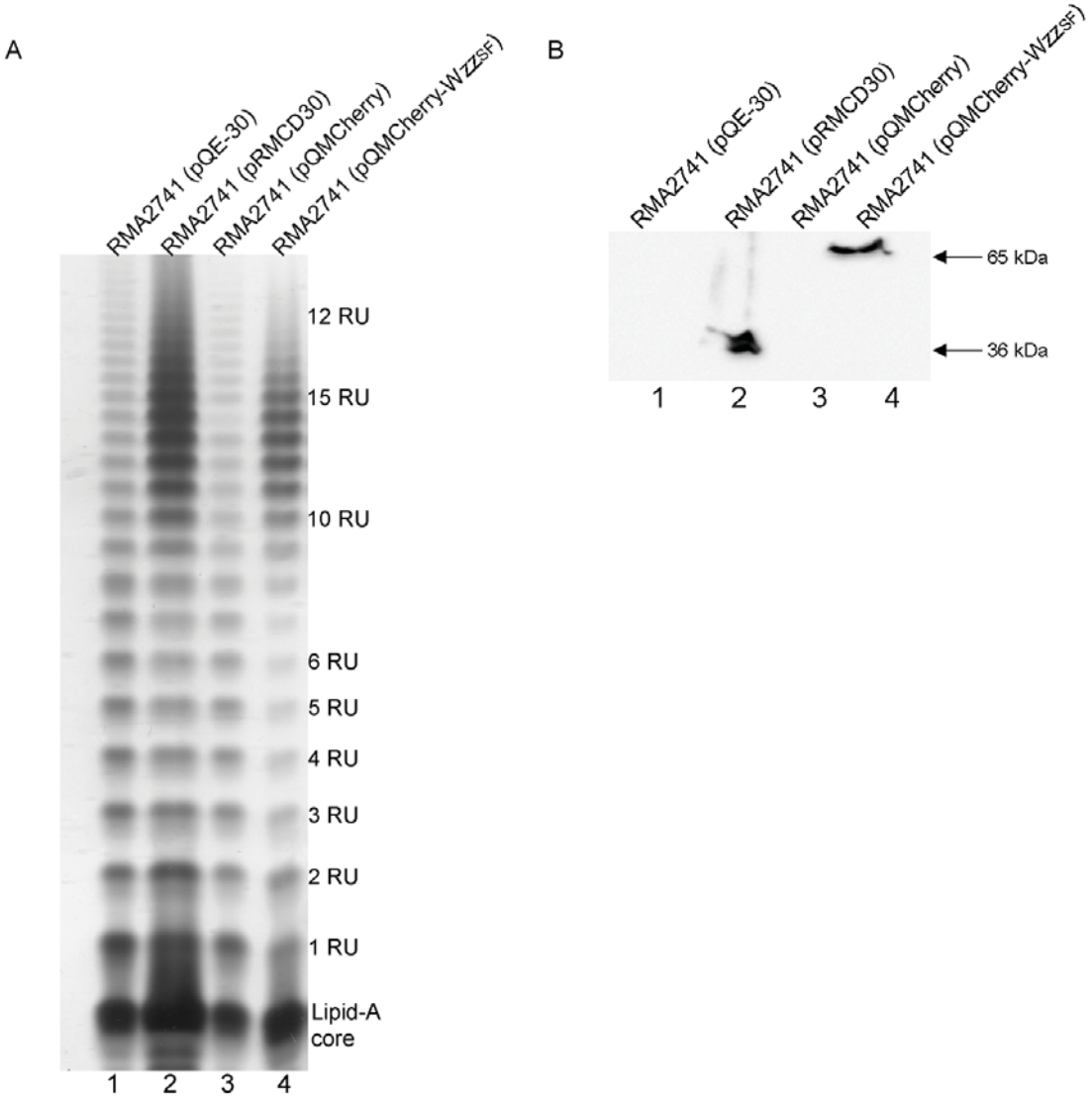

1 CTCGAGAAATCATAAAAAATTTATTTGCTTTGTGAGCGGATAACAATTATAATAGATTCA
1 L E K S * K I Y L L C E R I T I I I D S
EcoRI
61 ATTGTGAGCGGATAACAATTTCACACAGAATTCATTAAAGAGGAGAAATTAACTATGAGA
21 I V S G * Q F H T E F I K E E K L T M R
Met Start codon
His6 tag
BamHI
121 GGATCGCATCACCATCACCATCACGGATCCGTGAGCAAGGGCGAGGAGGATAACATGGCC
41 G S H H H H H H G S V S K G E E D N M A
181 ATCATCAAGGAGTTCATGCGCTTCAAGGTGCACATGGAGGGCTCCGTGAACGGCCACGAG
61 I I K E F M R F K V H M E G S V N G H E
mCherry coding region
241 TTCGAGATCGAGGGCGAGGGCGAGGGCCGCCCCTACGAGGGCACCCAGACCGCCAAGCTG
81 F E I E G E G E G R P Y E G T Q T A K L
301 AAGGTGACCAAGGGTGGCCCCCTGCCCTTCGCCTGGGACATCCTGTCCCCTCAGTTCATG
101 K V T K G G P L P F A W D I L S P Q F M
361 TACGGCTCCAAGGCCTACGTGAAGCACCCCGCCGACATCCCCGACTACTTGAAGCTGTCC
121 Y G S K A Y V K H P A D I P D Y L K L S
421 TTCCCCGAGGGCTTCAAGTGGGAGCGCGTGATGAACTTCGAGGACGGCGGCGTGGTGACC
141 F P E G F K W E R V M N F E D G G V V T
481 GTGACCCAGGACTCCTCCCTGCAGGACGGCGAGTTCATCTACAAGGTGAAGCTGCGCGGC
161 V T Q D S S L Q D G E F I Y K V K L R G
541 ACCAACTTCCCCTCCGACGGCCCCGTAATGCAGAAGAAGACCATGGGCTGGGAGGCCTCC
181 T N F P S D G P V M Q K K T M G W E A S
601 TCCGAGCGGATGTACCCCGAGGACGGCGCCCTGAAGGGCGAGATCAAGCAGAGGCTGAAG
201 S E R M Y P E D G A L K G E I K Q R L K
661 CTGAAGGACGGCGGCCACTACGACGCTGAGGTCAAGACCACCTACAAGGCCAAGAAGCCC
221 L K D G G H Y D A E V K T T Y K A K K P
721 GTGCAGCTGCCCGGCGCCTACAACGTCAACATCAAGTTGGACATCACCTCCCACAACGAG
241 V Q L P G A Y N V N I K L D I T S H N E
781 GACTACACCATCGTGGAACAGTACGAACGCGCCGAGGGCCGCCACTCCACCGGCGGCATG
261 D Y T I V E Q Y E R A E G R H S T G G M
841 GACGAGCTGTACAAGACTGGTGGACAGCAAATGGGTCGGGATCTGGCATGCGAGCTCAGA
281 D E L Y K T G G Q Q M G R D L A C E L R
Wzz$_{SF}$ coding region
10 aa T7 linker
SphI site
SacI site
901 GTAGAAAATAATAATGTTTCTGGGCAAAACCATGACCCGGAACAGATTGATTTGATTGAT
301 V E N N N V S G Q N H D P E Q I D L I D
961 TTACTAGTGCAGTTGTGGCGTGGCAAGATGACAATTATCATTTCCGTCATTGTGGCTATT
321 L L V Q L W R G K M T I I I S V I V A I
1021 GCCCTGGCTATTGGTTATTTGGCAGTAGCGAAGGAGAAATGGACGTCAACAGCAATTATC
341 A L A I G Y L A V A K E K W T S T A I I

```
1081    ACTCAGCCCGACGTGGGGCAAATTGCTGGCTATAACAATGCCATGAATGTTATCTATGGT
361      T  Q  P  D  V  G  Q  I  A  G  Y  N  N  A  M  N  V  I  Y  G

1141    CAGGCTGCACCGAAAGTATCGGATTTGCAGGAGACGTTAATTGGTCGCTTCAGTTCTGCC
381      Q  A  A  P  K  V  S  D  L  Q  E  T  L  I  G  R  F  S  S  A

1201    TTCTCTGCATTAGCAGAAACGCTGGATAATCAGGAAGAGCCAGAAAAACTTACCATCGAA
401      F  S  A  L  A  E  T  L  D  N  Q  E  E  P  E  K  L  T  I  E

1261    CCTTCTGTTAAGAACCAGCAATTACCATTGACTGTTTCTTATGTTGGGCAAACTGCAGAG
421      P  S  V  K  N  Q  Q  L  P  L  T  V  S  Y  V  G  Q  T  A  E

1321    GGCGCACAAATGAAGTTGGCCCAATACATTCAGCAAGTTGATGATAAAGTGAATCAAGAG
441      G  A  Q  M  K  L  A  Q  Y  I  Q  Q  V  D  D  K  V  N  Q  E

1381    CTAGAAAAGGATCTCAAGGACAACATTGCTCTGGGACGGAAAAACTTGCAGGACTCTTTA
461      L  E  K  D  L  K  D  N  I  A  L  G  R  K  N  L  Q  D  S  L

1441    AGAACCCAGGAAGTGGTCGCGCAGGAGCAGAAAGATCTGCGTATCCGTCAGATTCAGGAA
481      R  T  Q  E  V  V  A  Q  E  Q  K  D  L  R  I  R  Q  I  Q  E

1501    GCGTTGCAGTATGCGAATCAGGCGCAGGTGACAAAGCCACAGGTTCAGCAGACTGAAGAT
501      A  L  Q  Y  A  N  Q  A  Q  V  T  K  P  Q  V  Q  Q  T  E  D

1561    GTGACGCAAGATACGTTGTTCCTTCTAGGGAGCGAAGCGCTGGAGTCGATGATTAAGCAT
521      V  T  Q  D  T  L  F  L  L  G  S  E  A  L  E  S  M  I  K  H

1621    GAAGCGACTCGTCCGTTGGTGTTCTCACCAAACTACTATCAGACACGTCAAAACCTGTTG
541      E  A  T  R  P  L  V  F  S  P  N  Y  Y  Q  T  R  Q  N  L  L

1681    GATATTGAAAAATTAAAGTTTGATGATCTTGATATTCATGCTTACCGCTATGTGATGAAA
561      D  I  E  K  L  K  F  D  D  L  D  I  H  A  Y  R  Y  V  M  K

1741    CCGACGTTACCTATTCGTCGCGATAGTCCGAAAAAGGCAATCACCTTGATTCTGGCAGTG
581      P  T  L  P  I  R  R  D  S  P  K  K  A  I  T  L  I  L  A  V

1801    CTTCTGGGCGGCATGGTTGGCGCGGGGATTGTGTTGGGGCGTAACGCTCTGCGTAATTAC
601      L  L  G  G  M  V  G  A  G  I  V  L  G  R  N  A  L  R  N  Y

1861    AACGCGAAGTAATATTATTGTGCATTTAAGAGAAACGGGCAGGGTGGTGACACCATGCCC
621       N  A  K  *  Y  Y  C  A  F  K  R  N  G  Q  G  G  D  T  M  P

1921    GTTTTTTTTGCCGGATGCGATGCTGGCGCATCTTATCCGGCCTACGTGTGTTGAGATAAT
641      V  F  F  A  G  C  D  A  G  A  S  Y  P  A  Y  V  C  *  D  N

1981    GTGTAGGCACGATAAGTTTGCGCATCGGGCAATGGCTCCGGGTGTGACAACAACATCACA
661      V  *  A  R  *  V  C  A  S  G  N  G  S  G  C  D  N  N  I  T

2041    CCTGCTCCCCGGGTCGACCTGCAGCCAAGCTTAATTAGCTGAGCTTGGACTCCTGTTGAT
681      P  A  P  R  V  D  L  Q  P  S  L  I  S  *  A  W  T  P  V  D
```

***Sma*I site**

Figure 5.3 Protein detection and LPS Oag modal length conferred by pQMCherry-Wzz_{SF}

S. flexneri RMA2741 strains harbouring pQE-30-based plasmids were grown in LB + Amp, induced with 0.5 mM IPTG for 1.5 h (section 2.10.3), and LPS and whole cell lysate samples were prepared for silver staining (section 2.11) and Western immunoblotting (section 2.10.1). Samples were subjected to SDS-PAGE 15%. A) Lanes are as follows: 1) RMA2741 (pQE-30), 2) RMA2741 (pRMCD30), 3) RMA2741 (pQMCherry), and 4) RMA2741 (pQMCherry-Wzz_{SF}). Each lane contained approximately 1.3 x 10^8 cells. The lipid A-core and number of repeat units (RUs) are indicated on the right. B) Western samples were detected with affinity purified polyclonal anti-Wzz_{SF}, at a dilution of 1:1000. The lanes are as follows: 1) RMA2741 (pQE-30), 2) RMA2741 (pRMCD30), 3) RMA2741 (pQMCherry), and 4) RMA2741 (pQMCherry-Wzz_{SF}). Prestained Benchmark protein marker was used to determine protein sizes (Invitrogen). Each lane contains approximately 2 x 10^8 cells.


A
RMA2741 (pQE-30)
RMA2741 (pRMCD30)
RMA2741 (pQMCherry)
RMA2741 (pQMCherry-WZZSF)
12 RU
15 RU
10 RU
6 RU
5 RU
4 RU
3 RU
2 RU
1 RU
Lipid-A core
1 2 3 4
B
RMA2741 (pQE-30)
RMA2741 (pRMCD30)
RMA2741 (pQMCherry)
RMA2741 (pQMCherry-WZZSF)
65 kDa
36 kDa
1 2 3 4

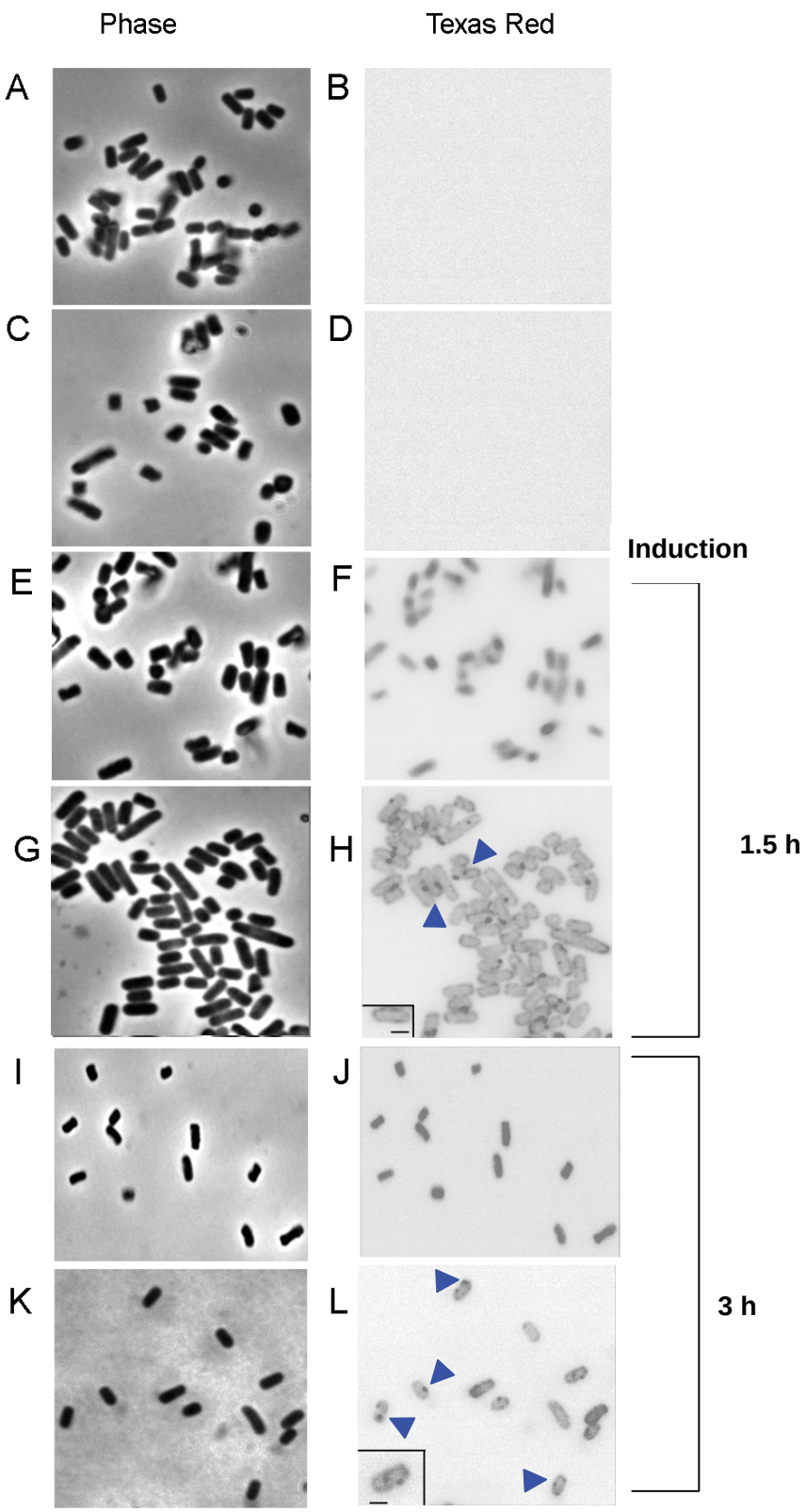
mCherry-Wzz_{SF} using anti-Wzz_{SF} polyclonal antibodies; strains were grown, induced and whole cell lysates were prepared. The resulting immunoblot showed that when pQE-30 and pQMCherry were expressed in RMA2741, no Wzz-related protein was detected, as expected (Figure 5.3B lanes 1 and 2, respectively), and that RMA2741 (pRMCD30) produced Wzz_{SF} (Figure 5.3B, lane 3). RMA2741 (pQMCherry-Wzz_{SF}) had a band of ~66 kDa (Figure 5.3B, lane 4), which is the size expected for the mCherry-Wzz_{SF} protein. The immunoblot also showed that no breakdown products and/or truncated forms of mCherry-Wzz_{SF} could be detected.

5.4 mCHERRY-WZZ LOCALISATION

To visualise intracellular Wzz_{SF} localisation, RMA2741 derivatives harbouring pQE-30, pQMCherry and pQMCherry-Wzz were prepared for microscopy. The bacteria were cultured and induced with amounts of IPTG (section 2.10.3) ranging from 0.2 mM to 1 mM, fixed with formaldehyde and mounted on slides as described in Section 2.14. Bacteria were visualised via phase contrast epi-fluorescence microscopy and, and fluorescence was detected using the Texas Red filter (section 2.13). The data indicated that low concentrations of IPTG were insufficient to induce detectable fluorescence in RMA2741 (pQMCherry) or RMA2741 (pQMCherry-Wzz) (data not shown). However, when induced with 0.5 mM IPTG for 1.5 and 3 h (section 2.10.3.1), expression of fluorescent protein was detected in RMA2741 (pQMCherry) (Figure 5.4F and J). Similarly, strain RMA2741 (pQMCherry-Wzz) also exhibited fluorescence under these experimental conditions (Figure 5.4H and L), however, unlike the diffuse fluorescence detected in cells expressing mCherry alone, discreet subcellular regions of localised mCherry-Wzz_{SF} were detected, predominantly at the peripheral regions of the cell (Figure 5.4H). Furthermore, the localisation appeared more concentrated in cells which had undergone 3 h induction with IPTG (Figure 5.4 L). Bacterial cells from strains RMA2741 (pQE-30) and RMA2741 (pRMCD30) were visible under phase

Figure 5.4 Epi-fluorescent microscopy detection of mCherry-Wzz$_{SF}$ localisation

S. flexneri RMA2741 strains harbouring pQE-30-based plasmids were grown in LB + Amp, induced with 0.5 mM IPTG for 1.5 or 3 h, washed in 1 x PBS and fixed with 3.7% (w/v) formaldehyde (section 2.14). Bacteria were visualised with epi-fluorescence microscopy at 100 x magnification. Images A, C, E, G, I and K were taken in phase contrast, whilst B, D, F, H, J and K were visualised with the Texas Red filter set. Images are as follows: A) RMA2741 (pQE-30), B) RMA2741 (pQE-30) Texas Red, C) RMA2741 (pRMCD30), D) RMA2741 (pRMCD30), E) RMA2741 (pQMCherry), F) RMA2741 (pQMCherry), G) RMA2741 (pQMCherry-Wzz), H) RMA2741 (pQMCherry-Wzz), I) RMA2741 (pQMCherry), J) RMA2741 (pQMCherry), K) RMA2741 (pQMCherry-Wzz), L) RMA2741 (pQMCherry-Wzz). Strains from panels I) to L) were induced for 3 h, whilst A) to H) were induced for 1.5 h. Scale bars indicate approximately 1 μm. Individual clusters of mCherry-Wzz$_{SF}$ fluorescence are indicated with blue arrowheads in panels H and L. Images on the right have been converted to grayscale and contrast-inverted to enhance clarity.


Phase
Texas Red
A
B
C
D
E
F
G
H
I
J
K
L
Induction
1.5 h
3 h

contrast, but no observable fluorescence was detected (Figure 5.4A and B, 5.4 C and D respectively).

5.5 SUMMARY

In this chapter, the construction of the fluorescent fusion protein mCherry-Wzz_{SF} was undertaken, and subsequent protein functionality and detection was investigated. The findings indicated that mCherry-Wzz_{SF} was able to fully complement a *wzz* mutation in *S. flexneri*, and showed that mCherry-Wzz_{SF} could be detected by Western immunoblotting. Epi-fluorescence microscopy was used to visualise mCherry-Wzz_{SF} with the protein localising to the periphery of the cell.

CHAPTER SIX

STUDIES ON WZY FUSION PROTEINS

6.1 INTRODUCTION

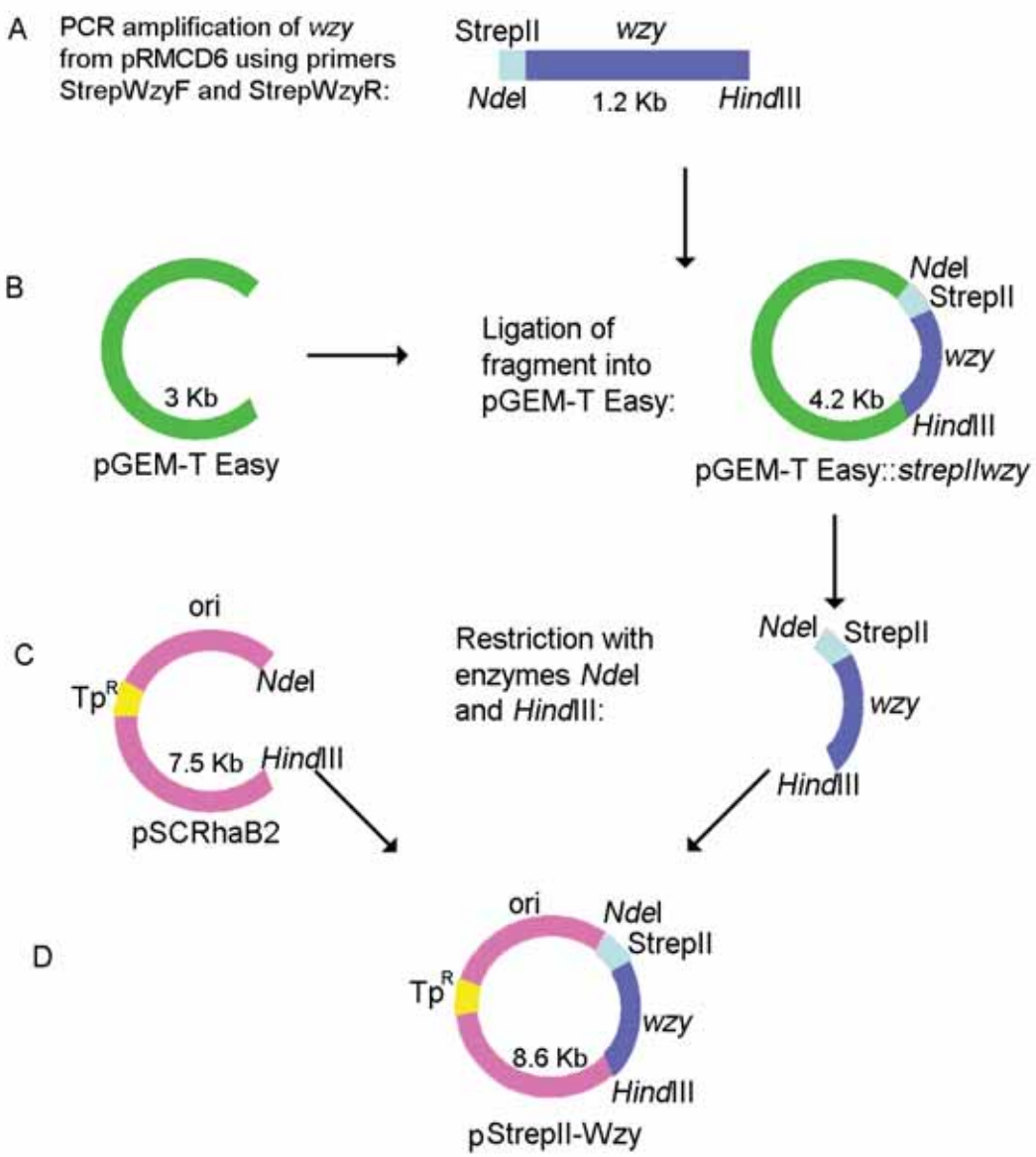
It has been previously proposed that Wzz proteins may interact with other Oag processing proteins (Marolda *et al.*, 2006). As described in Chapter 1, the exact mechanism of Wzz LPS Oag modal length control is unknown, although the formation of an Oag complex, possibly with the Wzy polymerase, WaaL ligase or the Wzx flippase, is considered to be a likely component of the modal length regulation mechanism (Marolda *et al.*, 2006). However, previous work with the Wzy polymerase has proven to be particularly difficult, as wild-type Wzy is poorly expressed and difficult to detect by Western immunoblotting (Daniels *et al.*, 1998). However, new reagents have considerably increased the sensitivity of Western immunoblotting in recent years (Kurien and Scofield, 2009). Therefore, the aim of this portion of the study was to construct a tagged Wzy protein capable of being detected using immunoblotting.

6.2 CONSTRUCTION OF pSTREPII-WZY$_{SF}$

Primers StrepWzyF and StrepWzyRev (Table 2.4) which incorporated the *Nde*I restriction site, the StrepII tag and the *Hind*III restriction site, respectively, were used to amplify the *wzy* reading frame and also incorporated a unique GGCGCC restriction site prior to *strepII* to function as a linker site (Figure 6.1). The primers were used to amplify a previously mutated *wzy* coding region, possessing three changes at codons 4, 9 and 23, from plasmid pRMCD6 (pBluescript::*wzy*, Daniels, 1999). Codons 4, 9 and 23 are rare codons present in the translation initiation region, and contribute to the low expression of wild-type

Figure 6.1 The construction of the StrepII-tagged Wzy plasmid

The plasmid pStrepII-Wzy was constructed as described in section 6.2. Briefly, forward and reverse primers incorporating *Nde*I (and an N-terminal StrepII tag) and *Hind*III restriction sites, respectively, were used to amplify *wzy* (also known as *rfc*, genbank accession number X71970.1) from plasmid pRMCD6 (A) (Daniels, 1999). The amplified fragment was then ligated into pGEM-T Easy (B), excised from the plasmid with an *Nde*I/*Hind*III double digest (C), and ligated into the similarly digested vector pSCRhaB2 (which possesses trimethoprim (Tp) resistance, a P_{rhaB} promoter, *ori*pBBR1 origin of replication and rhamnose induction) in order to generate pStrepII-Wzy (D). Indicated in this figure is *wzy* in blue, pGEM-T easy in green, the StrepII tag in sky blue, pSCRhaB2 in magenta and Tp^R in yellow.


A
PCR amplification of wzy from pRMCD6 using primers StrepWzyF and StrepWzyR:
StrepII
wzy
NdeI
1.2 Kb
HindIII
B
3 Kb
pGEM-T Easy
Ligation of fragment into pGEM-T Easy:
NdeI
StrepII
wzy
4.2 Kb
HindIII
pGEM-T Easy::strepIIwzy
C
ori
TpR
NdeI
7.5 Kb
HindIII
pSCRhaB2
Restriction with enzymes NdeI and HindIII:
NdeI
StrepII
wzy
HindIII
D
ori
NdeI
StrepII
TpR
wzy
8.6 Kb
HindIII
pStrepII-Wzy

Wzy_{SF} (Daniels *et al.*, 1998). These codons, ATA, ATA and AGA at 4, 9 and 23 were altered to ATT, ATT and CGT, respectively, and were shown to increase the Wzy expression, and its subsequent detection by radiolabelling (Daniels *et al.*, 1998). The amplified *wzy* was cloned into pGEM-T Easy, generating plasmid pGEM-T Easy::*strepII-wzy* (Figure 6.1), and was followed by restriction digest with enzymes *Nde*I and *Hind*III (Figure 6.1). This fragment was cloned into pSCRhaB2 (a vector comprising trimethoprim resistance, *ori*pBBR1 origin of replication and a P_{rhaB} promoter allowing for rhamnose induction), which had been similarly digested, resulting in the plasmid pStrepII-Wzy (Figure 6.1). This region was verified by DNA sequencing, and the sequence of the insert in pStrepII-Wzy is shown in Figure 6.2.

6.3 ANALYSIS OF LPS COMPLEMENTATION BY STREPII-WZY_{SF}

To determine if the predicted StrepII-Wzy_{SF} protein was functional, complementation analysis was undertaken. Strain RMM109 (Table 2.1) is a *S. flexneri wzy* mutant containing a frameshift mutation at position 9214 in the *wzy* gene sequence (*wzy/rfc* genbank accession number X71970.1). This mutation results in premature termination of Wzy synthesis and was therefore suitable for use as a host strain in this experiment (Morona *et al.*, 1994). Due to the lack of Oag polymerisation in the cell, only a single Oag repeat unit is ligated to the lipid A-core, resulting in a semi-rough phenotype. Control strains of PE638 and RMM109 harbouring plasmid pGP1-2 and pBluescript, or the control plasmid pRMCD6 (pBluescript encoding *wzy*) were grown and LPS samples were subjected to SDS-PAGE and silver staining. The resulting LPS profiles showed that PE638 (pGP1-2/pBluescript) had smooth LPS as expected, and RMM109 (pGP1-1/pBluescript) had SR- LPS, of only one Oag repeat unit (Figure 6.3B, lanes 1 and 2 respectively). However, RMM109 (pGP1-2/pRMCD6) had smooth LPS, and also exhibited Oag modal chain length if 11-17 RUs as did PE638 harbouring pGP1-2/pBluescript (Figure 6.3B, lane 3). This indicated that plasmid pRMCD6 completely restored the *wzy* defect in RMM109, as was previously shown (Daniels, 1999).

Start codon

StrepII

*Sfo*I

```
1          CATATGAGGAGAATATACAAAATGGCTAGCTGGAGCCACCCGCAGTTCGAAAAAGGCGCC
1           H  M  R  R  I  Y  K  M  A  S  W  S  H  P  Q  F  E  K  G  A

61         AATAATATTAATAAAATTTTTATTACATTTTTATGTATTGAACTGATTATTGGTGGTGGT   wzy_SF
21          N  N  I  N  K  I  F  I  T  F  L  C  I  E  L  I  I  G  G  G

121        GGACGTTTACTGGAGCCATTGGGAATATTCCCTTTGCGATATTTATTATTTGTATTTAGT
41          G  R  L  L  E  P  L  G  I  F  P  L  R  Y  L  L  F  V  F  S

181        TTTATACTTTTAATTTTTAATTTAGTTACATTCAATTTTTCAATCACCCAAAAATGTGTC
61          F  I  L  L  I  F  N  L  V  T  F  N  F  S  I  T  Q  K  C  V

241        AGTCTTTTTATATGGTTGCTTTTATTTCCTTTTTATGGCTTCTTTGTCGGCTTATTAGCT
81          S  L  F  I  W  L  L  L  F  P  F  Y  G  F  F  V  G  L  L  A

301        GGTAATAAAATAAATGATATACTGTTTGATGTGCAACCATACCTTTTTATGCTGTCACTT
101         G  N  K  I  N  D  I  L  F  D  V  Q  P  Y  L  F  M  L  S  L

361        ATATATCTATTTACACTAAGATATACTTTAAAAGTATTTTCATGTGAGATTTTTATTAAA
121         I  Y  L  F  T  L  R  Y  T  L  K  V  F  S  C  E  I  F  I  K

421        ATAGTTAATGCATTTGCATTATATGGATCACTGTTATATATTTCATACATAATTTTGTTG
141         I  V  N  A  F  A  L  Y  G  S  L  L  Y  I  S  Y  I  I  L  L

481        AATTTCGGTTTGTTAAATTTTAATTTAATTTATGAACACTTATCATTGACTAGCGAGTTC
161         N  F  G  L  L  N  F  N  L  I  Y  E  H  L  S  L  T  S  E  F

541        TTTTTTCGTCCCGATGGGGCTTTTTTTTCCAAATCCTTCTACTTTTTTGGTGTCGGTGCG
181         F  F  R  P  D  G  A  F  F  S  K  S  F  Y  F  F  G  V  G  A

601        ATTATCAGTTTTGTCGACAAAAAATATTTAAAATGTCTCATAATAGTGCTTGCGATATTA
201         I  I  S  F  V  D  K  K  Y  L  K  C  L  I  I  V  L  A  I  L

661        TTGACAGAATCAAGAGGTGTATTACTTTTTACAACATTATCACTGTTATTAGCCAGTTTT
221         L  T  E  S  R  G  V  L  L  F  T  T  L  S  L  L  L  A  S  F

721        AAATTACATAAGCTATATTTAAATACTATTATAATAATATTGGGCAGCGTTCTATTTATA
241         K  L  H  K  L  Y  L  N  T  I  I  I  I  L  G  S  V  L  F  I

781        ATTATGCTTTACATGGTCGGATCACGCAGTGAAGATTCTGACTCTGTTAGATTTAATGAT
261         I  M  L  Y  M  V  G  S  R  S  E  D  S  D  S  V  R  F  N  D

841        TTATATTTTTATTATAAAAATGTTGATTTAGCGACGTTCTTGTTTGGAAGAGGATTTGGT
281         L  Y  F  Y  Y  K  N  V  D  L  A  T  F  L  F  G  R  G  F  G

901        TCATTTATATTAGATCGATTAAGGATTGAAATAGTACCTCTTGAGATACTTCAGAAAACA
301         S  F  I  L  D  R  L  R  I  E  I  V  P  L  E  I  L  Q  K  T


961        GGCGTTATTGGTGTATTTATATCATTAGTTCCTATGTTGCTTATCTTTTTGAAAGGCTAT
321         G  V  I  G  V  F  I  S  L  V  P  M  L  L  I  F  L  K  G  Y
```

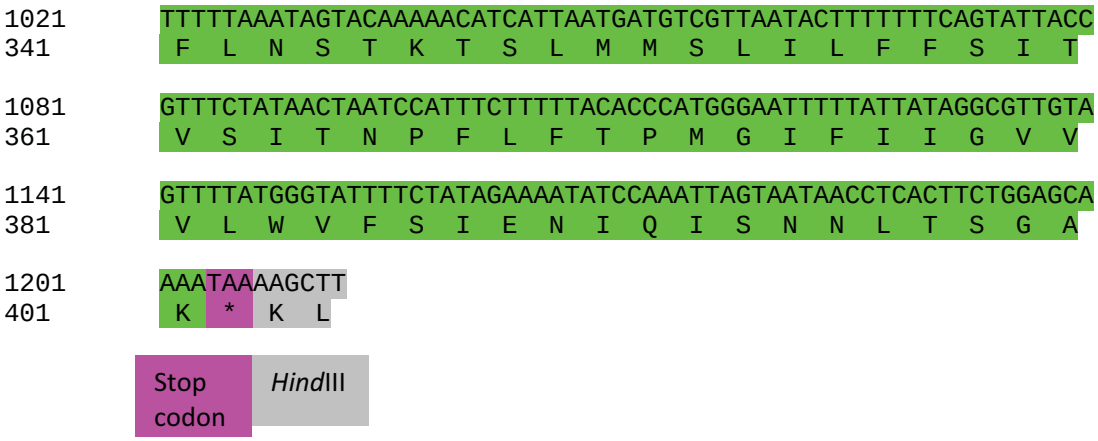


Figure 6.2 DNA sequence and predicted amino acid sequence of the StrepII-Wzy fragment

Presented above is the DNA sequence (determined by DNA sequencing) and corresponding amino acid sequence of the *strepII-wzy* fragment cloned into pSCRhaB2. Indicated are the start and stop codons in magenta, and the StrepII tag indicated in red and the linker region (which is a unique *Sfo*I site) in teal. Also indicated is the wzy_{SF} sequence in green (genbank accession number X71970.1), and the *Hind*III sequence in grey. The *wzy* substitutions in codons 4, 9 and 23 are indicated in blue text.

PE638 (pSCRhaB2), RMM109 (pStrepII-Wzy) were also grown, induced (section 2.10.5), and LPS samples were prepared and subjected to SDS-PAGE and silver staining. The resulting LPS profiles indicated that RMM109 (pStrepII-Wzy) was functional and able to partially complement the Wzy deficiency in RMM109, although SR-LPS was present (Figure 6.3A, lane 3, F). Its LPS however lacked any Oag modal chain length control i.e., it was not restored to the wild-type 11-17 RUs as was the case for pRMCD6, which indicated that StrepII-Wzy-mediated polymerization was not subject to modal chain length control by Wzz (Figure 6.3A, lane 3 and 6.B, lane 6). As expected, PE638 (pSCRhaB2) had smooth LPS with wild-type Oag modal chain length of 11-17 RUs (Figure 6.3A, lane 1 and 6.3B, lane 4) and RMM109 (pSCRhaB2) had SR-LPS (Figure 6.3A, lane 2 and 6.3B lane 5). These data indicate that the StrepII tag and additional residues at the NH_2 terminal end of Wzy interferes with Oag modal chain length control.

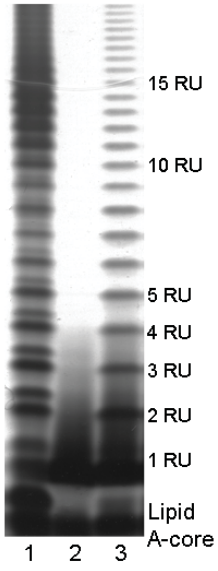
6.4 DETECTION OF STREP-WZY$_{SF}$ BY WESTERN IMMUNOBLOTTING

RMM109 strains harbouring either pSCRhaB2 or StrepII-Wzy were grown and induced with 0.2% (w/v) rhamnose for 2 h (section 2.10.5). Whole cell lysates were prepared as described previously (section 2.10.1), electrophoresed on a SDS 15% polyacrylamide gel, and were then subjected to Western immunoblotting to detect StrepII using an anti-StrepII antibody (Novagen/Merck) (Table 2.5). Samples prepared from a variety of growth temperatures were assessed following several different sample preparation methods, including an enhanced method for detecting membrane proteins (Abeyrathne and Lam, 2007). However, the detection of the StrepII tag proved to be non-reproducible and at or below the limits of detection for this assay and so could not reliably be used (data not shown). Although, as StrepII-Wzy exhibited function and complemented the *wzy* deficient strain, albeit with a modal chain length deficiency, an alternative approach was taken to enhance the detection of Wzy using this plasmid.

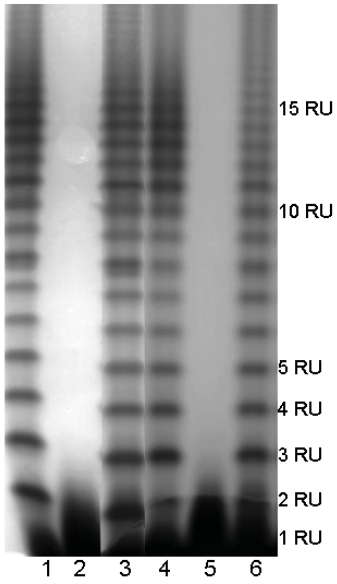
Figure 6.3 The LPS Oag modal length conferred by StrepII-Wzy

S. flexneri PE638 and RMM109 strains harbouring plasmids indicated below were grown in LB + Kan/Amp (for pBluescript-based plasmids), or in MH + Tp, and induced with rhamnose (1.5 h) for pSCRhaB2-based plasmids (section 2.10.5). 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) PE638 (pSCRhaB2), 2) RMM109 (pSCRhaB2) and 3) RMM109 (pStrepII-Wzy). B) 1) PE638 (pGP1-2)(pBluescript), 2) RMM109 (pGP1-2)(pBluescript), 3) RMM109 (pGP1-2)(pRMCD6), 4) PE638 (pSCRhaB2), 5) RMM109 (pSCRhaB2) and 6) RMM109 (pStrepII-Wzy). 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


15 RU
10 RU
5 RU
4 RU
3 RU
2 RU
1 RU
Lipid
A-core
1 2 3

B


15 RU
10 RU
5 RU
4 RU
3 RU
2 RU
1 RU
1 2 3 4 5 6

6.5 CONSTRUCTION OF GFP$^+$-TAGGED WZY PROTEINS

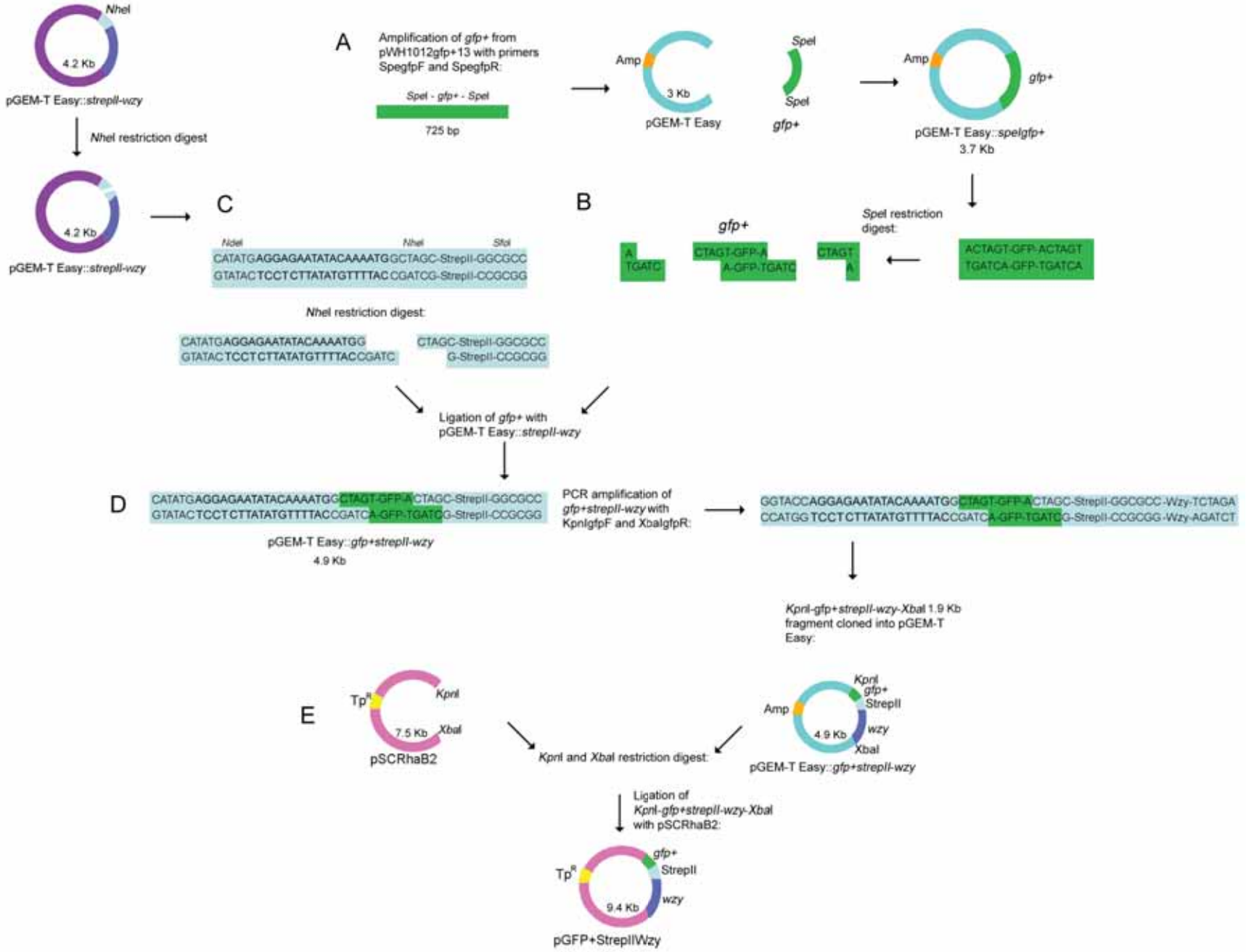
As mCherry-Wzz$_{SF}$ had full functionality (section 5.3), and its fluorescence could be detected via fluorescent microscopy, the construction of a complementary fluorescent-protein tagged Wzy was undertaken. Two approaches were devised; one to excise the StrepII-tag and allow the insertion of GFP$^+$, and the other to insert a GFP$^+$ tag prior to the StrepII tag. Both approaches took advantage of the unique *Sfo*I GGCGCC restriction site located at the 5' end of *strepII* (the coding sequence for the StrepII tag) (Figure 6.1). Both approaches also utilised plasmid pGEM-T Easy::*strepII-wzy* (section 6.2), as it contains a *Nhe*I site preceding *strepII* serving as a linker for the tag (Figure 6.2), and is a unique site in pGEM-T Easy. Multiple primers incorporating several restriction sites were designed to amplify *gfp*$^+$ from plasmid pWH1012gfp+13 (Scholz *et al.*, 2000).

6.5.1 CONSTRUCTION OF pGFP$^+$-STREPII-WZY

The objective for this cloning approach was to insert *gfp*$^+$ at the *Nhe*I site preceding *strepII*. Forward and reverse primers incorporating *Spe*I sites were designed (Table 2.4), which were used to amplify *gfp*$^+$ from plasmid pWH1012gfp+13 (Figure 6.4A). *Spe*I sites were selected because when *Spe*I and *Nhe*I sites are digested with their respective enzymes, complementary overhangs are generated (Figure 6.4B). This PCR amplified fragment was cloned into pGEM-T Easy and excised with *Spe*I, generating 5' overhangs. The construct pGEM-T::*strepII-wzy* was digested with *Nhe*I (Figure 6.4C) and was subsequently ligated with the *Spe*I-restricted *gfp*$^+$ fragment, resulting in plasmid pGEM-T Easy::*gfp*$^+$*strepII-wzy* (Figure 6.4D). Forward and reverse primers (KpnIgfpF and XbaIgfpR, Table 2.4) incorporating *Kpn*I and *Xba*I sites respectively, were designed to amplify the *gfp*$^+$-*strepII-wzy* region (Figure 6.4D). This PCR amplified region was cloned into pGEM-T Easy, then

Figure 6.4 Construction of pGFP^{+}-StrepII-Wzy

The plasmid pGFP^{+}-StrepII-Wzy was constructed as described in section 6.5.1. Briefly, primers SpegfpF and SpegfpR which incorporated *Spe*I restriction sites were used to amplify the *gfp*$^{+}$ coding region from template pWH1012Gfp+13 (Scholz *et al*., 2000). This amplified fragment was then cloned into pGEM-T Easy (B) and restricted with *Spe*I, generating a fragment with a 5' overhang. Simultaneous restriction digest with the plasmid pGEM-T Easy::*strepII-wzy* (section 6.2) using *Nhe*I (C) yielded a complimentary 5' overhang, and the *Spe*I-digested *gfp*$^{+}$ fragment was ligated with pGEM-T Easy::*strepII-wzy* to create the plasmid pGEM-T Easy::*strepII-gfp*$^{+}$*-wzy* (D). The entire *gfp*$^{+}$*-strepII-wzy* region was then amplified with new forward and reverse primers (KpnIgfpF and XbaIgfpR) which incorporated *Kpn*I and *Xba*I restriction sites respectively, and was cloned into pGEM-T Easy, creating plasmid pGEM-T Easy::*gfp*$^{+}$*-strepII-wzy*. Simultaneous restriction digest of this plasmid, and of pSCRhaB2 (E) was undertaken, and the *Kpn*I-*gfp*$^{+}$*-strepII-wzy-Xba*I fragment was ligated into pSCRhaB2 to generate the plasmid pGFP^{+}-StrepII-Wzy. The *wzy* fragment is illustrated in dark blue, *gfp*$^{+}$ is illustrated in green, pGEM-T Easy in cyan, pSCRhaB2 in magenta, and the StrepII tag is illustrated in sky blue. Amp resistance conferred by pGEM-T Easy is indicated in orange, and trimethoprim in yellow.


NheI
4.2 Kb
pGEM-T Easy::strepII-wzy
NheI restriction digest
4.2 Kb
pGEM-T Easy::strepII-wzy
A
Amplification of gfp+ from pWH1012gfp+13 with primers SpegfpF and SpegfpR:
SpeI - gfp+ - SpeI
725 bp
Amp
3 Kb
pGEM-T Easy
SpeI
SpeI
gfp+
Amp
gfp+
pGEM-T Easy::speIgfp+
3.7 Kb
B
SpeI restriction digest:
ACTAGT-GFP-ACTAGT
TGATCA-GFP-TGATCA
gfp+
A
TGATC
CTAGT-GFP-A
A-GFP-TGATC
CTAGT
A
C
NdeI
NheI
SfoI
CATATGAGGAGAATATACAAAATGGCTAGC-StrepII-GGCGCC
GTATACTCCTCTTATATGTTTTACCGATCG-StrepII-CCGCGG
NheI restriction digest:
CATATGAGGAGAATATACAAAATGG
GTATACTCCTCTTATATGTTTTACCGATC
CTAGC-StrepII-GGCGCC
G-StrepII-CCGCGG
Ligation of gfp+ with pGEM-T Easy::strepII-wzy
D
CATATGAGGAGAATATACAAAATGGCTAGT-GFP-ACTAGC-StrepII-GGCGCC
GTATACTCCTCTTATATGTTTTACCGATCA-GFP-TGATCG-StrepII-CCGCGG
pGEM-T Easy::gfp+strepII-wzy
4.9 Kb
PCR amplification of gfp+strepII-wzy with KpnIgfpF and XbaIgfpR:
GGTACCAGGAGAATATACAAAATGGCTAGT-GFP-ACTAGC-StrepII-GGCGCC-Wzy-TCTAGA
CCATGGTCCTCTTATATGTTTTACCGATCA-GFP-TGATCG-StrepII-CCGCGG-Wzy-AGATCT
KpnI-gfp+strepII-wzy-XbaI 1.9 Kb fragment cloned into pGEM-T Easy:
E
TpR
KpnI
7.5 Kb
XbaI
pSCRhaB2
KpnI and XbaI restriction digest:
Amp
KpnI
gfp+
StrepII
4.9 Kb
wzy
XbaI
pGEM-T Easy::gfp+strepII-wzy
Ligation of KpnI-gfp+strepII-wzy-XbaI with pSCRhaB2:
TpR
gfp+
StrepII
9.4 Kb
wzy
pGFP+StrepIIWzy

digested with *Kpn*I and *Xba*I, and ligated into the similarly digested with pSCRhaB2, generating plasmid pGFP$^+$-StrepII-Wzy (Figure 6.4E). This region was DNA sequenced to verify the cloning of the *gfp*$^+$-*strepII-wzy* region (Figure 6.5).

6.5.2 CONSTRUCTION OF pGFP$^+$-WZY

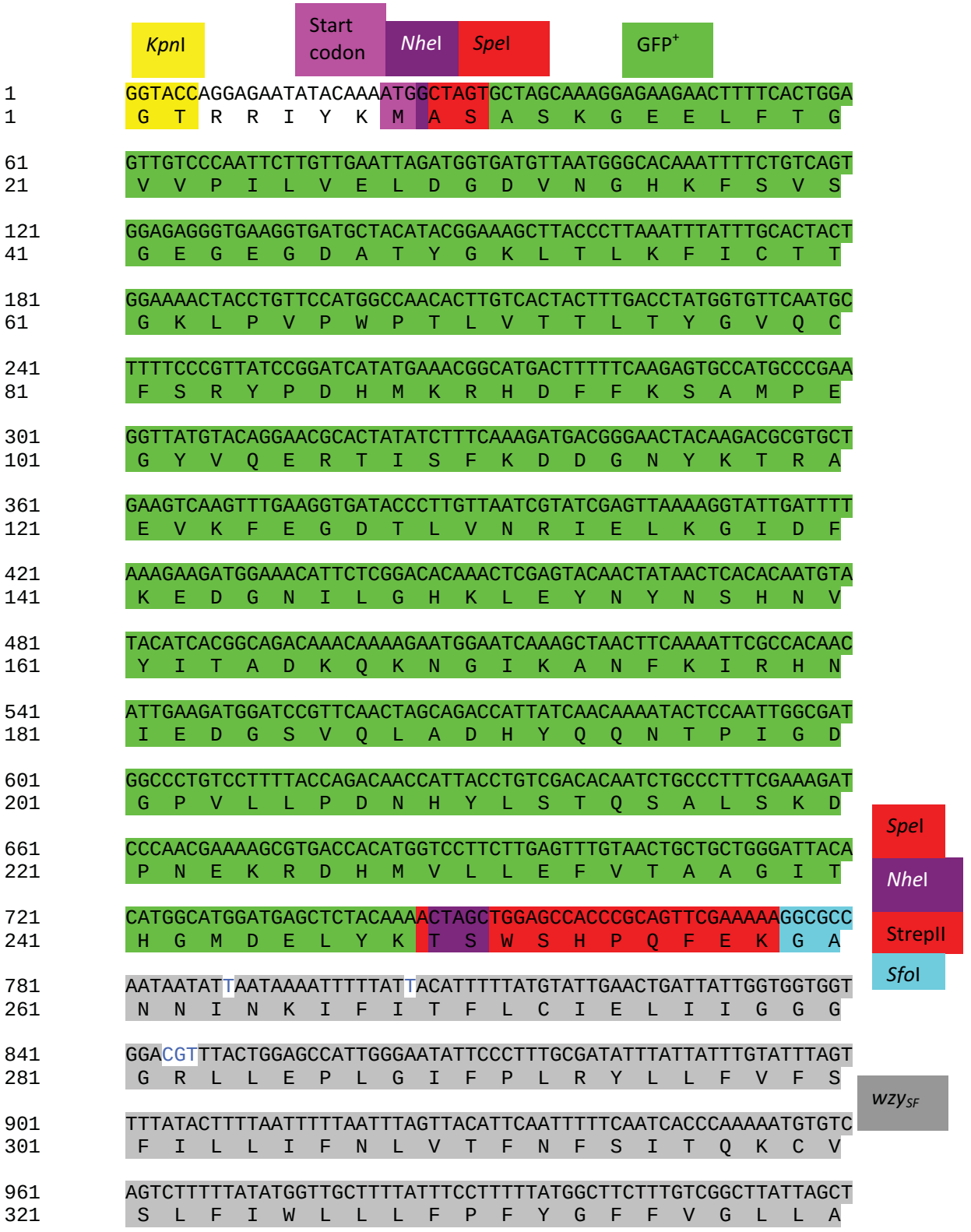
The objective for this cloning approach was to insert *gfp*$^+$ and excise the StrepII tag. Forward and reverse primers (SpegfpF and SfogfpR, Table 2.4) incorporating *Spe*I and *Sfo*I sites respectively, were designed to amplify *gfp*$^+$ from pWH1012gfp+13 (Scholz *et al.*, 2000) (Figure 6.6A). This was cloned into pGEM-T Easy and digested with *Sfo*I and *Spe*I, which resulted in a DNA fragment possessing a blunt end and a 5' overhang complimentary to overhangs of fragments digested with *Nhe*I respectively (Figure 6.6B). pGEM-T Easy::*strepII-wzy* was digested with enzymes *Nhe*I and *Sfo*I, also generating 5' overhangs and blunt ends, but unlike the previously constructed GFP$^+$-StrepII-Wzy protein, this double digest completely excised the StrepII tag (Figure 6.6B). Once excised, the *gfp*$^+$ fragment was cloned between *Nhe*I and *Sfo*I sites, which resulted in a hybrid site and the regeneration of the *Sfo*I restriction site (Figure 6.6C). Primers KpnIgfpF and XbaIgfpR incorporating *Kpn*I and *Xba*I sites were used to amplify this region and this PCR product was ligated into similarly digested pSCRhaB2 (Figure 6.6D), generating pGFP$^+$-Wzy (Figure 6.6E). DNA sequencing was undertaken to confirm the *gfp*$^+$-*wzy* sequence (Figure 6.7).

6.6 LPS PHENOTYPES OF GFP$^+$-TAGGED WZY PROTEINS

In order to assess the function of the GFP$^+$-tagged Wzy proteins, RMM109 cells were transformed with pGFP$^+$-StrepII-Wzy and pGFP$^+$-Wzy, and grown alongside RMM109 (pStrepII-Wzy) and the appropriate vector control. LPS samples were prepared from these cultures and subjected to SDS-PAGE and silver staining. The results showed that RMM109

Figure 6.5 DNA sequence and predicted amino acid sequence of the GFP$^+$-StrepII-Wzy fragment

Shown above is the DNA sequence, verified by DNA sequencing, and the corresponding amino acid sequence of the *gfp*$^+$*-Strep-wzy* sequence cloned into pSCRhaB2. The *Kpn*I site is indicated in yellow, the start and stop codons in magenta, the GFP$^+$ sequence in green, and the StrepII tag in red. The *Spe*I sites destroyed via the cloning are indicated in red, and the *Nhe*I sites are indicated in violet, and wzy_{SF} in grey. Also indicated is the *Xba*I site yellow. The *wzy* substitutions in codons 4, 9 and 23 are indicated in blue text.


KpnI
Start codon
NheI
SpeI
GFP+
1 GGTACCAGGAGAATATACAAAATGGCTAGTGCTAGCAAAGGAGAAGAACTTTTCACTGGA
1 G T R R I Y K M A S A S K G E E L F T G
61 GTTGTCCCAATTCTTGTTGAATTAGATGGTGATGTTAATGGGCACAAATTTTCTGTCAGT
21 V V P I L V E L D G D V N G H K F S V S
121 GGAGAGGGTGAAGGTGATGCTACATACGGAAAGCTTACCCTTAAATTTATTTGCACTACT
41 G E G E G D A T Y G K L T L K F I C T T
181 GGAAAACTACCTGTTCCATGGCCAACACTTGTCACTACTTTGACCTATGGTGTTCAATGC
61 G K L P V P W P T L V T T L T Y G V Q C
241 TTTTCCCGTTATCCGGATCATATGAAACGGCATGACTTTTTCAAGAGTGCCATGCCCGAA
81 F S R Y P D H M K R H D F F K S A M P E
301 GGTTATGTACAGGAACGCACTATATCTTTCAAAGATGACGGGAACTACAAGACGCGTGCT
101 G Y V Q E R T I S F K D D G N Y K T R A
361 GAAGTCAAGTTTGAAGGTGATACCCTTGTTAATCGTATCGAGTTAAAAGGTATTGATTTT
121 E V K F E G D T L V N R I E L K G I D F
421 AAAGAAGATGGAAACATTCTCGGACACAAACTCGAGTACAACTATAACTCACACAATGTA
141 K E D G N I L G H K L E Y N Y N S H N V
481 TACATCACGGCAGACAAACAAAAGAATGGAATCAAAGCTAACTTCAAAATTCGCCACAAC
161 Y I T A D K Q K N G I K A N F K I R H N
541 ATTGAAGATGGATCCGTTCAACTAGCAGACCATTATCAACAAAATACTCCAATTGGCGAT
181 I E D G S V Q L A D H Y Q Q N T P I G D
601 GGCCCTGTCCTTTTACCAGACAACCATTACCTGTCGACACAATCTGCCCTTTCGAAAGAT
201 G P V L L P D N H Y L S T Q S A L S K D
661 CCCAACGAAAAGCGTGACCACATGGTCCTTCTTGAGTTTGTAACTGCTGCTGGGATTACA
221 P N E K R D H M V L L E F V T A A G I T
721 CATGGCATGGATGAGCTCTACAAAACTAGCTGGAGCCACCCGCAGTTCGAAAAAGGCGCC
241 H G M D E L Y K T S W S H P Q F E K G A
781 AATAATATTAATAAAATTTTTATTACATTTTTATGTATTGAACTGATTATTGGTGGTGGT
261 N N I N K I F I T F L C I E L I I G G G
841 GGACGTTTACTGGAGCCATTGGGAATATTCCCTTTGCGATATTTATTATTTGTATTTAGT
281 G R L L E P L G I F P L R Y L L F V F S
901 TTTATACTTTTAATTTTTAATTTAGTTACATTCAATTTTTCAATCACCCAAAAATGTGTC
301 F I L L I F N L V T F N F S I T Q K C V
961 AGTCTTTTTATATGGTTGCTTTTATTTCCTTTTTATGGCTTCTTTGTCGGCTTATTAGCT
321 S L F I W L L L F P F Y G F F V G L L A
SpeI
NheI
StrepII
SfoI
wzySF

```
1021    GGTAATAAAATAAATGATATACTGTTTGATGTGCAACCATACCTTTTTATGCTGTCACTT
341      G  N  K  I  N  D  I  L  F  D  V  Q  P  Y  L  F  M  L  S  L

1081    ATATATCTATTTACACTAAGATATACTTTAAAAGTATTTTCATGTGAGATTTTTATTAAA
361      I  Y  L  F  T  L  R  Y  T  L  K  V  F  S  C  E  I  F  I  K

1141    ATAGTTAATGCATTTGCATTATATGGATCACTGTTATATATTTCATACATAATTTTGTTG
381      I  V  N  A  F  A  L  Y  G  S  L  L  Y  I  S  Y  I  I  L  L

1201    AATTTCGGTTTGTTAAATTTTAATTTAATTTATGAACACTTATCATTGACTAGCGAGTTC
401      N  F  G  L  L  N  F  N  L  I  Y  E  H  L  S  L  T  S  E  F

1261    TTTTTTCGTCCCGATGGGGCTTTTTTTTCCAAATCCTTCTACTTTTTTGGTGTCGGTGCG
421      F  F  R  P  D  G  A  F  F  S  K  S  F  Y  F  F  G  V  G  A

1321    ATTATCAGTTTTGTCGACAAAAAATATTTAAAATGTCTCATAATAGTGCTTGCGATATTA
441      I  I  S  F  V  D  K  K  Y  L  K  C  L  I  I  V  L  A  I  L

1381    TTGACAGAATCAAGAGGTGTATTACTTTTTACAACATTATCACTGTTATTAGCCAGTTTT
461      L  T  E  S  R  G  V  L  L  F  T  T  L  S  L  L  L  A  S  F

1441    AAATTACATAAGCTATATTTAAATACTATTATAATAATATTGGGCAGCGTTCTATTTATA
481      K  L  H  K  L  Y  L  N  T  I  I  I  I  L  G  S  V  L  F  I

1501    ATTATGCTTTACATGGTCGGATCACGCAGTGAAGATTCTGACTCTGTTAGATTTAATGAT
501      I  M  L  Y  M  V  G  S  R  S  E  D  S  D  S  V  R  F  N  D

1561    TTATATTTTTATTATAAAAATGTTGATTTAGCGACGTTCTTGTTTGGAAGAGGATTTGGT
521      L  Y  F  Y  Y  K  N  V  D  L  A  T  F  L  F  G  R  G  F  G

1621    TCATTTATATTAGATCGATTAAGGATTGAAATAGTACCTCTTGAGATACTTCAGAAAACA
541      S  F  I  L  D  R  L  R  I  E  I  V  P  L  E  I  L  Q  K  T

1681    GGCGTTATTGGTGTATTTATATCATTAGTTCCTATGTTGCTTATCTTTTTGAAAGGCTAT
561      G  V  I  G  V  F  I  S  L  V  P  M  L  L  I  F  L  K  G  Y

1741    TTTTTAAATAGTACAAAAACATCATTAATGATGTCGTTAATACTTTTTTTCAGTATTACC
581      F  L  N  S  T  K  T  S  L  M  M  S  L  I  L  F  F  S  I  T

1801    GTTTCTATAACTAATCCATTTCTTTTTACACCCATGGGAATTTTTATTATAGGCGTTGTA
601      V  S  I  T  N  P  F  L  F  T  P  M  G  I  F  I  I  G  V  V

1861    GTTTTATGGGTATTTTCTATAGAAAATATCCAAATTAGTAATAACCTCACTTCTGGAGCA
621      V  L  W  V  F  S  I  E  N  I  Q  I  S  N  N  L  T  S  G  A

1921    AAATAATCTAGA
641      K  *  S  R
```

Stop codon

*Xba*I

Figure 6.6 The construction of pGFP^{+}-Wzy$_{SF}$

The pGFP^{+}-Wzy$_{SF}$ plasmid was created as described in section 6.5.2. Briefly, primers SpegfpF and SfogfpR incorporating *Spe*I and *Sfo*I sites respectively were used to amplify gfp+ from the vector pWH1012gfp+13 (Scholz *et al*., 2000). This amplified fragment was cloned into pGEM-T Easy and restricted with *Spe*I and *Sfo*I, generating a 5' end and a blunt end (C). The plasmid pGEM-T Easy::StrepII-wzy was digested with *Nhe*I and *Sfo*I, which also generated 5' and blunt ends respectively. The *Spe*I-*gfp*$^{+}$-*Sfo*I restricted fragment from pGEM-T Easy::*speI-gfp*$^{+}$*-sfoI* was ligated into pGEM-T Easy::s*trepII-wzy*, to create the plasmid pGEM-T Easy::*gfp*$^{+}$*-wzy*. The *gfp*$^{+}$*wzy* region was amplified with forward and reverse primers (KpnIgfpF and XbaIgfpR) which incorporated *Kpn*I and *Xba*I restriction sites respectively, and was cloned into pGEM-T Easy, creating plasmid pGEM-T Easy::*gfp*$^{+}$*-wzy*. Simultaneous restriction digest of this plasmid, and of pSCRhaB2 (E) was undertaken, and the *Kpn*I-*gfp*$^{+}$*-wzy-Xba*I fragment was ligated into pSCRhaB2 to generate the plasmid pGFP^{+}-Wzy. The *wzy* fragment is illustrated in dark blue, *gfp*$^{+}$ is illustrated in green, pGEM-T Easy in cyan, and pSCRhaB2 in magenta. Amp resistance conferred by pGEM-T Easy is indicated in orange, and trimethoprim in yellow.

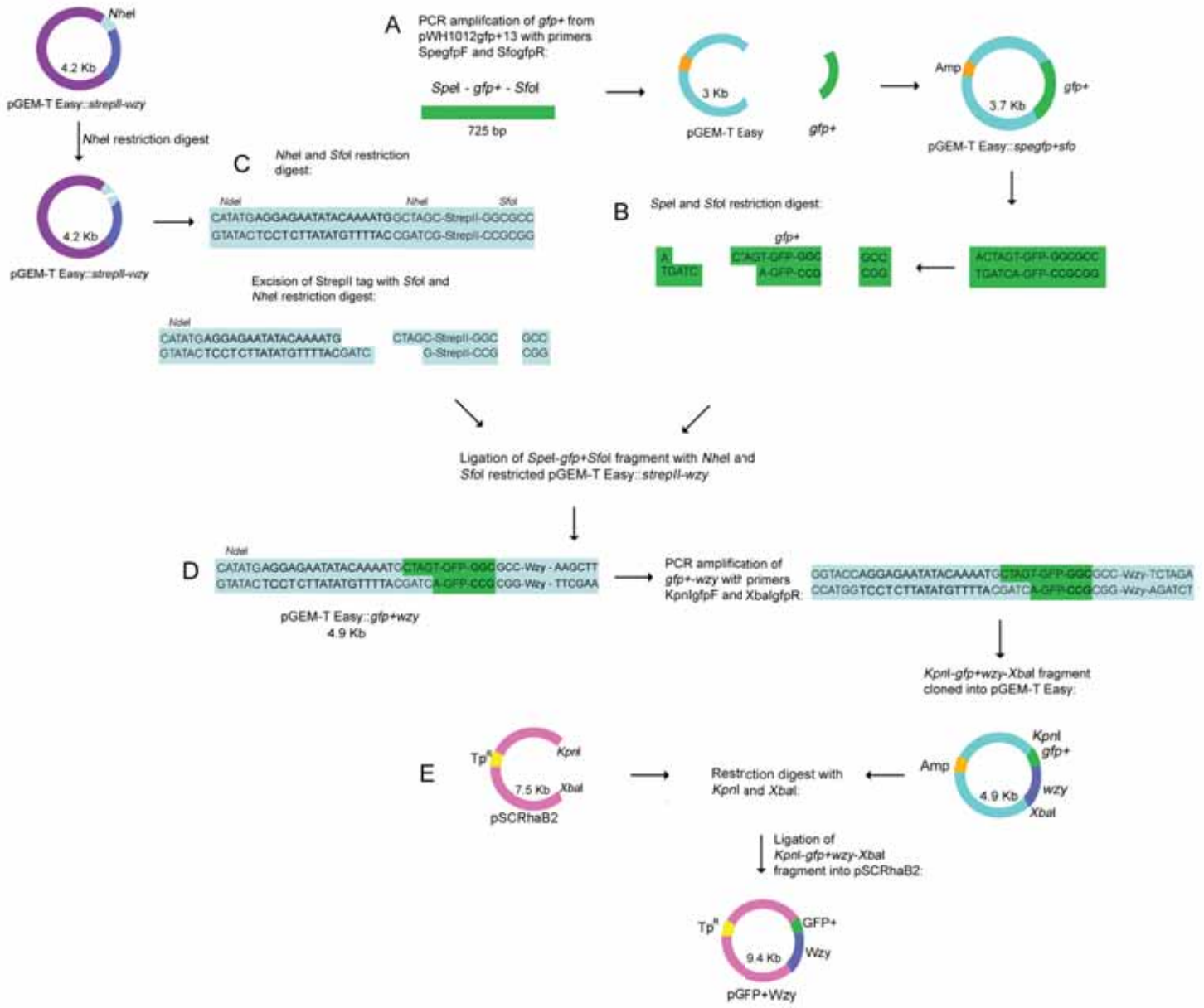
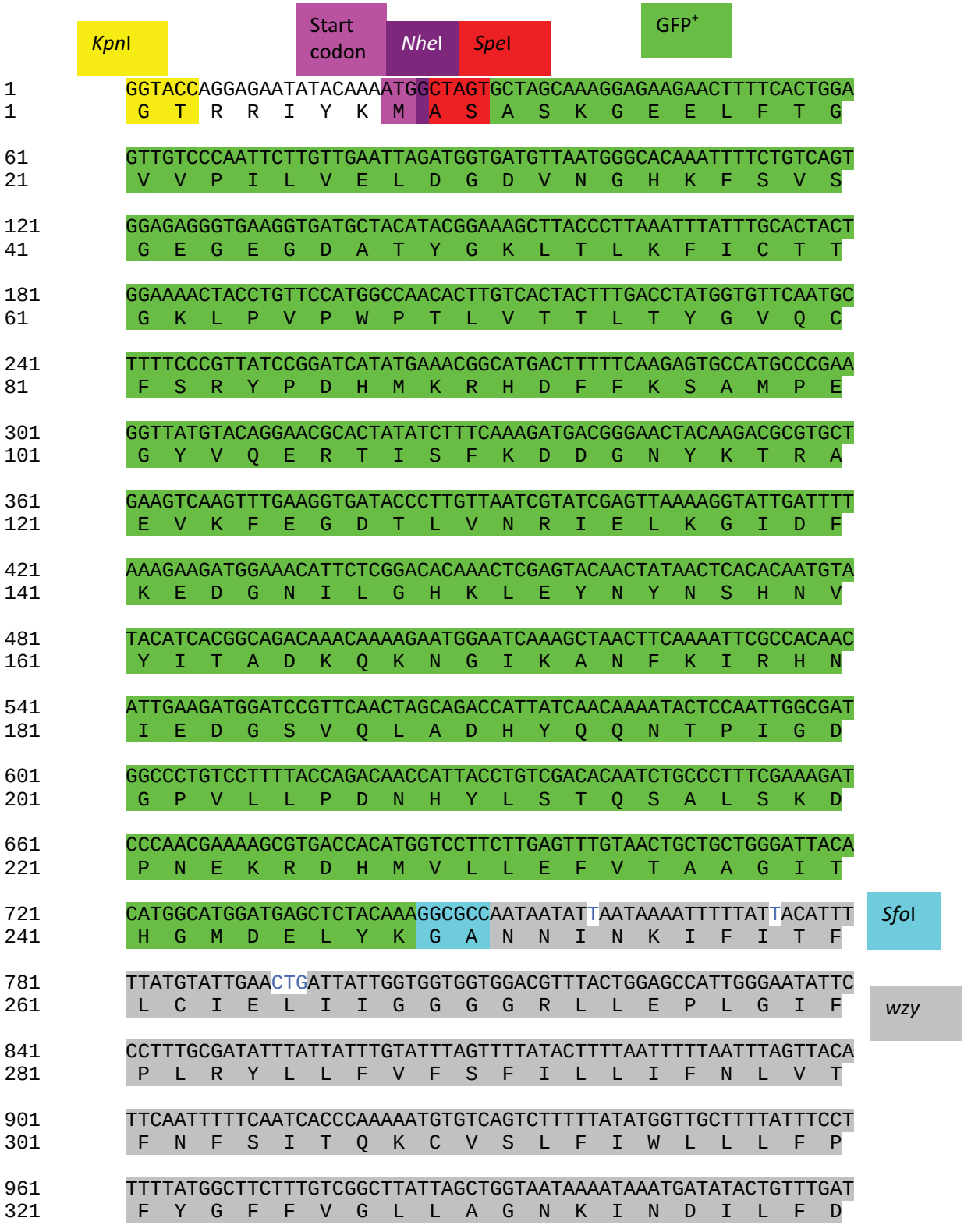

NheI
4.2 Kb
pGEM-T Easy::strepII-wzy
NheI restriction digest
4.2 Kb
pGEM-T Easy::strepII-wzy
A
PCR amplifcation of gfp+ from pWH1012gfp+13 with primers SpegfpF and SfogfpR:
SpeI - gfp+ - SfoI
725 bp
3 Kb
pGEM-T Easy
gfp+
Amp
3.7 Kb
gfp+
pGEM-T Easy::spegfp+sfo
C
NheI and SfoI restriction digest:
NdeI
NheI
SfoI
CATATGAGGAGAATATACAAAATGCTAGC-StrepII-GGCGCC
GTATACTCCTCTTATATGTTTTACCGATCG-StrepII-CCGCGG
B
SpeI and SfoI restriction digest:
gfp+
ACTAGT-GFP-GGCGCC
TGATCA-GFP-CCGCGG
Excision of StrepII tag with SfoI and NheI restriction digest:
NdeI
CATATGAGGAGAATATACAAAATG
GTATACTCCTCTTATATGTTTTACGATC
CTAGC-StrepII-GGC
G-StrepII-CCG
GCC
CGG
Ligation of SpeI-gfp+SfoI fragment with NheI and SfoI restricted pGEM-T Easy::strepII-wzy
D
NdeI
CATATGAGGAGAATATACAAAATGCTAGT-GFP-GGCGCC-Wzy - AAGCTT
GTATACTCCTCTTATATGTTTTACGATCA-GFP-CCGCGG-Wzy - TTCGAA
pGEM-T Easy::gfp+wzy
4.9 Kb
PCR amplification of gfp+-wzy with primers KpnIgfpF and XbaIgfpR:
GGTACCAGGAGAATATACAAAATGCTAGT-GFP-GGCGCC-Wzy-TCTAGA
CCATGGTCCTCTTATATGTTTTACGATCA-GFP-CCGCGG-Wzy-AGATCT
KpnI-gfp+wzy-XbaI fragment cloned into pGEM-T Easy:
E
Tp^R
KpnI
7.5 Kb
XbaI
pSCRhaB2
Restriction digest with KpnI and XbaI:
Amp
KpnI
gfp+
4.9 Kb
wzy
XbaI
Ligation of KpnI-gfp+wzy-XbaI fragment into pSCRhaB2:
Tp^R
GFP+
9.4 Kb
Wzy
pGFP+Wzy

Figure 6.7 DNA sequence and predicted amino acid sequence of the pGFP^{+}-Wzy fragment

Shown here is the sequence of the insert in pGFP^{+}-Wzy (determined by DNA sequencing), and the corresponding amino acid sequence. The *Kpn*I site is indicated in yellow, the start and stop codons in magenta, the GFP^{+} sequence in green, and the StrepII tag in red. The *Spe*I sites destroyed via the cloning are indicated in red, and the *Nhe*I sites are indicated in violet, and wzy_{SF} in grey. Also indicated is the *Xba*I site in yellow. The *wzy* substitutions in codons 4, 9 and 23 are indicated in blue text.


KpnI
Start codon
NheI
SpeI
GFP+
1 GGTACCAGGAGAATATACAAAATGGCTAGTGCTAGCAAAGGAGAAGAACTTTTCACTGGA
1 G T R R I Y K M A S A S K G E E L F T G
61 GTTGTCCCAATTCTTGTTGAATTAGATGGTGATGTTAATGGGCACAAATTTTCTGTCAGT
21 V V P I L V E L D G D V N G H K F S V S
121 GGAGAGGGTGAAGGTGATGCTACATACGGAAAGCTTACCCTTAAATTTATTTGCACTACT
41 G E G E G D A T Y G K L T L K F I C T T
181 GGAAAACTACCTGTTCCATGGCCAACACTTGTCACTACTTTGACCTATGGTGTTCAATGC
61 G K L P V P W P T L V T T L T Y G V Q C
241 TTTTCCCGTTATCCGGATCATATGAAACGGCATGACTTTTTCAAGAGTGCCATGCCCGAA
81 F S R Y P D H M K R H D F F K S A M P E
301 GGTTATGTACAGGAACGCACTATATCTTTCAAAGATGACGGGAACTACAAGACGCGTGCT
101 G Y V Q E R T I S F K D D G N Y K T R A
361 GAAGTCAAGTTTGAAGGTGATACCCTTGTTAATCGTATCGAGTTAAAAGGTATTGATTTT
121 E V K F E G D T L V N R I E L K G I D F
421 AAAGAAGATGGAAACATTCTCGGACACAAACTCGAGTACAACTATAACTCACACAATGTA
141 K E D G N I L G H K L E Y N Y N S H N V
481 TACATCACGGCAGACAAACAAAAGAATGGAATCAAAGCTAACTTCAAAATTCGCCACAAC
161 Y I T A D K Q K N G I K A N F K I R H N
541 ATTGAAGATGGATCCGTTCAACTAGCAGACCATTATCAACAAAATACTCCAATTGGCGAT
181 I E D G S V Q L A D H Y Q Q N T P I G D
601 GGCCCTGTCCTTTTACCAGACAACCATTACCTGTCGACACAATCTGCCCTTTCGAAAGAT
201 G P V L L P D N H Y L S T Q S A L S K D
661 CCCAACGAAAAGCGTGACCACATGGTCCTTCTTGAGTTTGTAACTGCTGCTGGGATTACA
221 P N E K R D H M V L L E F V T A A G I T
721 CATGGCATGGATGAGCTCTACAAAGGCGCCAATAATATTAATAAAATTTTTATTACATTT
241 H G M D E L Y K G A N N I N K I F I T F
781 TTATGTATTGAACTGATTATTGGTGGTGGTGGACGTTTACTGGAGCCATTGGGAATATTC
261 L C I E L I I G G G G R L L E P L G I F
841 CCTTTGCGATATTTATTATTTGTATTTAGTTTTATACTTTTAATTTTTAATTTAGTTACA
281 P L R Y L L F V F S F I L L I F N L V T
901 TTCAATTTTTCAATCACCCAAAAATGTGTCAGTCTTTTTATATGGTTGCTTTTATTTCCT
301 F N F S I T Q K C V S L F I W L L L F P
961 TTTTATGGCTTCTTTGTCGGCTTATTAGCTGGTAATAAAATAAATGATATACTGTTTGAT
321 F Y G F F V G L L A G N K I N D I L F D
SfoI
wzy

```
1021        GTGCAACCATACCTTTTTATGCTGTCACTTATATATCTATTTACACTAAGATATACTTTA
341          V  Q  P  Y  L  F  M  L  S  L  I  Y  L  F  T  L  R  Y  T  L

1081        AAAGTATTTTCATGTGAGATTTTTATTAAAATAGTTAATGCATTTGCATTATATGGATCA
361          K  V  F  S  C  E  I  F  I  K  I  V  N  A  F  A  L  Y  G  S

1141        CTGTTATATATTTCATACATAATTTTGTTGAATTTCGGTTTGTTAAATTTTAATTTAATT
381          L  L  Y  I  S  Y  I  I  L  L  N  F  G  L  L  N  F  N  L  I

1201        TATGAACACTTATCATTGACTAGCGAGTTCTTTTTTCGTCCCGATGGGGCTTTTTTTTCC
401          Y  E  H  L  S  L  T  S  E  F  F  F  R  P  D  G  A  F  F  S

1261        AAATCCTTCTACTTTTTTGGTGTCGGTGCGATTATCAGTTTTGTCGACAAAAAATATTTA
421          K  S  F  Y  F  F  G  V  G  A  I  I  S  F  V  D  K  K  Y  L

1321        AAATGTCTCATAATAGTGCTTGCGATATTATTGACAGAATCAAGAGGTGTATTACTTTTT
441          K  C  L  I  I  V  L  A  I  L  L  T  E  S  R  G  V  L  L  F

1381        ACAACATTATCACTGTTATTAGCCAGTTTTAAATTACATAAGCTATATTTAAATACTATT
461          T  T  L  S  L  L  L  A  S  F  K  L  H  K  L  Y  L  N  T  I

1441        ATAATAATATTGGGCAGCGTTCTATTTATAATTATGCTTTACATGGTCGGATCACGCAGT
481          I  I  I  L  G  S  V  L  F  I  I  M  L  Y  M  V  G  S  R  S

1501        GAAGATTCTGACTCTGTTAGATTTAATGATTTATATTTTTATTATAAAAATGTTGATTTA
501          E  D  S  D  S  V  R  F  N  D  L  Y  F  Y  Y  K  N  V  D  L

1561        GCGACGTTCTTGTTTGGAAGAGGATTTGGTTCATTTATATTAGATCGATTAAGGATTGAA
521          A  T  F  L  F  G  R  G  F  G  S  F  I  L  D  R  L  R  I  E

1621        ATAGTACCTCTTGAGATACTTCAGAAAACAGGCGTTATTGGTGTATTTATATCATTAGTT
541          I  V  P  L  E  I  L  Q  K  T  G  V  I  G  V  F  I  S  L  V

1681        CCTATGTTGCTTATCTTTTTGAAAGGCTATTTTTTAAATAGTACAAAAACATCATTAATG
561          P  M  L  L  I  F  L  K  G  Y  F  L  N  S  T  K  T  S  L  M

1741        ATGTCGTTAATACTTTTTTTCAGTATTACCGTTTCTATAACTAATCCATTTCTTTTTACA
581          M  S  L  I  L  F  F  S  I  T  V  S  I  T  N  P  F  L  F  T

1801        CCCATGGGAATTTTTATTATAGGCGTTGTAGTTTTATGGGTATTTTCTATAGAAAATATC
601          P  M  G  I  F  I  I  G  V  V  V  L  W  V  F  S  I  E  N  I

1861        CAAATTAGTAATAACCTCACTTCTGGAGCAAAATAATCTAGA
621          Q  I  S  N  N  L  T  S  G  A  K  *  S  R
```

Stop codon

*Xba*I

(pGFP$^+$-StrepII-Wzy) and RMM109 (pGFP$^+$-Wzy) exhibited smooth LPS, indicating complementation of the Wzy deficiency in RMM109 (Figure 6.8A lanes 4 and 5 respectively). However, although all three Wzy-fusion proteins (StrepII-Wzy, GFP$^+$-StrepII-Wzy and GFP$^+$-Wzy) complement the Wzy deficiency in RMM109, there are differences between the conferred LPS profiles in the three strains; StrepII-Wzy exhibits partial activity (due to significant amount of SR-LPS) and while conferring smooth LPS, wild-type 11-17 RU Oag modal chain length control is absent (Figure 6.3A lane 3, 6.3B lane 6). Partial activity is also seen in RMM109 (pGFP$^+$-Wzy). In this strain less of the SR-LPS is present, and modal chain length appears to be more strongly regulated than in the strain harbouring pStrepII-Wzy. This was indicated by a region of intense banding around the 11-17 RU level (Figure 6.8A, lane 5), which has not been generally observed in strains carrying pStrepII-Wzy. Surprisingly, the LPS of RMM109 (pGFP$^+$-StrepII-Wzy) had regulated Oag modal chain length, and little or no SR-LPS was observed, and this strain had an LPS profile similar to the wild-type PE638 (Figure 6.8A, lane 1). Thus pGFP$^+$-StrepII-Wzy had the most efficient complementation of the *wzy* mutation in strain RMM109 (Figure 6.8A, lane 4; Figure 6.8B, lane 4).

6.7 COLICIN SENSITIVITY OF STRAINS EXPRESSING TAGGED WZY PROTEINS

In order to further investigate the functionality of the tagged Wzy proteins, cultures of the RMM109 strains harbouring plasmids pSCRhaB2, pStrepII-Wzy, pGFP$^+$-Wzy and pGFP$^+$-StrepII-Wzy were assessed in a colicin sensitivity assay (section 2.12). Overnight cultures were subcultured and induced with 0.2% (w/v) rhamnose and streaked onto a plate containing a killed single streak of strain RMA2782, a colicin E2 producing strain. Colicin sensitive *S. flexneri* strains exhibit a significant killing zone bordering the colicin-E2 zone, whilst the resistant strains exhibit minimal or absent killing zones. Figure 6.9 displays the

Figure 6.8 LPS Oag modal chain length conferred by the pSCRhaB2-based plasmids

S. flexneri PE638 and RMM109 strains harbouring pSCRhaB2-based plasmids were grown in MH + Tp and induced with 0.2% (w/v) rhamnose for 1.5 h (section 2.10.5). LPS samples were prepared, electrophoresed on a SDS 15% polyacrylamide gel and silver stained (section 2.11). A) Strains in each lane are as follows: 1) PE638 [pSCRhaB2], 2) RMM109 [pSCRhaB2], 3) RMM109 [pStrepII-Wzy], 4) RMM109 [pGFP^{+}-StrepII-Wzy], and 5) RMM109 [pGFP^{+}-Wzy]. The lipid-A core and number of repeat units (RUs) are indicated on the right. Each lane contains approximately 1.6 x 10^{8} cells. Oag modal length regions are indicated in brackets B) A magnified and contrast-inverted image of the LPS Oag modal chain lengths conferred by the pSCRhaB2-based plasmids presented in A).

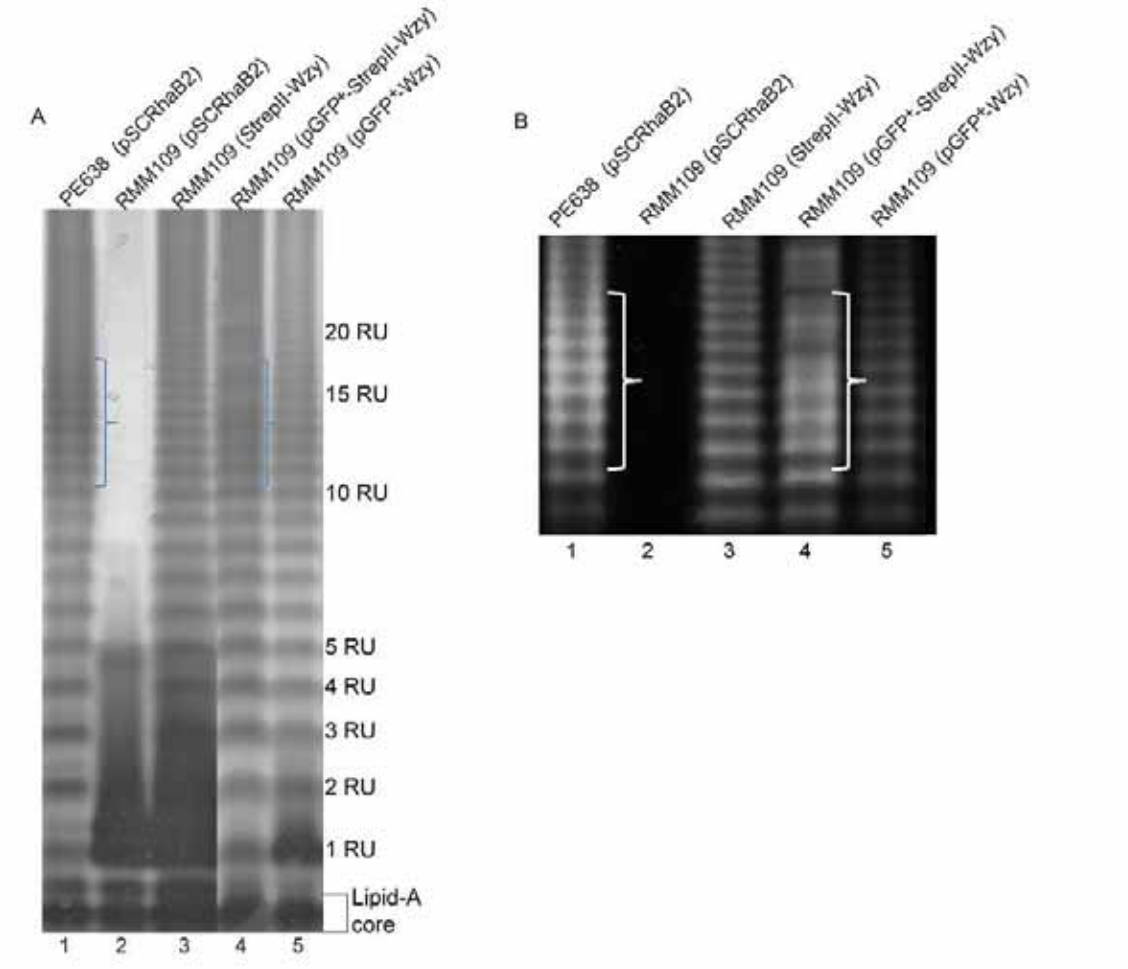
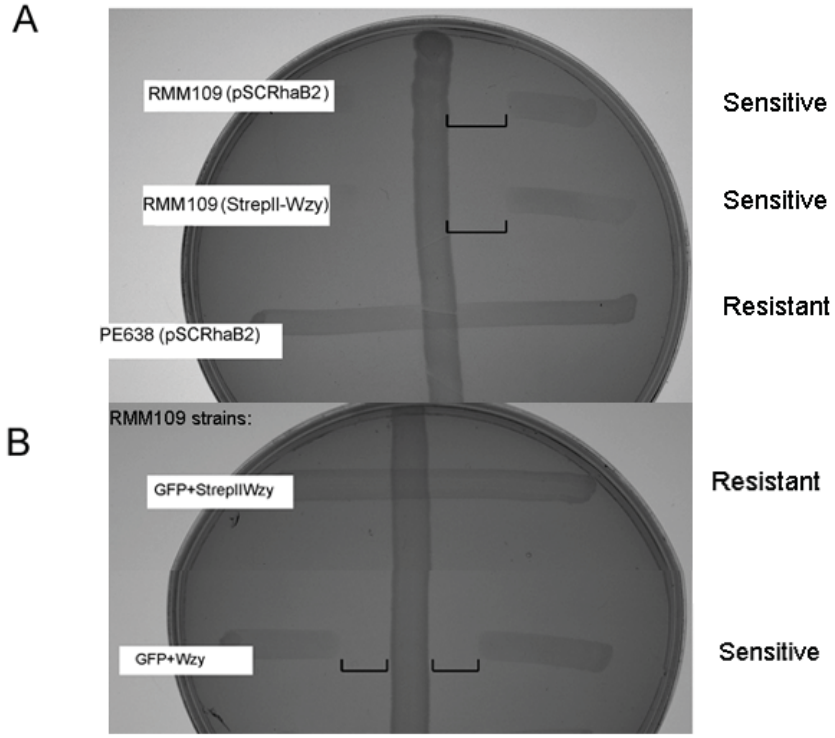

A
PE638 (pSCRhaB2)
RMM109 (pSCRhaB2)
RMM109 (StrepII-Wzy)
RMM109 (pGFP*-StrepII-Wzy)
RMM109 (pGFP*-Wzy)
20 RU
15 RU
10 RU
5 RU
4 RU
3 RU
2 RU
1 RU
Lipid-A core
1 2 3 4 5
B
PE638 (pSCRhaB2)
RMM109 (pSCRhaB2)
RMM109 (StrepII-Wzy)
RMM109 (pGFP*-StrepII-Wzy)
RMM109 (pGFP*-Wzy)
1 2 3 4 5

Figure 6.9 Susceptibility of RMM109 strains encoding tagged Wzy protein to colicin

RMM109 strains carrying pSCRhaB2-based plasmids were grown in MH + Tp, induced with 0.2% (w/v) rhamnose (section 2.10.5) and streaked on the upper layer of MH agar (the lower level was contained a single streak of RMA2782, a colicin E2-producing strain grown on LB), section 2.12. Zones of inhibitory growth of strains proximal to the RMA2782 streak are indicated with black brackets. Strains in panel A are as follows: RMM109 (pSCRhaB2), RMM109 (StrepII-Wzy), and PE638 (pSCRhaB2). Strains in panel B are as follows: RMM109 (pGFP^{+}-StrepII-Wzy), and RMM109 (pGFP^{+}-Wzy).


A
RMM109 (pSCRhaB2)
Sensitive
RMM109 (StrepII-Wzy)
Sensitive
Resistant
PE638 (pSCRhaB2)
B
RMM109 strains:
GFP+StrepIIWzy
Resistant
GFP+Wzy
Sensitive

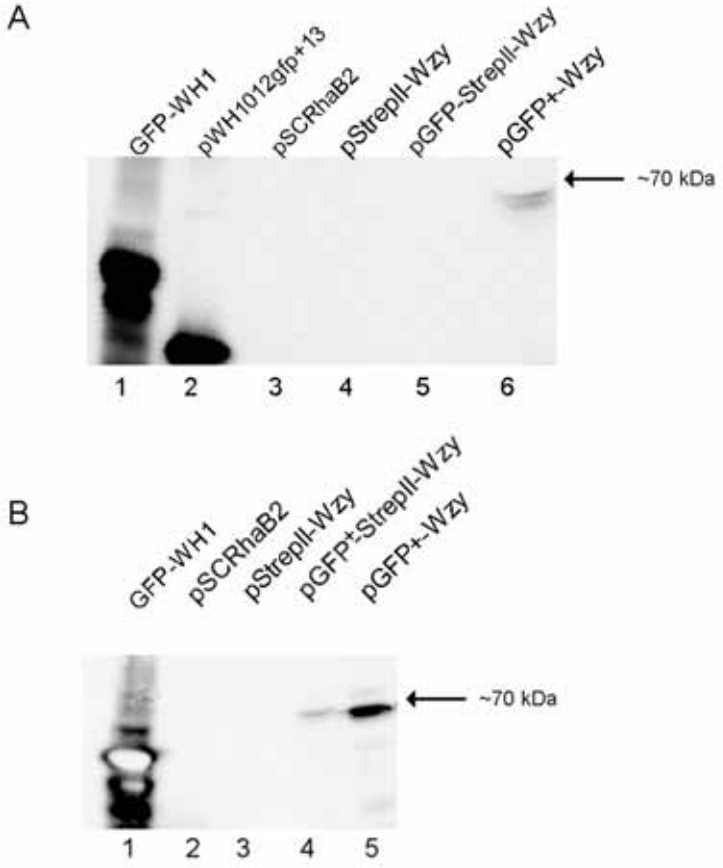
resulting colicin sensitivity assay which indicated that PE638 harbouring pSCRhaB2 exhibited no killing zone, and hence appeared to be completely resistant to the action of colicin (Figure 6.9A). However, RMM109 harbouring the pSCRhaB2 vector control exhibited a significant killing zone, and was sensitive to killing by colicin (Figure 6.9A). Interestingly, despite possessing a partially functional Wzy protein (based on LPS profiles, Figures 6.3 and 6.8), the strain harbouring pStrepII-Wzy was sensitive to colicin E2 (Figure 5.13A). Similarly, the resulting killing zone for pGFP^{+}-Wzy indicated substantial killing comparable to the RMM109 carrying either pSCRhaB2 or pStrepII-Wzy (Figure 6.9B). Conversely, RMM109 (pGFP^{+}-StrepII-Wzy) exhibited wild-type resistance to colicin killing (Figure 6.8B), thereby supporting the data obtained from the LPS profiling, suggesting that GFP^{+}-StrepII-Wzy exhibits near wild-type functionality compared to the other Wzy-tagged proteins. These data show that LPS with wild-type Oag modal chain length distribution confers resistance to colicin E2, whilst LPS with significant SR-LPS confers sensitivity.

6.8 DETECTION OF GFP^{+}-WZY PROTEINS

Detection of GFP^{+}-tagged Wzy proteins was attempted in order to assess whether subsequent experiments such as pull down assays and/or co-localisation microscopy assays with mCherry-Wzz$_{SF}$ were possible. A variety of Western immunoblotting techniques were used to attempt detection of both GFP^{+}-StrepII-Wzy and GFP^{+}-Wzy using the anti-GFP antibody (Roche). RMM109 strains harbouring the GFP^{+}-tagged constructs with the appropriate vector controls were grown, induced with 0.2% (w/v) rhamnose, solubilised with 1% (w/v) DDM (section 2.10.5) and the resulting soluble fractions were subjected to SDS-PAGE and immunoblotting (section 2.10). The negative controls pSCRhaB2 and StrepII-Wzy resulted in no detectable protein when probed with anti-GFP (Figure 6.10A, lanes 3 and 4 respectively) as expected. The positive control strain RMA2611 (E1315 (pWH1012gfp+13) (GFP^{+})) (Figure 6.10A, lane 1) and the other positive control, CV-1 lysate containing GFP-

Figure 6.10 Western immunoblotting of GFP$^+$-tagged Wzy proteins

Strains were grown in LB + Kan (RMA2611) or MH + Tp (pSCRhaB2-based plasmids), and strains harbouring pSCRhaB2-based plasmids were induced with 0.2% (w/v) rhamnose for 2 h, and whole membranes were prepared (section 2.10.5), treated with 1% (w/v) DDM and soluble fractions were electrophoresed on a SDS 15% polyacrylamide gel (section 2.10.1). Approximately 30 μL from each fraction (and ~15 μL of the GFP positive control, CV-1 lysate containing GFP-WH1 protein (section 2.10.5)) was loaded in each well. Whole cell lysate samples of RMA2611 (E1315 harbouring pWH1012gfp+13) were prepared and electrophoresed along the *S. flexneri* soluble fractions (section 2.10.1). Samples were detected with anti-GFP monoclonal antibodies (Roche) diluted 1:1000. Strains in each lane are as follows: A) 1) 15 μL of CV-1 lysate containing GFP-WH1 protein, 2) RMA2611 (E1315 (pWH1012gfp+13)), 3) RMM109 (pSCRhaB2), 4) RMM109 (pStrepII-Wzy), 5) RMM109 (pGFP$^+$-StrepII-Wzy) and 6) RMM109 (pGFP$^+$-Wzy). Prestained Benchmark protein marker was used to determine protein sizes (Invitrogen). B) *S. flexneri* strain RMM109 harbouring pSCRhaB2-based constructs was grown in MH + Tp, induced at 25°C for 8 h with 0.2% (w/v) rhamnose, and solubilised with DDM (section 2.10.5). Samples were subjected to 15% SDS-PAGE and detected with anti-GFP monoclonal antibodies at a dilution of 1:1000. Strains in each lane are as follows: 1) 15 μL of CV-1 lysate containing GFP-WH1 protein 2) RMM109 (pSCRhaB2), 3) RMM109 (pStrepII-Wzy), 4) (pGFP$^+$-StrepII-Wzy) and 5) RMM109 (pGFP$^+$-Wzy).


A
GFP-WH1
pWH1012gfp+13
pSCRhaB2
pStrepII-Wzy
pGFP-StrepII-Wzy
pGFP+-Wzy
~70 kDa
1 2 3 4 5 6
B
GFP-WH1
pSCRhaB2
pStrepII-Wzy
pGFP+StrepII-Wzy
pGFP+-Wzy
~70 kDa
1 2 3 4 5

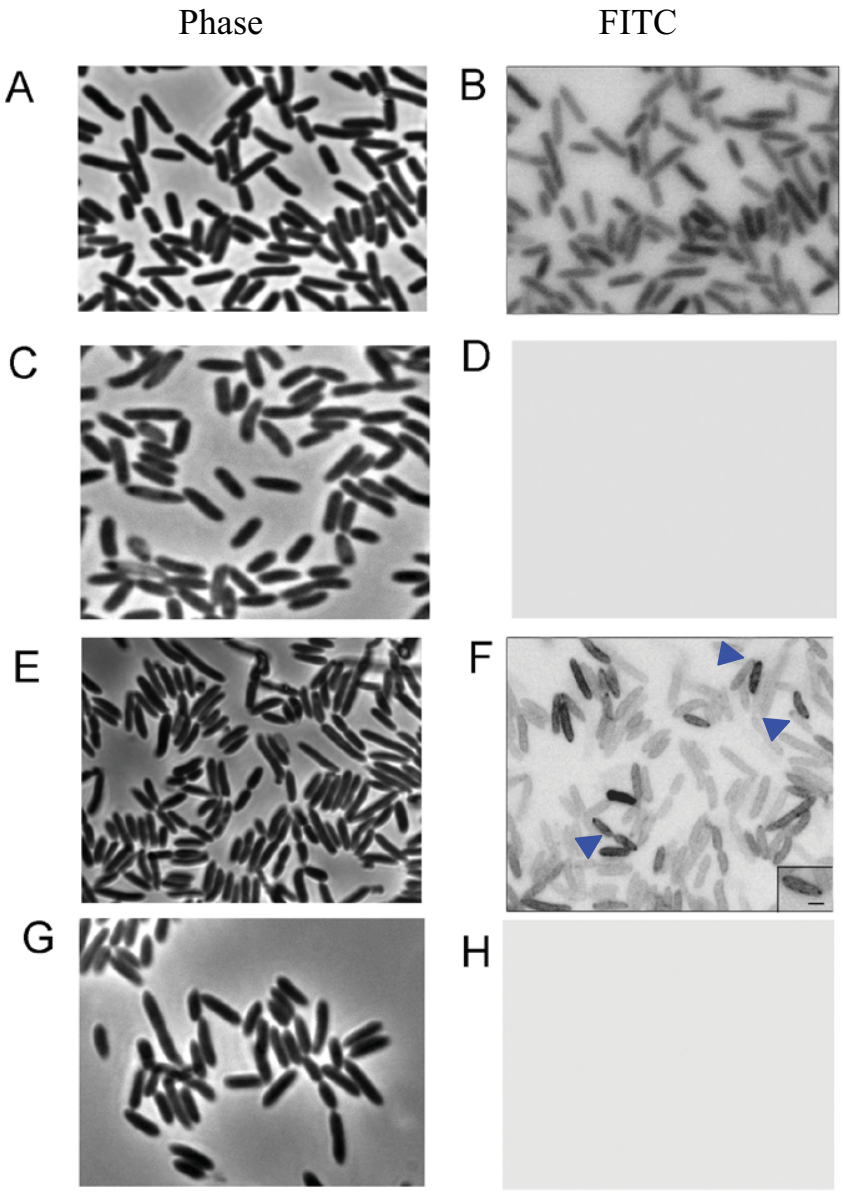
WH1 protein (section 2.10.5), had reactive bands (Figure 6.10A, lane 2). Furthermore, pGFP$^+$-StrepII-Wzy was also not detected under these experimental conditions (Figure 6.10A, lane 5). RMM109 carrying pGFP$^+$-Wzy conversely, exhibited a band that migrated slightly smaller than the 72 kDa predicted protein for GFP$^+$-Wzy (Figure 6.10, lane 6). It appeared that GFP$^+$-Wzy was significantly easier to detect than GFP$^+$-StrepII-Wzy in many of the different growth conditions and Western transfer conditions attempted, including the optimised method for inner membrane detection described by Abeyrathne *et al.* (2007) (data not shown). Because of the difficulty encountered detecting the pGFP$^+$-StrepII-Wzy protein, Western immunoblotting was then conducted with RMM109 strains that were prepared after an 8 h induction period at 25°C. The induced cultures were solubilised as described in section 2.10.5, and the resulting soluble fractions from both GFP$^+$-StrepII-Wzy and GFP$^+$-Wzy were subjected to SDS-PAGE and Western immunoblotting. The results indicated that whilst the GFP$^+$-tagged control protein was observed, RMM109 harbouring pSCRhaB2 and StrepII-Wzy did not have detectable bands (Figure 6.10B, lanes 2 and 3 respectively) as expected. However, bands were detected for both GFP$^+$-StrepII-Wzy and GFP$^+$-Wzy proteins (Figure 6.10B, lanes 4 and 5), and GFP$^+$-Wzy was more readily detected than GFP$^+$-StrepII-Wzy (Figure 6.10B, lane 5), perhaps due to a higher level of expression.

6.9 LOCALISATION OF GFP$^+$-TAGGED WZY PROTEINS

GFP$^+$-Wzy protein fluorescence was investigated by epi-fluorescence microscopy (section 2.13). RMA2611 (E1315 (pWH1012gfp+13)) and RMM109 harbouring pSCRhaB2, pGFP$^+$-Wzy and pGFP$^+$-StrepII-Wzy were grown, induced with 0.2% (w/v) rhamnose for 2 h, and the bacteria were mounted onto slides as described in section 2.13. The positive control strain RMA2611 possessing GFP$^+$ displayed significant fluorescence while the strain carrying the pSCRhaB2 vector control exhibited negligible fluorescence, as expected (Figure 6.11B and D respectively). RMM109 (pGFP$^+$-StrepII-Wzy) exhibited negligible fluorescence under

Figure 6.11 Epi-fluorescent microscopy detection of GFP$^+$-tagged Wzy in *S. flexneri*

RMM109 strains were grown in MH + Tp, induced with 0.2% (w/v) rhamnose for 2 h (section 2.10.5), washed in 1 x PBS and fixed with 3.7% (w/v) formaldehyde (section 2.13). RMA2611 (E1315 (pWH1012gfp+13)) was grown in LB + Kan, and cells were washed in 1 x PBS and fixed with 3.7% (w/v) formaldehyde as described above. Bacteria were visualised with epi-fluorescence microscopy at 100 x magnification. Images A, C, E, and G were taken in phase, whilst B, D, F and H were visualised with the FITC filter set. The images are as follows: A) RMA2611, B) RMA2611 FITC, C) RMM109 (pSCRhaB2), D) RMM109 (pSCRhaB2) FITC, E) RMM109 (pGFP$^+$-Wzy), F) RMM109 (pGFP$^+$-Wzy) FITC, G) RMM109 (pGFP$^+$-StrepII-Wzy) and H) RMM109 (pGFP$^+$-StrepII-Wzy) FITC. Scale bars indicate approximately 1 μm. Individual clusters of GFP$^+$-Wzy are indicated with blue arrowheads in panel F. Panels on the right were converted to greyscale and contrast-inverted to enhance clarity.


Phase
FITC
A
B
C
D
E
F
G
H

the experimental conditions tested (Figure 6.11H). However, the strain containing pGFP$^+$-Wzy displayed readily detectable fluorescence (Figure 6.11F) and in ~10% of cells, GFP$^+$-Wzy appeared to be at the cell periphery (Figure 6.11F).

6.10 SUMMARY

This chapter describes the construction of multiple Wzy-tagged proteins; the StrepII tag was utilised to generate an N-terminal tagged Wzy protein which exhibited partial function and was able to complement the *wzy* mutation in RMM109, although detection of StrepII-Wzy proved to be unreliable. Using the pStrepII-Wzy construct, unique restriction sites were targeted for insertion of GFP$^+$ which resulted in two proteins, one which had GFP$^+$ inserted prior to the StrepII tag (GFP$^+$-StrepII-Wzy) and one which excised the StrepII tag altogether (GFP$^+$-Wzy). These proteins exhibited functionality to different degrees, with GFP$^+$-StrepII-Wzy conferring functionality close to that of wild-type, GFP$^+$-Wzy conferring intermediate functionality, and StrepII-Wzy exhibiting partial functionality. Western immunoblotting was performed and detection of both GFP$^+$-tagged proteins was achieved, although GFP$^+$-Wzy was significantly more readily detected. Preliminary microscopy revealed that the partially active GFP$^+$-Wzy was readily detectable in approximately 10% of cells with GFP$^+$-Wzy appearing to locate to the cell periphery. However detection of GFP$^+$-StrepII-Wzy (which confers a higher level of wild-type functionality) with epi-fluorescence microscopy was not achieved.

CHAPTER SEVEN

DISCUSSION

7.1 INTRODUCTION

Regulation of Oag polysaccharide chain length is controlled by the Wzz protein, a member of the polysaccharide co-polymerase (PCP1a) family. *S. flexneri* Wzz (Wzz_{SF}) confers an average chain modal length of 10-17 Oag RUs. Many mutagenesis studies targeting residues throughout Wzz indicate that function is an overall property of the protein and not limited to one particular region. Recently, the periplasmic domain structures of a collection of PCP proteins have been solved, showing that these structures exhibit similarly shaped protomers and oligomers (Tocilj *et al.*, 2008). Wzz_{ST} forms pentameric oligomers, with WzzE assembling into octameric oligomers and FepE assembling into nonameric structures. Multiple studies have indicated that Wzz has the ability to form high order oligomers, suggesting that oligomerisation is important in function. Marolda *et al.* (2006) have provided genetic evidence that proteins involved in Oag/ECA biosynthesis and processing may function as a complex, as ECA-associated Wzx can fully complement an LPS Oag associated Wzx deficient mutant if the remaining ECA gene cluster is deleted (Marolda *et al.*, 2006). However, despite studies conducted to probe the Wzz structure-function relationship and identify a regulation mechanism, little is known about the mode of action of these proteins in determining Oag modal chain length. Several models of the likely mechanisms of Oag chain regulation have been proposed. Bastin *et al.* (1993) initially suggested that Wzz acts as a molecular timer, allowing polymerization to occur to a particular point, hence increasing the number of repeat units added to the chain. An alternative model proposed by Morona *et al.* (1995) suggested that Wzz acts as a molecular chaperone, facilitating the interaction between Wzy and WaaL and modal length is the result of the ratio

of Wzy and WaaL. Published data indicate that the ratio of Wzy and Wzz was important in determining Oag modal chain length, which is supportive of the latter model. 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.

7.2 RESULTS

In Chapter 3, in-frame linker mutagenesis was undertaken to generate a collection of Wzz_i proteins. The resulting proteins were classified into five categories (section 3.3); Class I mutants appeared to be non-functional (conferring random LPS Oag modal chain length), Class II mutants exhibited VS-type LPS Oag modal chain length (2-10 RUs), whilst Class III mutants conferred shorter than wild type LPS Oag chain length (8-14 RUs). Class IV mutants conferred near wild-type LPS Oag modal chain length (11-19 RUs), and Class V mutants increased the resulting LPS Oag modal chain length to 16-25 RUs (equivalent to L-type). Cross-linking analyses demonstrated 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 Wzz_i mutants from Classes I, II and III. Strains expressing the mutants were also assessed for their sensitivity to colicin E2; this showed that there was a strong correlation between LPS Oag modal chain length and susceptibility to colicin, as strains expressing Classes IV and V mutants had greater resistance to the lethal action of colicin E2 compared to strains expressing Class II and III mutants. 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 4 showed that 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 mono-modal LPS Oag modal chain length. Plasmids expressing FLAG-tagged $Wzz_{G305A/G311A}$, Wzz_{SF} and Wzz_{ST} were constructed, and strains expressing these plasmids emulated the LPS Oag modal chain lengths observed with pET-17b based constructs, including the bimodality observed with co-expression of $Wzz_{G305A/G311A}$ with wild-type Wzz_{SF}. 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 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 investigated, and yielded the similarly observed bimodal phenotype exhibited by co-expression of wild-type Wzz_{SF} and $Wzz_{G305A/G311A}$. This suggests that these mutants may 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 most likely interacting with wild-type Wzz_{SF}.

Chapter 5 described the construction of the fluorescent fusion protein mCherry-Wzz_{SF}, and indicated that mCherry-Wzz_{SF} was able to fully complement a *wzz* mutation in *S. flexneri*, and showed that mCherry-Wzz_{SF} could be detected by Western immunoblotting. Epi-fluorescence microscopy was used to visualise mCherry-Wzz_{SF}, demonstrating that the protein localised to distinct regions at the cell periphery.

Chapter 6 described the construction of several tagged Wzy proteins; the StrepII tag was utilised to generate an N-terminal tagged Wzy protein which exhibited partial function (smooth LPS lacking Oag chain modal length control) when used to complement the

RMM109 *wzy* mutation, although detection of StrepII-Wzy proved unreliable. Using the pStrepII-Wzy construct, unique restriction sites were targeted for insertion of GFP^+ which resulted in two proteins, one with GFP^+ inserted prior to the StrepII tag (GFP^+-StrepII-Wzy) and one which excised the StrepII tag (GFP^+-Wzy). These proteins exhibited different degrees of functionality with GFP^+-StrepII-Wzy having activity close to that of wild-type, while GFP^+-Wzy had reduced activity. Western immunoblotting was performed and detection of both GFP^+-tagged proteins was achieved, although GFP^+-Wzy was significantly more readily detected. Preliminary epi-fluorescence microscopy revealed that the partially active GFP^+-Wzy was readily detectable in approximately 10% of cells with GFP^+-Wzy appearing to locate to the cell periphery. However GFP^+-StrepII-Wzy (which had a higher level of wild-type activity) could not be detected.

7.3 CORRELATION OF WZZi LOCATION AND FUNCTION

Previous studies have examined regions of Wzz to investigate whether particular locations were responsible for determining LPS Oag modal chain length. Franco *et al.* (1998) investigated amino acid variations existing in *E. coli* Wzz proteins that confer either intermediate (10-18 RUs) LPS Oag or Short type (7-16 RUs) LPS Oag, and discovered that some residues, e.g. glycine 221, do not affect the resulting modal chain length, whilst others do, e.g. isoleucine 224 (Franco *et al.*, 1998). They also showed that chimeric *wzz* genes which incorporate a segment of the Short type-conferring *wzz* into the Intermediate type-conferring *wzz*, reduced the resulting Oag modal chain length from an Intermediate to Short type. In the same study, it was also shown that the N-terminal region of Wzz conferring Long-type LPS Oag (16-25 RUs) differ significantly to the Short- and Intermediate-conferring Wzz proteins, and that chimeric genes incorporating the first 424 bases (encoding the first 140 amino acids) of Long-type conferring *wzz* resulted in a Long-type Oag chain length. Also, point mutations were made in Intermediate type-conferring *E. coli wzz* at aa residues that were substantially

different to those present in Long-type *E. coli* Wzz and *S. enterica* Wzz (section 1.4.2). The results indicated that one point mutation shifted the modal length from Intermediate to Long-type, and one point mutation reduced the modal length, whilst two other changes had no effect (section 1.4.2). Further studies by Daniels *et al.* (1999) investigated a K267N substitution in Wzz_{SF} via site-directed mutagenesis, which resulted in an increase in S type Oag (11-17 RUs) to I-type Oag (13-20 RUs), therefore illustrating that this residue is important in Oag chain length variation, but does not singularly drive Oag modal length to Long-type (Daniels and Morona, 1999). Together, these data suggest that Wzz function results from the complex interaction of numerous critically positioned amino acids. Taking into account that Wzz has such a prominent regulatory function, and that many studies (including those presented here) suggest that functionality is a result of the 3D protein structure as a whole, it might be predicted that the insertion of 5-aa would abolish regulation of Oag chain length in most mutants. In contrast, the results presented in Chapter 3 indicate that 5-aa insertions throughout Wzz result in a range of effects on the resulting Oag modal chain length. These data support the previous findings that there is no one particular region of Wzz that drives Oag modal chain length, but that Wzz function is a complex and intricate result of the protein as a whole entity.

7.4 WZZi INSERTIONS MAPPED TO MONOMERIC AND OLIGOMERIC STRUCTURES

The insertions in each Wzz_i mutant were mapped onto the 3D PCP oligomeric structures to determine if the interrupted regions appeared to be structurally critical and/or important in either intra- or inter-monomer interactions. Despite the fact that each mutant exhibits a unique 5-aa insertion, and the different amino acids may be affecting the structure and resulting phenotype, Classes II-V generally consist of insertions that are predicted to be located in a similar region (the predicted insertion positions relative to the Wzz 3D structure

for the Class I mutants span many regions of the structure). The seven Class I mutants were unable to restore wild-type LPS Oag modal chain length, thus it appears as though protein function is sensitive to 5-aa insertions at these locations. The α/β base domain, comprised of β sheets β1 - β4, α helices α1-α2, and the lowest region of α6 (in Wzz_{ST}), and α1-α4 and the lowest region of α6 (in FepE and WzzE), appears to play an important role in inter-monomeric interactions and oligomerisation (Tocilj *et al.*, 2008). The mutations that are predicted to be located within this region are: i66, i80, i81, i92, i128, i131 and i290. There are noteworthy differences in the LPS phenotype of these mutations – the Class I i66 and i290 mutations do not impart any Oag modal chain length control, whereas the others do. Mutation i66 is predicted to exist on the periphery of the monomer, hence may be in a position to influence hydrophobic interactions between monomers, thus affecting stability and oligomerisation. The β strands comprising the central feature of the α/β base domain are key structural elements in bringing the N- and C-termini closer together, which may also bring the TM segments closer. Mutation i279, predicted to be located on the final β strand (β4), may also be interfering with this process, perhaps by steric hindrance. However, in reference to Class IV i80 and i81 mutations, it is unclear why this region allows a 5-aa insertion with little alteration in resulting Oag chain length, although the insertions appear to be located in a region embedded within α2 and α8 and may not be destabilising intra-monomeric interactions. Mutation i92 (Class III) is predicted to be located close to the top of an α helix, however i92 differs from both i80 and i81 in that the α2 helix is more exposed to the inner cavity than α1, hence possibly being less tolerant to a 5-aa insertion. The other Class III $_{i138}$ insertion is predicted to be located close to the base of the oligomeric structure, towards the end of β3. Considering its predicted proximity to the α/β base domain, it is surprising that the resulting phenotype is merely a slight decrease in the LPS Oag modal chain length. This mutant insertion, however, is predicted to be located towards the end of β3 (Wzz_{ST}) or on the loop between β3 and α6 (FepE), and this position may be able to sustain extra amino acids

without drastically altering nearby structural features, as it appears to be on the cusp of the oligomer. In the case of i290, as it is close to TM2, the insertion may be influencing a key region needed for function. In previous Wzz_{SF} mutagenesis studies, P292, one of the highly conserved prolines present in the proline-glycine rich motif, was found to knock out Wzz function when converted to an alanine. This proline is theorised to be in *cis* conformation, hence providing rigidity and stability, and influencing the orientation of the base domain and TM2 (Morona *et al.*, 2009; Tocilj *et al.*, 2008). Mutation i290, being located so closely to this conserved residue, may undermine this critical arrangement. It is also possible that this mutation destabilises any anchoring of the oligomer at the membrane surface, as residues close to the inner membrane are considered to interact with lipid head groups (Tocilj *et al.*, 2008).

There are a number of mutations predicted to be located on the long extended hairpin α6, such as i161, i191 and i199. Mutant i191 is in Class II, and i161 and i199 are Class I mutants. From the 3D crystal structure analyses, it appears as though the α6 helix is involved in maintaining intra-monomeric stability, by interacting with α2 via conserved hydrophobic residues (Tocilj *et al.*, 2008). Indeed, all mutations predicted to exist within α6 are in Classes I and II and have similar phenotypes. The upper region of the oligomer is speculated to be involved in interacting with outer membrane proteins such as those involved in LPS export. It is possible that perturbation in the central α-helix either results in disruption of local protein conformation or affects interaction with outer membrane proteins. Various conserved residues, including I237 and L240 in FepE are present which form a leucine zipper motif. If this leucine zipper is critical for monomeric or oligomeric stability and/or interactions in Wzz_{SF}, it is possible that the insertion in i199, predicted to be located within two turns of this region, may severely disrupt those interactions.

Mutations i247 and i255 are predicted to be mapped to the upper region of the oligomer, within or close to α7 and α8. These α-helices play a role in inter monomeric

stability, as they interact with the long extended α6 helix on neighbouring monomers (Tocilj *et al.*, 2008). The mutations i247 and i255 (both Class II mutants), are located on the outermost region of the oligomer, facing outwards and high up on the structure. It is interesting to note that mutations that are located on the exterior region of the oligomer can play such a role in Wzz function. Another aforementioned mutagenesis study conducted on Wzz_{SF} indicates that a K267N mutation, predicted to be located on the outer side of the oligomer on the lower region of α8, results in an increase in Oag modal chain length. Hence, it appears that residues on the exterior face of Wzz oligomers have the ability to influence Wzz function, as previously proposed (Tocilj *et al.*, 2008).

The Class V mutants i128 and i131, conferring longer Oag chain length modality, were predicted to be located in the loop between β2 and β3 in the α/β base domain, directly in the central cavity of the oligomer. The phenotype resulting from the 5-aa insertions at this location is not observed for any other Wzz_i mutant. Previous Wzz_{SF} mutagenesis studies have not yielded mutants which increase the Oag modal chain length to this degree. It can be speculated that the cause for such a dramatic modal length change may be attributed to the change this 5-aa insertion exerts on the cavity width, and that the increase in the number of aa residues within this cavity is widening it by increasing the size of the α/β base domain. In general it appears as though Class V, Class IV and Class III mutants are mapped to internal regions on the oligomeric structure, whereas Class II mutants have their insertions mapped exclusively to external regions, and Class I mutant insertions are mapped to both internal and external regions.

7.5 PROTEIN DETECTION

Chapter 3 assessed Wzz protein expression by the 18 mutants. Whilst mutants i32, i66, i80, i81, i92, i128, i131, i199, i219 and i290 were detected at a level comparable to wild-

type Wzz_{SF}, mutants i138, i161, i231 and i247 were detected at a lower level than wild-type Wzz_{SF}. Strains harbouring mutants i163, i191, i255 and i279 produced protein below the limit of detection. From these data, it would appear that less than wild-type protein detection does not correlate with a particular resulting Oag modal length, as these mutants are from a range of different phenotypic classes. It is not impossible that these 5-aa insertions have disrupted, concealed or altered particular epitopes that the Wzz antibody binds to, hence altering protein detection in some mutants. This is still possible, despite the fact that the anti-Wzz antibody used was polyclonal. It is also possible that several Class I mutants lack control of Oag modal chain length due to low protein production and/or protein misfolding. However, it does seem that a very small amount of Wzz protein is required to establish a regulated Oag modal chain length, consistent with other studies.

7.6 SUSCEPTIBILITY TO COLICIN E2

Section 3.5 described the differences observed by each Wzz_i Class in susceptibility to the lethal action of Colicin E2. RMA2741 (pQE-30), Classes I, II and III exhibited large zones of clearance, and Classes IV, V and RMA2741 (pRMCD30) exhibited smaller zones of clearance, thus showing a strong correlation between LPS Oag modal chain length and susceptibility to Colicin E2. It appears that wild-type or longer LPS Oag modal chain length is required for greater resistance to Colicin E2. Chapter 6 compared the susceptibility of RMM109 strains encoding Wzy variants (StrepII-Wzy, GFP^+-Wzy and GFP^+-StrepII-Wzy) to colicin E2, and showed that only the strain expressing GFP^+-StrepII-Wzy was resistant to colicin, indicating that the presence of SR-LPS renders the strain susceptible to the action of colicin E2. Hence, it appears that the presence of less than wild-type Oag levels and/or modal length on the cell surface allows access for colicin E2, and hence increased susceptibility.

CHAPTER SEVEN

7.7 CROSSLINKING ANALYSES

Previous cross-linking studies have shown that Wzz can oligomerise. Recent studies have suggested that Wzz *E. coli* O86:H2 can form tetrameric oligomers (Guo *et al.*, 2006), while others have identified hexameric oligomerisation of Wzz (Daniels and Morona, 1999). In this study, higher order oligomers were easily detected in wild-type Wzz_{SF} and a number of selected mutants. It was shown that the mutants conferring longer Oag chain length have comparable cross-linking profiles to wild-type, whereas mutants resulting in random or shorter chain lengths do not appear to oligomerise as well. The lack of detectable oligomers in the resulting cross-linking profiles of i92, i219 and i290 may also be attributed to weak stability of the mutant oligomers (Figure 3.5). It is possible that various Wzz_i mutant proteins may be able to form oligomers (e.g. i290), however they either cannot be stably maintained or are perhaps incapable of successful interactions with other Oag processing proteins (or putative OM binding partners) to confer wild-type Oag modal chain length. An unusual feature of mutant i290 is the presence of the extra 30-kDa band in the non cross-linked sample. This band appears to be Wzz-related as it is readily detected by the Wzz antibody and is present in all other mutants (including wild-type Wzz), although never detected in the absence of cross-linking. The 30-kDa variation might have an altered conformation and may be non-functional, and it is possible that the presence of this variant in the non cross-linked sample of i290 is linked to the fact that i290 appears to be non-functional (Figure 3.5). In contrast, monomers and dimers of Class V Wzz_i mutants were detected at a much lower intensity when subjected to cross-linking than other mutants, and higher order oligomers were easily detected. Previous findings have shown that Wzz_{SF} dimers appear to be very stable, as heating to 100°C in the presence of SDS does not cause complete disassociation. In this study, Wzz_{SF} and Wzz_i cross-linked and non cross-linked samples were heated to 100°C for 5 minutes and subjected to Western immunoblotting to ascertain whether or not the Wzz_i mutant dimers exhibit the similar stable trait of the wild-type dimers by being able to

withstand dissociation in the presence of SDS at 100°C. Only the wild-type Wzz_{SF} and Class V i128 and i131 mutant dimers were detected after this treatment (Figure 3.6). From these data, it appears as though there is a positive correlation between dimeric stability and wild-type or longer Oag modal chain length determination. These experiments show that the mutants which exhibit higher order oligomers as judged by *in vivo* formaldehyde cross-linking form stable dimers, and it may be possible that this feature is a key factor in the ability to form oligomers, perhaps by providing a scaffolding element.

7.8 CO-EXPRESSION STUDIES OF FLAG-WZZ AND HIS_6-WZZ

In Chapter 4, we investigated previously constructed Wzz mutants in co-expression assays with wild-type Wzz_{SF} to observe the resulting LPS Oag modal chain length, and discovered that several phenotypes were exhibited by different mutants. K31A (pRMCD119) and P292A were non-functional, as previously established (Daniels and Morona, 1999). In the same study, P286A was reported to confer diminished levels of wild-type LPS Oag modal chain length, however in this study P286A was found to be completely non-functional (section 4.2). K267N (pRMCD108), whilst previously shown to increase the Oag modal chain length (13- 20 RUs), resulted in a minimal increase in this study (12-18 RUs). However, these subtle effects may be caused by differences in the strain backgrounds used. All of these mutant proteins appeared to exhibit wild-type LPS Oag modal chain length when co-expressed with wild-type Wzz_{SF}, however it is difficult to distinguish subtle differences in Oag modality in this situation. It cannot be determined from these co-expression assays whether the aforementioned proteins are able to interact with wild-type Wzz_{SF}.

$Wzz_{G305A/G311A}$ (encoded by pRMCD113) conferred VS-type LPS Oag modal chain length, as was previously shown (Daniels and Morona, 1999) and exhibited bimodal chain length when co-expressed with wild-type Wzz_{SF} (section 4.2). The VS defect was not corrected by the function of wild-type Wzz_{SF}, and the presence of two modal chain lengths in

one strain suggests that these two Wzz proteins are able to function independently and confer distinct Oag modal chains on LPS. In contrast, while $Wzz_{M32T/I35C}$ (pRMCD122) conferred a VS type LPS Oag modal chain length like $Wzz_{G305A/G311A}$, when co-expressed with wild-type Wzz_{SF}, only one LPS Oag modal chain length was observed. In this case, the VS defect is corrected by the function of wild-type Wzz_{SF}, and M32T/I35C function appears to have no effect on the function of Wzz_{SF}, indicating that wild-type Wzz_{SF} was dominant. These data suggested that $Wzz_{M32T/I35C}$ interacts with wild-type Wzz_{SF}, whilst $Wzz_{G305A/G311A}$ does not so. Consistent with the observed $Wzz_{G305A/G311A}$ functionality, previous studies investigating oligomerisation of Wzz mutants using *in vivo* formaldehyde cross linking indicated that $Wzz_{G305A/G311A}$ can form oligomers (Daniels and Morona, 1999).

Wzz_{ST} (encoded by pRMCD80) conferred L-type LPS Oag modal chain length (17-26 RUs), and when co-expressed with wild-type Wzz_{SF} that confers a modal length of 10-17 RUs, the resultant LPS had a single Oag chain length which was slightly longer than wild-type LPS at 12-20 RUs. The presence of one LPS modal chain length suggests that Wzz_{ST} is capable of interacting with wild-type Wzz_{SF}. This was interesting to observe as $Wzz_{G305A/G311A}$ differs to wild-type by 2-aa, however Wzz_{SF} and Wzz_{ST} differ by 23% overall. The hybrid protein $Wzz_{ST/SF}$ (pRMCD104) conferred an LPS Oag modal chain length of 11-23 RUs, although this protein appeared to exhibit ~50 % functionality to other Wzz proteins (section 4.2). When co-expressed with wild-type Wzz_{SF}, the resulting LPS Oag modal chain length was near wild-type 10-18 RUs. This protein appeared to have a minimal effect on wild-type Wzz_{SF} function, and $Wzz_{ST/SF}$ function was corrected by wild-type Wzz_{SF}. The hybrid protein $Wzz_{SF/ST}$ (encoded by pRMCD106) conferred an LPS Oag modal chain length of 14-25 RUs, longer than that conferred by wild-type Wzz_{SF}, and closer to the L-type Oag modal chain length conferred by Wzz_{ST}. As suggested previously, the increase in LPS Oag modal chain length conferred by $Wzz_{SF/ST}$ indicates that the C-terminus contains residues that have a major impact on the Oag modal chain length (Daniels and Morona, 1999). When co-expressed with wild-type Wzz_{SF}, the resulting LPS Oag modal chain length

was 10-21 RUs, which is slightly longer than wild-type (10-17 RUs). From these data, we formulated a hypothesis that LPS Oag bimodal chain length distribution is an indication of Wzz function conferred by two independently acting Wzz proteins, and that $Wzz_{G305A/G311A}$ is unable to interact with wild-type Wzz_{SF} (see section 7.9). Conversely, when the expression of two functional proteins results in a single LPS Oag modal chain length, this is suggestive of Wzz::Wzz interactions occurring.

7.9 COPURIFICATION OF WZZ PROTEINS AND WZZ::WZZ INTERACTION

Further investigation was undertaken of the hypothesis that bimodal LPS Oag chain length was the result of two different Wzz proteins failing to interact when co-expressed and conversely to determine if the single Oag chain length observed when Wzz_{ST} and Wzz_{SF} were co-expressed correlated with an interaction of these two proteins. Characterisation of the pBAD-Wzz_{SF}, pBAD-Wzz_{ST} and pBAD-$Wzz_{G305A/G311A}$ proteins encoded by these plasmids conferred LPS Oag chain modal lengths as previously observed for the pET-17b-based constructs (described in section 4.4); pBAD-Wzz_{ST} conferred L-type LPS Oag modal chain length, pBAD-$Wzz_{G305A/G311A}$ conferred VS type LPS Oag modal chain length, and co-expression with wild-type resulted in single LPS Oag modal chain length and bimodality, respectively. However, an interesting observation was that of the LPS Oag modal chain length when FLAG-$Wzz_{G305A/G311A}$ was co-expressed with His_6-Wzz_{SF} and induced with IPTG, thus overexpressing His_6-Wzz_{SF} (Figure 4.8C). Bimodality appeared to disappear in the presence of high levels of His_6-Wzz_{SF}. It was shown that FLAG-tagged Wzz expression (of FLAG-Wzz_{SF}, FLAG-Wzz_{ST} and FLAG-$Wzz_{G305A/G311A}$) was weaker than that of the pQE-30-based expression of His_6-Wzz_{SF} (Figure 4.10), thus the function and effect of His_6-Wzz_{SF} appeared to swamp (or override) the effect of the FLAG-tagged proteins. A co-purification assay was performed to determine if FLAG-Wzz_{ST} or FLAG-$Wzz_{G305A/G311A}$ was co-purified with His_6-Wzz_{SF}, in order to definitively establish whether wild-type His_6-Wzz_{SF} was able to interact

with Wzz_{ST} and $Wzz_{G305A/G311A}$. The results indicated that wild-type Wzz was capable of interacting with FLAG-Wzz_{ST}, as it was readily detected in the elution fraction containing purified His_6-Wzz_{SF} (Figure 4.12). In contrast, only very low levels of FLAG-$Wzz_{G305A/G311A}$ were detected in the elution fraction containing purified His_6-Wzz_{SF}. As hypothesized above, the data acquired from the co-expression and co-purification experiments suggest that whereas Wzz_{ST} and Wzz_{SF} are capable of efficient interaction, $Wzz_{G305A/G311}$ and Wzz_{SF} are not. Therefore, showing that FLAG-$Wzz_{G305A/G311A}$ and His_6-Wzz_{SF} interact poorly illustrates that the masking of bimodality when high levels of His_6-Wzz_{SF} are present in the cells is possibly an effect of His_6-Wzz_{SF} swamping the Oag processing system (and potentially the Oag processing complex), and may be prohibiting $Wzz_{G305A/G311A}$ oligomers from taking part in the Oag processing mechanism, or diminishing the levels of $Wzz_{G305A/G311A}$ oligomers from interacting with other Oag processing proteins. A mechanism of bimodality exists in many bacteria, i.e., wild-type *S. flexneri* exhibits VL-type LPS Oag conferred by Wzz_{PHS-2} and also S-type LPS Oag modal chain length conferred by Wzz_{SF}, hence there must be a mechanism in place to enable two competing assembly systems/mechanisms to operate simultaneously. What determines this is possibly due to particular residues and/or structural features of the two proteins. As Wzz_{PHS-2} and Wzz_{SF} only exhibit ~22% amino acid sequence identity, it is difficult to identify the critical elements that establish bimodal distribution. It appears that from the co-expression of $Wzz_{G305A/G311A}$ and wild-type Wzz_{SF}, and co-purification data of FLAG-$Wzz_{G305A/G311A}$ and His_6-Wzz_{SF}, the changes that confer VS-type LPS Oag modal chain length has also resulted in a similar separation of the Oag processing systems.

It is possible to speculate on the mode of chain length regulation occurring from these results. A single modal chain length may result from two or more compatible Wzz monomers being able to form oligomers together. The chain length may be a result of different sized oligomers, creating hybrid oligomers, or may be proportional to the diameter of the cavity the oligomeric structure, as was discussed in section 7.4. Conversely the two amino acid changes in the glycines present in TM2 of $Wzz_{G305A/G311}$ have drastically altered

not only the resulting Oag modal chain length (to decrease it from 11-17 RUs conferred by Wzz_{SF} down to 3-8 RUs), but have also inhibited/abolished effective interactions between wild-type Wzz and $Wzz_{G305A/G311A}$, resulting in the observed LPS bimodality exhibited when these two proteins are co-expressed. This implies that one, or the other, or both of these two glycines either directly or indirectly are involved in interactions between Wzz protein monomers. Future experiments encompassing the identical G305A/G311A substitutions in Wzz_{ST} would therefore be expected to abolish interactions between $Wzz_{STG305A/G311A}$ and Wzz_{SF}. It is possible that TM interactions are hindered or weakened and result in the poor interaction detected between wild-type Wzz and $Wzz_{G305A/G311A.}$ However, it is also possible that these glycine to alanine substitutions have affected function indirectly, and that the resulting lack of interaction may be an issue of steric hindrance or accessibility of particular amino acids. It is possible that the glycine residues assist positional spacing of other critical aa residues which affect interaction. The high frequency of glycines in transmembrane domains of many membrane proteins has been observed, and this has generally suggested a structural role (Javadpour *et al.*, 1999). Despite the fact that alanine residues are considered to stabilise the helical conformation relative to glycine because of reduced backbone entropy and it buries more apolar area upon folding, glycine residues are suggested to not affect the secondary structure of helical transmembrane segments, and instead function as molecular 'indentations' to facilitate helix packing. Therefore, despite alanine and glycine both being non polar neutral amino acids and exhibiting similar characteristics, the differences between the two amino acids may be adequate to affect structural function. As previously observed, Wzz_{SF} possesses a GXXXG motif in the TM2 region (G305-G309), which is a well known sequence for mediating interactions between adjacent transmembrane regions and is involved in membrane protein folding (Senes *et al.*, 2004; Tocilj *et al.*, 2008). Previous studies have suggested the importance of tightly packed GXXXG motifs within TM regions in oligomer formation (Jenei *et al.*, 2009), and if indeed this motif similarly is responsible for establishing or stabilising oligomeric Wzz_{SF} structures, then it is possible that the G to A substitutions in the TM2

region are capable of undermining this significant arrangement. The hypothesis formulated in light of these data, and the data acquired in Chapter 3 demonstrating greater dimer stability of wild-type/Longer LPS-conferring Wzz, is that interactions between TM2 regions are needed to establish a dimeric Wzz and this G305A/G311A mutation blocks interaction with G305 and G311 in wild-type Wzz_{SF}, however interactions are still capable of occurring between G305A/G311A in $Wzz_{G305A/G311A}$.

7.10 VS-CONFERRING WZZi CO-EXPRESSION WITH WILD-TYPE WZZ, AND THEIR FUNCTIONAL ACTIVITY

The possibility existed that bimodality was a characteristic of all VS conferring Wzz proteins, and not just specifically due to the G to A changes in the TM2 region of Wzz. As shown in Chapter 4, the resulting LPS profiles indicated that bimodal Oag chain length was observed for all five Class II mutants tested, although the degrees of modal control varied between them. The Class II co-expression assays suggest that biochemical analyses between Class II mutants and wild-type Wzz_{SF} would likely yield the same result as that seen in FLAG-$Wzz_{G305A/G311A}$, in that Class II mutants may not co-purify alongside wild-type Wzz_{SF}, although due to lack of time, this was not investigated. These results are surprising as the locations of the mutations are drastically different; Class II mutations are generally predicted to be located on the upper side of the monomer and oligomer. Inspection of the PCP structures determined by Tocilj *et al.* (2008) suggests that the Wzz protein conferring the longest LPS Oag modal chain length (FepE) exhibits closely packed helical regions in the FepE nonamers, whilst fewer interactions exist between these helical regions in WzzE (which was shown to assemble into octamers), and very few exist for Wzz_{ST} (which assemble into pentamers) (Tocilj *et al.*, 2008). The diameter of the FepE oligomer cavity is 50 Å, whilst the WzzE and Wzz_{ST} oligomeric cavities were 23 Å and 9 Å in diameter, respectively (Tocilj *et al.*, 2008). If the resulting modal chain length conferred is a result of cavity size and how

tightly packed the helical regions are, this affords the possibility that the 5-aa insertions in Wzz have resulted in the VS-conferring Wzz proteins possessing smaller sized cavities and/or that they exhibit fewer interactions between their helical regions.

Previous it was theorized that the upper region of oligomers may be responsible for interacting with outer membrane proteins (Tocilj *et al.*, 2008), although there is currently no evidence for this. As the Class II mutants are predicted to be located on the long α6 helix, it is possible that the altered structure in the upper monomeric region of Wzz has affected putative interactions with outer membrane proteins, and that this may be the basis for VS-conferred modality. However, G305A/G311A are not located on the upper region of the oligomeric structure, but located in the TM2 region; and the nature of the mutations is extremely different, as Class II mutants consist of 5-aa insertions within Wzz_{SF}, and G305A/G311A has two single aa substitutions. Despite the complete difference in the nature and location of the alterations in the Class II mutants compared to G305A/G311A, the resulting phenotypes are identical. It is possible that the G to A changes in TM2 results in VS modal chain length due to conformational changes that result in the mutant Wzz monomers being unable to interact with wild-type monomers, and that amino acid changes in the upper area of the monomers also equates to VS modal length. It is possible that the resulting monomers have conformational changes in the structure that make it unable to interact with wild-type monomers, and this is applicable to both the Class II mutants and $Wzz_{G305A/G311A}$. Taken together, these data suggest that Wzz::Wzz interactions are determined by residues in the TM2 region, and those residues affected by the Class II 5-aa insertions. In the case of the Class II Wzz_i proteins, these residues also affect oligomer stability, as shown in cross-linking analyses (Figure 3.5) and analyses assessing dimeric stability (Figure 3.6). Based on this, it is predicted that $Wzz_{G305A/G311A}$ will also exhibit hindered dimeric stability, and also as shown by $Wzz_{G305A/G311A}$, Class II Wzz_i mutants will exhibit limited ability to interact with wild-type Wzz_{SF} due to bimodality in co-expression assays. Due to time constraints, this could not be investigated.

Interestingly, there is not always a correlation between VS type LPS Oag and the proposed inability to interact with wild type Wzz_{SF}. The LPS phenotype conferred by $Wzz_{M32T/I35C}$ is the same as that conferred by Class II Wzz_i proteins and $Wzz_{G305A/G311A}$, however when co-expression with wild-type Wzz_{SF} the resulting LPS Oag has a mono-modal length. This predicts that whilst some mutations resulting in VS modal length (G305A/G311A, and Class II Wzz_i mutations) appear to prevent interaction between mutant Wzz proteins with Wzz_{SF} wild type, others such as M32T/I35C do not exhibit this possible phenomenon. The M32T/I35C mutation thus appears to only affect control of Oag modal chain length.

7.11 SUBCELLULAR LOCALISATION OF WZZ AND WZY

Chapter 5 and 6 described the construction and characterisation of mCherry-Wzz and three tagged Wzy proteins. Sections 5.1 to 5.4 describe the construction of pQMCherry-Wzz, a pQE-30 based plasmid with *wzz_{SF}* and an N-terminal in frame *mcherry* tag (Shaner *et al.*, 2004). mCherry-Wzz_{SF} was shown to be fully functional, as the LPS produced by the Wzz mutant complemented with pQMCherry-Wzz_{SF} showed restoration of the 11-17 RU wild-type Oag modal chain length (Figure 5.3). mCherry-Wzz was also detected by Western immunoblotting and epi-fluorescent microscopy. The fluorescent microscopy data indicate that mCherry-Wzz showed a particular pattern of distribution in the cells that was not observed when mCherry was expressed alone (Figure 5.4). The pattern of distribution appeared to indicate mCherry-Wzz_{SF} located in the periphery of the cell, suggesting that Wzz_{SF} is not evenly distributed in the cell. It is possible that Oag processing proteins such as Wzy, WaaL and Wzx would also be localised in similar regions as those of Wzz and perhaps Oag processing occurs in localised regions. Future experiments to investigate this may include fluorescence co-localisation or FRET assays to determine conclusively which Oag processing proteins Wzz interacts with directly, and also co-purification of tagged Wzy or

other Oag processing proteins. The construction of three Wzy proteins was undertaken in preparation for the latter experiments. Chapter 6 described the construction of the three Wzy_{SF} proteins; initially, a StrepII-tagged Wzy protein was generated, and shown to be partially functional. However, detection of this protein proved to be unreliable, and subsequently two GFP^+-tagged Wzy proteins were constructed; GFP^+-StrepII-Wzy and GFP^+-Wzy (section 6.6). These two proteins showed increased functionality compared to StrepII-Wzy, with GFP^+-StrepII-Wzy having near wild-type activity (Figure 6.8), however, this protein also proved difficult to detect (Figure 6.10). While GFP^+-StrepII-Wzy was successfully detected following a longer than normal induction period at 25°C, this protein was not detected using epi-fluorescence microscopy under any of the experimental conditions tested (Figure 6.11). GFP^+-Wzy, however, was readily detectable by Western immunoblotting and was observed using epi-fluorescence microscopy. Colicin sensitivity assays indicated that strains expressing GFP^+-StrepII-Wzy exhibited resistance to colicin similar to wild-type strains, however GFP^+-Wzy did not, further confirming the wild-type activity of GFP^+-StrepII-Wzy and the lack thereof by GFP^+-Wzy. The relative ease with which GFP^+-Wzy was detected is inversely related to the ability of this protein to effectively complement the *wzy* mutation in RMM109 suggests that GFP^+-Wzy may be misfolded or that GFP^+ interferes with its function but renders it detectable via epi-fluorescence microscopy. However, both GFP^+-Wzy and GFP^+-Strep-Wzy were only detected in membrane fractions when DDM was used as a solubilising agent, illustrating that these proteins are likely to be located in the inner membrane (Figure 6.10).

7.12 PHENOTYPIC VARIATION BETWEEN TAGGED WZY PROTEINS

All three tagged Wzy proteins exhibit different degrees of functionality. StrepII-Wzy, whilst conferring smooth Oag chains in the resulting LPS profile does not restore wild-type average modal chain length of 11-17 RUs and the complemented *wzy* mutant strain still

exhibits a partial semi-rough phenotype. The GFP$^+$-Wzy protein, while also restoring smooth Oag, exhibits greater functionality than StrepII-Wzy, in that less of the SR- LPS profile is visible (i.e., prominent detection of the 1st RU) (Figure 6.8). However this protein does not appear to have full functionality, as a wild-type level of modal chain length control was not restored. The strain expressing GFP$^+$-StrepII-Wzy protein exhibited the closest LPS profile to wild-type, i.e., a distinct banding of Oag modal chain length 11-17 RUs and wild-type levels of Oag with a single RU, with little or no SR-LPS.

It is noteworthy that StrepII-Wzy exhibits partial functionality, and exhibited significantly improved functionality with the addition of the GFP$^+$ tag. However, it seemed more likely that the addition of ~26 kDa (GFP protein) to the Wzy protein would have affected stability or diminished function. The data implicate the N-terminal end of the protein in Wzy function and also in Wzz interaction. The absence of modal length control for StrepII-Wzy suggests that interaction of this protein with Wzz is deficient, and is an intriguing and novel possibility. An interesting observation was the improved function of GFP$^+$-Wzy correlated with the presence of the StrepII tag between GFP$^+$ and Wzy, as GFP$^+$-StrepII-Wzy had similar activity to wild-type Wzy. It is possible that the folding of the protein or stability is improved, although this is speculative. The microscopy results indicate that in a small percentage of cells, GFP$^+$-Wzy was located in the periphery of the cell (section 5.12). However, due to the partial functionality of this protein, these results must be interpreted with caution. As discussed above, the lack of wild-type function exhibited by this tagged protein for both LPS profiling and colicin sensitivity assays indicates that the positioning within the cell may be artificial. Further attempts at constructing a fully functional and readily detectable tagged Wzy, preferably utilizing a COOH-terminal tag, are therefore required to empirically determine the location of Wzy within the cell. Future experiments investigating the N-terminal region of Wzy with mutagenesis may also provide further insight into the role this region of Wzy has in LPS Oag modal chain length determination.

CHAPTER SEVEN

7.13 CONCLUSION

Collectively, the results presented within this thesis show that many regions spanning Wzz_{SF} affect the function of the protein, and that mutations located in the central cavity of Wzz_{SF} result in longer LPS Oag modal chain length than wild-type. Cross-linking analyses show that mutants conferring wild-type or longer LPS Oag modal chain length form readily detectable higher order oligomers, and exhibit stable dimer formation compared to mutants that result in VS-type or random LPS Oag modal chain length. Using co-purification assays, this thesis showed that G to A mutations in TM2 of Wzz_{SF} result in the inability of this mutant protein to interact with wild-type Wzz_{SF}. In contrast, whilst exhibiting significantly less sequence similarity, Wzz_{ST} was capable of interacting with wild-type Wzz_{SF}. This study has also found that Wzz_{SF} appears to localise to the periphery of the cell in concentrated clusters, and that the N-terminal region of Wzy appears to have a significant effect on the Oag polymerization activity. We can speculate that the defect in wild-type Oag modal chain length control observed when the *wzy* mutant strain was complemented with StrepII-Wzy is a result of the inability of StrepII-Wzy to interact with Wzz. Site-directed mutagenesis of this region may be used to explore the Wzy:Wzz interaction in future studies.

REFERENCES

Abeyrathne PD, Lam JS (2007) Conditions that allow for effective transfer of membrane proteins onto nitrocellulose membrane in Western blots. *Can J Microbiol* **53:** 526-532

Albert MJ, Singh KV, Murray BE, Erlich J (1990) Molecular epidemiology of *Shigella* infection in Central Australia. *Epidemiol Infect* **105:** 51-57

Alexander DC, Valvano MA (1994) Role of the *rfe* gene in the biosynthesis of the *Escherichia coli* O7-specific lipopolysaccharide and other O-specific polysaccharides containing N-acetylglucosamine. *J Bacteriol* **176:** 7079-7084

Allaoui A, Sansonetti PJ, Parsot C (1992) MxiJ, a lipoprotein involved in secretion of *Shigella* Ipa invasins, is homologous to YscJ, a secretion factor of the *Yersinia* Yop proteins. *J Bacteriol* **174:** 7661-7669

Allaoui A, Sansonetti PJ, Parsot C (1993) MxiD, an outer membrane protein necessary for the secretion of the *Shigella flexneri* lpa invasins. *Mol Microbiol* **7:** 59-68

Allison GE, Verma NK (2000) Serotype-converting bacteriophages and O-antigen modification in *Shigella flexneri*. *Trends Microbiol* **8:** 17-23

Amor PA, Whitfield C (1997) Molecular and functional analysis of genes required for expression of group IB K antigens in *Escherichia coli*: characterization of the his-region containing gene clusters for multiple cell-surface polysaccharides. *Mol Microbiol* **26:** 145-161

Anderson MS, Raetz CR (1987) Biosynthesis of lipid A precursors in *Escherichia coli*. A cytoplasmic acyltransferase that converts UDP-N-acetylglucosamine to UDP-3-O-(R-3-hydroxymyristoyl)-N-acetylglucosamine. *J Biol Chem* **262:** 5159-5169

Ashbolt R, Givney R, Gregory JE, Hall G, Hundy R, Kirk M, McKay I, Meuleners L, Millard G, Raupach J, Roche P, Prasopa-Plaizier N, Sama MK, Stafford R, Tomaska N, Unicomb L, Williams C (2002) Enhancing foodborne disease surveillance across Australia in 2001: the OzFoodNet Working Group. *Commun Dis Intell* **26:** 375-406

Ashida H, Toyotome T, Nagai T, Sasakawa C (2007) *Shigella* chromosomal IpaH proteins are secreted via the type III secretion system and act as effectors. *Mol Microbiol* **63:** 680-693

Ashkenazi S, Cleary KR, Pickering LK, Murray BE, Cleary TG (1990) The association of Shiga toxin and other cytotoxins with the neurologic manifestations of shigellosis. *J Infect Dis* **161:** 961-965

Ashkenazi S, Dinari G, Zevulunov A, Nitzan M (1987) Convulsions in childhood shigellosis. Clinical and laboratory features in 153 children. *Am J Dis Child* **141:** 208-210

Bastin DA, Stevenson G, Brown PK, Haase A, Reeves PR (1993) Repeat unit polysaccharides of bacteria: a model for polymerization resembling that of ribosomes and fatty acid synthetase, with a novel mechanism for determining chain length. *Mol Microbiol* **7:** 725-734

Becker A, Kuster H, Niehaus K, Puhler A (1995) Extension of the *Rhizobium meliloti* succinoglycan biosynthesis gene cluster: identification of the *exsA* gene encoding an ABC transporter protein, and the *exsB* gene which probably codes for a regulator of succinoglycan biosynthesis. *Mol Gen Genet* **249:** 487-497

Bennish ML (1991) Potentially lethal complications of shigellosis. *Rev Infect Dis* **13 Suppl 4:** S319-324

Bennish ML, Harris JR, Wojtyniak BJ, Struelens M (1990) Death in shigellosis: incidence and risk factors in hospitalized patients. *J Infect Dis* **161:** 500-506

Bernardini ML, Mounier J, d'Hauteville H, Coquis-Rondon M, Sansonetti PJ (1989) Identification of *icsA*, a plasmid locus of *Shigella flexneri* that governs bacterial intra- and intercellular spread through interaction with F-actin. *Proc Natl Acad Sci U S A* **86:** 3867-3871

Blocker A, Gounon P, Larquet E, Niebuhr K, Cabiaux V, Parsot C, Sansonetti P (1999) The tripartite type III secreton of *Shigella flexneri* inserts IpaB and IpaC into host membranes. *J Cell Biol* **147:** 683-693

Blocker A, Jouihri N, Larquet E, Gounon P, Ebel F, Parsot C, Sansonetti P, Allaoui A (2001) Structure and composition of the *Shigella flexneri* "needle complex", a part of its type III secreton. *Mol Microbiol* **39:** 652-663

Bone RC (1993) Gram-negative sepsis: a dilemma of modern medicine. *Clin Microbiol Rev* **6:** 57-68

Bourdet-Sicard R, Rudiger M, Jockusch BM, Gounon P, Sansonetti PJ, Nhieu GT (1999) Binding of the *Shigella* protein IpaA to vinculin induces F-actin depolymerization. *Embo J* **18:** 5853-5862

Bray D, Robbins PW (1967) The direction of chain growth in *Salmonella anatum* O-antigen biosynthesis. *Biochem Biophys Res Commun* **28:** 334-339

Buysse JM, Stover CK, Oaks EV, Venkatesan M, Kopecko DJ (1987) Molecular cloning of invasion plasmid antigen (*ipa*) genes from *Shigella flexneri*: analysis of *ipa* gene products and genetic mapping. *J Bacteriol* **169:** 2561-2569

Cardona ST, Valvano MA (2005) An expression vector containing a rhamnose-inducible promoter provides tightly regulated gene expression in *Burkholderia cenocepacia*. *Plasmid* **54:** 219-228

Carter JA, Blondel CJ, Zaldivar M, Alvarez SA, Marolda CL, Valvano MA, Contreras I (2007) O-antigen modal chain length in *Shigella flexneri* 2a is growth-regulated through RfaH-mediated transcriptional control of the *wzy* gene. *Microbiology* **153:** 3499-3507

Chen Y, Smith MR, Thirumalai K, Zychlinsky A (1996) A bacterial invasin induces macrophage apoptosis by binding directly to ICE. *Embo J* **15:** 3853-3860

Chng SS, Gronenberg LS, Kahne D (2010a) Proteins required for lipopolysaccharide assembly in *Escherichia coli* form a transenvelope complex. *Biochemistry* **49:** 4565-4567

Chng SS, Ruiz N, Chimalakonda G, Silhavy TJ, Kahne D (2010b) Characterization of the two-protein complex in *Escherichia coli* responsible for lipopolysaccharide assembly at the outer membrane. *Proc Natl Acad Sci U S A* **107:** 5363-5368

Clark MA, Hirst BH (2002) Expression of junction-associated proteins differentiates mouse intestinal M cells from enterocytes. *Histochem Cell Biol* **118:** 137-147

Clementz T, Raetz CR (1991) A gene coding for 3-deoxy-D-manno-octulosonic-acid transferase in *Escherichia coli*. Identification, mapping, cloning, and sequencing. *J Biol Chem* **266:** 9687-9696

Coleman WG, Jr. (1983) The *rfaD* gene codes for ADP-L-glycero-D-mannoheptose-6-epimerase. An enzyme required for lipopolysaccharide core biosynthesis. *J Biol Chem* **258:** 1985-1990

Collins LV, Hackett J (1991) Molecular cloning, characterization, and nucleotide sequence of the *rfc* gene, which encodes an O-antigen polymerase of *Salmonella typhimurium*. *J Bacteriol* **173:** 2521-2529

Cornelis GR (2006) The type III secretion injectisome. *Nat Rev Microbiol* **4:** 811-825

Cossart P (2000) Actin-based motility of pathogens: the Arp2/3 complex is a central player. *Cell Microbiol* **2:** 195-205

Creeger ES, Rothfield LI (1979) Cloning of genes for bacterial glycosyltransferases. I. Selection of hybrid plasmids carrying genes for two glucosyltransferases. *J Biol Chem* **254:** 804-810

Cunneen MM, Reeves PR (2008) Membrane topology of the *Salmonella enterica* serovar *Typhimurium* Group B O-antigen translocase Wzx. *FEMS Microbiol Lett* **287:** 76-84

Cuthbertson L, Mainprize IL, Naismith JH, Whitfield C (2009) Pivotal roles of the outer membrane polysaccharide export and polysaccharide copolymerase protein families in export of extracellular polysaccharides in gram-negative bacteria. *Microbiol Mol Biol Rev* **73:** 155-177

Daniels, C (1999) Characterisation of proteins involved in *Shigella flexneri* O-antigen biosynthesis. PhD thesis

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

Daniels C, Vindurampulle C, Morona R (1998) Overexpression and topology of the *Shigella flexneri* O-antigen polymerase (Rfc/Wzy). *Mol Microbiol* **28:** 1211-1222

Doerrler WT, Gibbons HS, Raetz CR (2004) MsbA-dependent translocation of lipids across the inner membrane of *Escherichia coli*. *J Biol Chem* **279:** 45102-45109

DuPont HL, Levine MM, Hornick RB, Formal SB (1989) Inoculum size in shigellosis and implications for expected mode of transmission. *J Infect Dis* **159:** 1126-1128

Egile C, Loisel TP, Laurent V, Li R, Pantaloni D, Sansonetti PJ, Carlier MF (1999) Activation of the CDC42 effector N-WASP by the *Shigella flexneri* IcsA protein promotes actin nucleation by Arp2/3 complex and bacterial actin-based motility. *J Cell Biol* **146:** 1319-1332

Feldman MF, Marolda CL, Monteiro MA, Perry MB, Parodi AJ, Valvano MA (1999) The activity of a putative polyisoprenol-linked sugar translocase (Wzx) involved in *Escherichia coli* O antigen assembly is independent of the chemical structure of the O repeat. *J Biol Chem* **274:** 35129-35138

Formal SB, Gemski P, Baron LS, Labrec EH (1970) Genetic Transfer of *Shigella flexneri* Antigens to *Escherichia coli* K-12. *Infect Immun* **1:** 279-287

Franco AV, Liu D, Reeves PR (1998) The wzz (cld) protein in *Escherichia coli*: amino acid sequence variation determines O-antigen chain length specificity. *J Bacteriol* **180:** 2670-2675

Galan JE, Collmer A (1999) Type III secretion machines: bacterial devices for protein delivery into host cells. *Science* **284:** 1322-1328

Galloway SM, Raetz CR (1990) A mutant of *Escherichia coli* defective in the first step of endotoxin biosynthesis. *J Biol Chem* **265:** 6394-6402

Gebert A, Rothkotter HJ, Pabst R (1996) M cells in Peyer's patches of the intestine. *Int Rev Cytol* **167:** 91-159

Goldberg MB (2001) Actin-based motility of intracellular microbial pathogens. *Microbiol Mol Biol Rev* **65:** 595-626, table of contents

Goldberg MB, Barzu O, Parsot C, Sansonetti PJ (1993) Unipolar localization and ATPase activity of IcsA, a *Shigella flexneri* protein involved in intracellular movement. *J Bacteriol* **175:** 2189-2196

Grangeasse C, P. Doublet, and A. J. Cozzone (2002) Tyrosine phosphorylation of protein kinase Wzc from *Escherichia coli* K-12 occurs through a two-step process. *J Biol Chem* **277:** 7127-7135

Guan S, Verma NK (1998) Serotype conversion of a *Shigella flexneri* candidate vaccine strain via a novel site-specific chromosome-integration system. *FEMS Microbiol Lett* **166:** 79-87

Guo H, Lokko K, Zhang Y, Yi W, Wu Z, Wang PG (2006) Overexpression and characterization of Wzz of *Escherichia coli* O86:H2. *Protein Expr Purif* **48:** 49-55

Guo H, Yi W, Song JK, Wang PG (2008) Current understanding on biosynthesis of microbial polysaccharides. *Curr Top Med Chem* **8:** 141-151

Gupta A, Polyak CS, Bishop RD, Sobel J, Mintz ED (2004) Laboratory-confirmed shigellosis in the United States, 1989-2002: epidemiologic trends and patterns. *Clin Infect Dis* **38:** 1372-1377

Guzman LM, Belin D, Carson MJ, Beckwith J (1995) Tight regulation, modulation, and high-level expression by vectors containing the arabinose PBAD promoter. *J Bacteriol* **177:** 4121-4130

Hale TL (1991) Genetic basis of virulence in *Shigella* species. *Microbiol Rev* **55:** 206-224

Hardie KR, Seydel A, Guilvout I, Pugsley AP (1996) The secretin-specific, chaperone-like protein of the general secretory pathway: separation of proteolytic protection and piloting functions. *Mol Microbiol* **22:** 967-976

Harrington AT, Hearn PD, Picking WL, Barker JR, Wessel A, Picking WD (2003) Structural characterization of the N terminus of IpaC from *Shigella flexneri*. *Infect Immun* **71:** 1255-1264

Heinrichs DE, Yethon JA, Amor PA, Whitfield C (1998) The assembly system for the outer core portion of R1- and R4-type lipopolysaccharides of *Escherichia coli*. The R1 core-specific beta-glucosyltransferase provides a novel attachment site for O-polysaccharides. *J Biol Chem* **273:** 29497-29505

Hong M, Payne SM (1997) Effect of mutations in *Shigella flexneri* chromosomal and plasmid-encoded lipopolysaccharide genes on invasion and serum resistance. *Mol Microbiol* **24:** 779-791

Javadpour MM, Eilers M, Groesbeek M, Smith SO (1999) Helix packing in polytopic membrane proteins: role of glycine in transmembrane helix association. *Biophys J* **77:** 1609-1618

Jenei ZA, Borthwick K, Zammit VA, Dixon AM (2009) Self-association of transmembrane domain 2 (TM2), but not TM1, in carnitine palmitoyltransferase 1A: role of GXXXG(A) motifs. *J Biol Chem* **284:** 6988-6997

Jennison AV, Verma NK (2004) *Shigella flexneri* infection: pathogenesis and vaccine development. *FEMS Microbiol Rev* **28:** 43-58

Jouihri N, Sory MP, Page AL, Gounon P, Parsot C, Allaoui A (2003) MxiK and MxiN interact with the Spa47 ATPase and are required for transit of the needle components MxiH and MxiI, but not of Ipa proteins, through the type III secretion apparatus of *Shigella flexneri*. *Mol Microbiol* **49:** 755-767

Kim DW, Lenzen G, Page AL, Legrain P, Sansonetti PJ, Parsot C (2005) The *Shigella flexneri* effector OspG interferes with innate immune responses by targeting ubiquitin-conjugating enzymes. *Proc Natl Acad Sci U S A* **102:** 14046-14051

Klee SR, Tzschaschel BD, Timmis KN, Guzman CA (1997) Influence of different rol gene products on the chain length of *Shigella dysenteriae* type 1 lipopolysaccharide O antigen expressed by *Shigella flexneri* carrier strains. *J Bacteriol* **179:** 2421-2425

Kosek M, Bern C, Guerrant RL (2003) The global burden of diarrhoeal disease, as estimated from studies published between 1992 and 2000. *Bull World Health Organ* **81:** 197-204

Koster F, Levin J, Walker L, Tung KS, Gilman RH, Rahaman MM, Majid MA, Islam S, Williams RC, Jr. (1978) Hemolytic-uremic syndrome after shigellosis. Relation to endotoxemia and circulating immune complexes. *N Engl J Med* **298:** 927-933

Koster M, Bitter W, de Cock H, Allaoui A, Cornelis GR, Tommassen J (1997) The outer membrane component, YscC, of the Yop secretion machinery of *Yersinia enterocolitica* forms a ring-shaped multimeric complex. *Mol Microbiol* **26:** 789-797

Kotloff KL, Winickoff JP, Ivanoff B, Clemens JD, Swerdlow DL, Sansonetti PJ, Adak GK, Levine MM (1999) Global burden of *Shigella* infections: implications for vaccine development and implementation of control strategies. *Bull World Health Organ* **77:** 651-666

Kraehenbuhl JP, Neutra MR (2000) Epithelial M cells: differentiation and function. *Annu Rev Cell Dev Biol* **16:** 301-332

Kubori T, Matsushima Y, Nakamura D, Uralil J, Lara-Tejero M, Sukhan A, Galan JE, Aizawa SI (1998) Supramolecular structure of the *Salmonella typhimurium* type III protein secretion system. *Science* **280:** 602-605

Kurien BT, Scofield RH (2009) A brief review of other notable protein detection methods on blots. *Methods Mol Biol* **536:** 557-571

Labrec EH, Schneider H, Magnani TJ, Formal SB (1964) Epithelial Cell Penetration as an Essential Step in the Pathogenesis of Bacillary Dysentery. *J Bacteriol* **88:** 1503-1518

Lan R, Reeves PR (2002) *Escherichia coli* in disguise: molecular origins of *Shigella*. *Microbes Infect* **4:** 1125-1132

Larue K, Kimber MS, Ford R, Whitfield C (2009) Biochemical and structural analysis of bacterial O-antigen chain length regulator proteins reveals a conserved quaternary structure. *J Biol Chem* **284:** 7395-7403

Lehrer J, Vigeant KA, Tatar LD, Valvano MA (2007) Functional characterization and membrane topology of *Escherichia coli* WecA, a sugar-phosphate transferase initiating the biosynthesis of enterobacterial common antigen and O-antigen lipopolysaccharide. *J Bacteriol* **189:** 2618-2628

Lett MC, Sasakawa C, Okada N, Sakai T, Makino S, Yamada M, Komatsu K, Yoshikawa M (1989) *virG*, a plasmid-coded virulence gene of *Shigella flexneri*: identification of the virG protein and determination of the complete coding sequence. *J Bacteriol* **171:** 353-359

Levine MM, Kotloff KL, Barry EM, Pasetti MF, Sztein MB (2007) Clinical trials of *Shigella* vaccines: two steps forward and one step back on a long, hard road. *Nat Rev Microbiol* **5:** 540-553

Liu B, Knirel YA, Feng L, Perepelov AV, Senchenkova SN, Wang Q, Reeves PR, Wang L (2008) Structure and genetics of *Shigella* O antigens. *FEMS Microbiol Rev* **32:** 627-653

Liu D, Cole RA, Reeves PR (1996) An O-antigen processing function for Wzx (RfbX): a promising candidate for O-unit flippase. *J Bacteriol* **178:** 2102-2107

Ma B, Reynolds CM, Raetz CR (2008) Periplasmic orientation of nascent lipid A in the inner membrane of an *Escherichia coli* LptA mutant. *Proc Natl Acad Sci U S A* **105:** 13823-13828

Macnab RM (1999) The bacterial flagellum: reversible rotary propellor and type III export apparatus. *J Bacteriol* **181:** 7149-7153

Macpherson DF, Manning PA, Morona R (1995) Genetic analysis of the *rfbX* gene of *Shigella flexneri*. *Gene* **155:** 9-17

Magdalena J, Goldberg MB (2002) Quantification of *Shigella* IcsA required for bacterial actin polymerization. *Cell Motil Cytoskeleton* **51:** 187-196

Marolda CL, Tatar LD, Alaimo C, Aebi M, Valvano MA (2006) Interplay of the Wzx translocase and the corresponding polymerase and chain length regulator proteins in the translocation and periplasmic assembly of lipopolysaccharide o antigen. *J Bacteriol* **188:** 5124-5135

Masi M, Vuong P, Humbard M, Malone K, Misra R (2007) Initial steps of colicin E1 import across the outer membrane of *Escherichia coli*. *J Bacteriol* **189:** 2667-2676

Maurelli AT, Baudry B, d'Hauteville H, Hale TL, Sansonetti PJ (1985) Cloning of plasmid DNA sequences involved in invasion of HeLa cells by *Shigella flexneri*. *Infect Immun* **49:** 164-171

May KL, Morona R (2008) Mutagenesis of the *Shigella flexneri* autotransporter IcsA reveals novel functional regions involved in IcsA biogenesis and recruitment of host neural Wiscott-Aldrich syndrome protein. *J Bacteriol* **190:** 4666-4676

Meier-Dieter U, Barr K, Starman R, Hatch L, Rick PD (1992) Nucleotide sequence of the *Escherichia coli rfe* gene involved in the synthesis of enterobacterial common antigen. Molecular cloning of the *rfe-rff* gene cluster. *J Biol Chem* **267:** 746-753

Meier U, Mayer H (1985) Genetic location of genes encoding enterobacterial common antigen. *J Bacteriol* **163:** 756-762

Menard R, Sansonetti P, Parsot C, Vasselon T (1994) Extracellular association and cytoplasmic partitioning of the IpaB and IpaC invasins of *S. flexneri*. *Cell* **79:** 515-525

Menard R, Sansonetti PJ, Parsot C (1993) Nonpolar mutagenesis of the *ipa* genes defines IpaB, IpaC, and IpaD as effectors of *Shigella flexneri* entry into epithelial cells. *J Bacteriol* **175:** 5899-5906

Meredith TC, Woodard RW (2003) *Escherichia coli* YrbH is a D-arabinose 5-phosphate isomerase. *J Biol Chem* **278:** 32771-32777

Monack DM, Theriot JA (2001) Actin-based motility is sufficient for bacterial membrane protrusion formation and host cell uptake. *Cell Microbiol* **3:** 633-647

Morita-Ishihara T, Ogawa M, Sagara H, Yoshida M, Katayama E, Sasakawa C (2006) *Shigella* Spa33 is an essential C-ring component of type III secretion machinery. *J Biol Chem* **281:** 599-607

Morona, R (1982) The tolC locus of *Escherichia coli* K-12 : gene, protein and function. PhD thesis

Morona R, Daniels C, Van Den Bosch L (2003) Genetic modulation of *Shigella flexneri* 2a lipopolysaccharide O antigen modal chain length reveals that it has been optimized for virulence. *Microbiology* **149:** 925-939

Morona R, Mavris M, Fallarino A, Manning PA (1994) Characterization of the *rfc* region of *Shigella flexneri*. *J Bacteriol* **176:** 733-747

Morona R, Purins L, Tocilj A, Matte A, Cygler M (2009) Sequence-structure relationships in polysaccharide co-polymerase (PCP) proteins. *Trends Biochem Sci* **34:** 78-84

Morona R, Van Den Bosch L (2003) Multicopy *icsA* is able to suppress the virulence defect caused by the *wzz(SF)* mutation in *Shigella flexneri*. *FEMS Microbiol Lett* **221:** 213-219

Morona R, Van Den Bosch L, Daniels C (2000) Evaluation of Wzz/MPA1/MPA2 proteins based on the presence of coiled-coil regions. *Microbiology* **146 (Pt 1):** 1-4

Morona R, van den Bosch L, Manning PA (1995) Molecular, genetic, and topological characterization of O-antigen chain length regulation in *Shigella flexneri*. *J Bacteriol* **177:** 1059-1068

Mounier J, Vasselon T, Hellio R, Lesourd M, Sansonetti PJ (1992) *Shigella flexneri* enters human colonic Caco-2 epithelial cells through the basolateral pole. *Infect Immun* **60:** 237-248

Muhlradt P (1969) Biosynthesis of *Salmonella* lipopolysaccharide. The *in vitro* transfer of phosphate to the heptose moiety of the core. *Eur J Biochem* **11:** 241-248

Muller JM, Ziegler-Heitbrock HW, Baeuerle PA (1993) Nuclear factor kappa B, a mediator of lipopolysaccharide effects. *Immunobiology* **187:** 233-256

Murray GL, Attridge SR, Morona R (2003) Regulation of *Salmonella typhimurium* lipopolysaccharide O antigen chain length is required for virulence; identification of FepE as a second Wzz. *Mol Microbiol* **47:** 1395-1406

Naide Y, Nikaido H, Maekelae PH, Wilkinson RG, Stocker BA (1965) Semirough strains of *Salmonella*. *Proc Natl Acad Sci U S A* **53:** 147-153

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

Neutra MR, Frey A, Kraehenbuhl JP (1996) Epithelial M cells: gateways for mucosal infection and immunization. *Cell* **86:** 345-348

Niebuhr K, Giuriato S, Pedron T, Philpott DJ, Gaits F, Sable J, Sheetz MP, Parsot C, Sansonetti PJ, Payrastre B (2002) Conversion of PtdIns(4,5)P(2) into PtdIns(5)P by the *S.flexneri* effector IpgD reorganizes host cell morphology. *Embo J* **21:** 5069-5078

Niyogi SK (2005) Shigellosis. *J Microbiol* **43:** 133-143

Ogawa M, Yoshimori T, Suzuki T, Sagara H, Mizushima N, Sasakawa C (2005) Escape of intracellular *Shigella* from autophagy. *Science* **307:** 727-731

Okada N, Sasakawa C, Tobe T, Talukder KA, Komatsu K, Yoshikawa M (1991a) Construction of a physical map of the chromosome of *Shigella flexneri* 2a and the direct assignment of nine virulence-associated loci identified by Tn5 insertions. *Mol Microbiol* **5:** 2171-2180

Okada N, Sasakawa C, Tobe T, Yamada M, Nagai S, Talukder KA, Komatsu K, Kanegasaki S, Yoshikawa M (1991b) Virulence-associated chromosomal loci of *Shigella flexneri* identified by random Tn5 insertion mutagenesis. *Mol Microbiol* **5:** 187-195

Okamura N, Nagai T, Nakaya R, Kondo S, Murakami M, Hisatsune K (1983) HeLa cell invasiveness and O antigen of *Shigella flexneri* as separate and prerequisite attributes of virulence to evoke keratoconjunctivitis in guinea pigs. *Infect Immun* **39:** 505-513

Okuda J, Toyotome T, Kataoka N, Ohno M, Abe H, Shimura Y, Seyedarabi A, Pickersgill R, Sasakawa C (2005) *Shigella* effector IpaH9.8 binds to a splicing factor U2AF(35) to modulate host immune responses. *Biochem Biophys Res Commun* **333:** 531-539

Olive AJ, Kenjale R, Espina M, Moore DS, Picking WL, Picking WD (2007) Bile salts stimulate recruitment of IpaB to the *Shigella flexneri* surface, where it colocalizes with IpaD at the tip of the type III secretion needle. *Infect Immun* **75:** 2626-2629

Paiment A, Hocking J, Whitfield C (2002) Impact of phosphorylation of specific residues in the tyrosine autokinase, Wzc, on its activity in assembly of group 1 capsules in *Escherichia coli*. *J Bacteriol* **184:** 6437-6447

Parsot C (2005) *Shigella* spp. and enteroinvasive *Escherichia coli* pathogenicity factors. *FEMS Microbiol Lett* **252:** 11-18

Perdomo OJ, Cavaillon JM, Huerre M, Ohayon H, Gounon P, Sansonetti PJ (1994) Acute inflammation causes epithelial invasion and mucosal destruction in experimental shigellosis. *J Exp Med* **180:** 1307-1319

Prossnitz E, Nikaido K, Ulbrich SJ, Ames GF (1988) Formaldehyde and photoactivatable cross-linking of the periplasmic binding protein to a membrane component of the histidine transport system of *Salmonella typhimurium*. *J Biol Chem* **263:** 17917-17920

Purins L, Van Den Bosch L, Richardson V, Morona R (2008) Coiled-coil regions play a role in the function of the *Shigella flexneri* O-antigen chain length regulator WzzpHS2. *Microbiology* **154:** 1104-1116

Raetz CR (1990) Biochemistry of endotoxins. *Annu Rev Biochem* **59:** 129-170

Raetz CR (1993) Bacterial endotoxins: extraordinary lipids that activate eucaryotic signal transduction. *J Bacteriol* **175:** 5745-5753

Raetz CR, Reynolds CM, Trent MS, Bishop RE (2007) Lipid A modification systems in gram-negative bacteria. *Annu Rev Biochem* **76:** 295-329

Raetz CR, Whitfield C (2002) Lipopolysaccharide endotoxins. *Annu Rev Biochem* **71:** 635-700

Ranallo RT, Kaminski RW, George T, Kordis AA, Chen Q, Szabo K, Venkatesan MM (2010) Virulence, inflammatory potential, and adaptive immunity induced by *Shigella flexneri msbB* mutants. *Infect Immun* **78:** 400-412

Reid AN, Whitfield C (2005) Functional analysis of conserved gene products involved in assembly of *Escherichia coli* capsules and exopolysaccharides: evidence for molecular recognition between Wza and Wzc for colonic acid biosynthesis. *J Bacteriol* **187:** 5470-5481

Reizer J, Reizer A, Saier MH, Jr. (1992) A new subfamily of bacterial ABC-type transport systems catalyzing export of drugs and carbohydrates. *Protein Sci* **1:** 1326-1332

Robbins JR, Monack D, McCallum SJ, Vegas A, Pham E, Goldberg MB, Theriot JA (2001) The making of a gradient: IcsA (VirG) polarity in *Shigella flexneri*. *Mol Microbiol* **41:** 861-872

Ruiz N, Gronenberg LS, Kahne D, Silhavy TJ (2008) Identification of two inner-membrane proteins required for the transport of lipopolysaccharide to the outer membrane of *Escherichia coli*. *Proc Natl Acad Sci U S A* **105:** 5537-5542

Samuel G, Reeves P (2003) Biosynthesis of O-antigens: genes and pathways involved in nucleotide sugar precursor synthesis and O-antigen assembly. *Carbohydr Res* **338:** 2503-2519

Sandlin RC, Lampel KA, Keasler SP, Goldberg MB, Stolzer AL, Maurelli AT (1995) Avirulence of rough mutants of *Shigella flexneri*: requirement of O antigen for correct unipolar localization of IcsA in the bacterial outer membrane. *Infect Immun* **63:** 229-237

Sansonetti PJ (2001) Microbes and microbial toxins: paradigms for microbial-mucosal interactions III. Shigellosis: from symptoms to molecular pathogenesis. *Am J Physiol Gastrointest Liver Physiol* **280:** G319-323

Sansonetti PJ, Arondel J, Huerre M, Harada A, Matsushima K (1999) Interleukin-8 controls bacterial transepithelial translocation at the cost of epithelial destruction in experimental shigellosis. *Infect Immun* **67:** 1471-1480

Sansonetti PJ, Kopecko DJ, Formal SB (1982) Involvement of a plasmid in the invasive ability of *Shigella flexneri*. *Infect Immun* **35:** 852-860

Sansonetti PJ, Phalipon A (1999) M cells as ports of entry for enteroinvasive pathogens: mechanisms of interaction, consequences for the disease process. *Semin Immunol* **11:** 193-203

Sansonetti PJ, Ryter A, Clerc P, Maurelli AT, Mounier J (1986) Multiplication of *Shigella flexneri* within HeLa cells: lysis of the phagocytic vacuole and plasmid-mediated contact hemolysis. *Infect Immun* **51:** 461-469

Sasakawa C, Adler B, Tobe T, Okada N, Nagai S, Komatsu K, Yoshikawa M (1989) Functional organization and nucleotide sequence of virulence Region-2 on the large virulence plasmid in *Shigella flexneri* 2a. *Mol Microbiol* **3:** 1191-1201

Sasakawa C, Kamata K, Sakai T, Makino S, Yamada M, Okada N, Yoshikawa M (1988) Virulence-associated genetic regions comprising 31 kilobases of the 230-kilobase plasmid in *Shigella flexneri* 2a. *J Bacteriol* **170:** 2480-2484

Sasakawa C, Komatsu K, Tobe T, Suzuki T, Yoshikawa M (1993) Eight genes in region 5 that form an operon are essential for invasion of epithelial cells by *Shigella flexneri* 2a. *J Bacteriol* **175:** 2334-2346

Scholz O, Thiel A, Hillen W, Niederweis M (2000) Quantitative analysis of gene expression with an improved green fluorescent protein *Eur J Biochem* **267:** 1565-1570

Schuch R, Maurelli AT (2001a) MxiM and MxiJ, base elements of the Mxi-Spa type III secretion system of *Shigella*, interact with and stabilize the MxiD secretin in the cell envelope. *J Bacteriol* **183:** 6991-6998

Schuch R, Maurelli AT (2001b) Spa33, a cell surface-associated subunit of the Mxi-Spa type III secretory pathway of *Shigella flexneri*, regulates Ipa protein traffic. *Infect Immun* **69:** 2180-2189

Senes A, Engel DE, DeGrado WF (2004) Folding of helical membrane proteins: the role of polar, GxxxG-like and proline motifs. *Curr Opin Struct Biol* **14:** 465-479

Shaner NC, Campbell RE, Steinbach PA, Giepmans BN, Palmer AE, Tsien RY (2004) Improved monomeric red, orange and yellow fluorescent proteins derived from *Discosoma* sp. red fluorescent protein. *Nat Biotechnol* **22:** 1567-1572

Sirisena DM, Brozek KA, MacLachlan PR, Sanderson KE, Raetz CR (1992) The *rfaC* gene of *Salmonella typhimurium*. Cloning, sequencing, and enzymatic function in heptose transfer to lipopolysaccharide. *J Biol Chem* **267:** 18874-18884

Sirisena DM, MacLachlan PR, Liu SL, Hessel A, Sanderson KE (1994) Molecular analysis of the *rfaD* gene, for heptose synthesis, and the *rfaF* gene, for heptose transfer, in lipopolysaccharide synthesis in *Salmonella typhimurium*. *J Bacteriol* **176:** 2379-2385

Skoudy A, Mounier J, Aruffo A, Ohayon H, Gounon P, Sansonetti P, Tran Van Nhieu G (2000) CD44 binds to the *Shigella IpaB* protein and participates in bacterial invasion of epithelial cells. *Cell Microbiol* **2:** 19-33

Sperandeo P, Cescutti R, Villa R, Di Benedetto C, Candia D, Deho G, Polissi A (2007) Characterization of *lptA* and *lptB*, two essential genes implicated in lipopolysaccharide transport to the outer membrane of *Escherichia coli*. *J Bacteriol* **189:** 244-253

Sperandeo P, Deho G, Polissi A (2009) The lipopolysaccharide transport system of Gram-negative bacteria. *Biochim Biophys Acta* **1791:** 594-602

Sperandeo P, Lau FK, Carpentieri A, De Castro C, Molinaro A, Deho G, Silhavy TJ, Polissi A (2008) Functional analysis of the protein machinery required for transport of lipopolysaccharide to the outer membrane of *Escherichia coli*. *J Bacteriol* **190:** 4460-4469

Stevenson G, Kessler A, Reeves PR (1995) A plasmid-borne O-antigen chain length determinant and its relationship to other chain length determinants. *FEMS Microbiol Lett* **125:** 23-30

Suits MD, Sperandeo P, Deho G, Polissi A, Jia Z (2008) Novel structure of the conserved gram-negative lipopolysaccharide transport protein A and mutagenesis analysis. *J Mol Biol* **380:** 476-488

Suzuki T, Lett MC, Sasakawa C (1995) Extracellular transport of VirG protein in *Shigella*. *J Biol Chem* **270:** 30874-30880

Suzuki T, Miki H, Takenawa T, Sasakawa C (1998) Neural Wiskott-Aldrich syndrome protein is implicated in the actin-based motility of *Shigella flexneri*. *Embo J* **17:** 2767-2776

Suzuki T, Mimuro H, Suetsugu S, Miki H, Takenawa T, Sasakawa C (2002) Neural Wiskott-Aldrich syndrome protein (N-WASP) is the specific ligand for *Shigella* VirG among the WASP family and determines the host cell type allowing actin-based spreading. *Cell Microbiol* **4:** 223-233

Suzuki T, Saga S, Sasakawa C (1996) Functional analysis of *Shigella* VirG domains essential for interaction with vinculin and actin-based motility. *J Biol Chem* **271:** 21878-21885

Tamano K, Aizawa S, Katayama E, Nonaka T, Imajoh-Ohmi S, Kuwae A, Nagai S, Sasakawa C (2000) Supramolecular structure of the *Shigella* type III secretion machinery: the needle part is changeable in length and essential for delivery of effectors. *Embo J* **19:** 3876-3887

Tamano K, Katayama E, Toyotome T, Sasakawa C (2002) *Shigella* Spa32 is an essential secretory protein for functional type III secretion machinery and uniformity of its needle length. *J Bacteriol* **184:** 1244-1252

Tocilj A, Munger C, Proteau A, Morona R, Purins L, Ajamian E, Wagner J, Papadopoulos M, Van Den Bosch L, Rubinstein JL, Fethiere J, Matte A, Cygler M (2008) Bacterial polysaccharide co-polymerases share a common framework for control of polymer length. *Nat Struct Mol Biol* **15:** 130-138

Tran, E (2008) Molecular characterisation of *Shigella flexneri* outer membrane protease IcsP. PhD thesis

Tran Van Nhieu G, Caron E, Hall A, Sansonetti PJ (1999) IpaC induces actin polymerization and filopodia formation during *Shigella* entry into epithelial cells. *Embo J* **18:** 3249-3262

Van den Bosch L, Manning PA, Morona R (1997) Regulation of O-antigen chain length is required for *Shigella flexneri* virulence. *Mol Microbiol* **23:** 765-775

Van den Bosch L, Morona R (2003) The actin-based motility defect of a *Shigella flexneri rmlD* rough LPS mutant is not due to loss of IcsA polarity. *Microb Pathog* **35:** 11-18

van den Broek JM, Roy SK, Khan WA, Ara G, Chakraborty B, Islam S, Banu B (2005) Risk factors for mortality due to shigellosis: a case-control study among severely-malnourished children in Bangladesh. *J Health Popul Nutr* **23:** 259-265

van der Goot FG, Tran van Nhieu G, Allaoui A, Sansonetti P, Lafont F (2004) Rafts can trigger contact-mediated secretion of bacterial effectors via a lipid-based mechanism. *J Biol Chem* **279:** 47792-47798

van Pelt W, de Wit MA, Wannet WJ, Ligtvoet EJ, Widdowson MA, van Duynhoven YT (2003) Laboratory surveillance of bacterial gastroenteric pathogens in The Netherlands, 1991-2001. *Epidemiol Infect* **130:** 431-441

Veenendaal AK, Sundin C, Blocker AJ (2009) Small-molecule type III secretion system inhibitors block assembly of the *Shigella* type III secreton. *J Bacteriol* **191:** 563-570

Venkatesan MM, Buysse JM, Oaks EV (1992) Surface presentation of *Shigella flexneri* invasion plasmid antigens requires the products of the spa locus. *J Bacteriol* **174:** 1990-2001

Vimont S, Dumontier S, Escuyer V, Berche P (1997) The *rfaD* locus: a region of rearrangement in Vibrio cholerae O139. *Gene* **185:** 43-47

Vincent C, Doublet P, Grangeasse C, Vaganay E, Cozzone AJ, Duclos B (1999) Cells of *Escherichia coli* contain a protein-tyrosine kinase, Wzc, and a phosphotyrosine-protein phosphatase, Wzb. *J Bacteriol* **181:** 3472-3477

von Seidlein L, Kim DR, Ali M, Lee H, Wang X, Thiem VD, Canh do G, Chaicumpa W, Agtini MD, Hossain A, Bhutta ZA, Mason C, Sethabutr O, Talukder K, Nair GB, Deen JL, Kotloff K, Clemens J (2006) A multicentre study of *Shigella* diarrhoea in six Asian countries: disease burden, clinical manifestations, and microbiology. *PLoS Med* **3:** e353

Watarai M, Funato S, Sasakawa C (1996) Interaction of Ipa proteins of *Shigella flexneri* with alpha5beta1 integrin promotes entry of the bacteria into mammalian cells. *J Exp Med* **183:** 991-999

West NP, Sansonetti P, Mounier J, Exley RM, Parsot C, Guadagnini S, Prevost MC, Prochnicka-Chalufour A, Delepierre M, Tanguy M, Tang CM (2005) Optimization of virulence functions through glucosylation of *Shigella* LPS. *Science* **307:** 1313-1317

Whitfield C (2006) Biosynthesis and assembly of capsular polysaccharides in *Escherichia coli*. *Annu Rev Biochem* **75:** 39-68

Woodward R, Yi W, Li L, Zhao G, Eguchi H, Sridhar PR, Guo H, Song JK, Motari E, Cai L, Kelleher P, Liu X, Han W, Zhang W, Ding Y, Li M, Wang PG (2010) *In vitro* bacterial polysaccharide biosynthesis: defining the functions of Wzy and Wzz. *Nat Chem Biol* **6:** 418-423

Wyckoff TJ, Lin S, Cotter RJ, Dotson GD, Raetz CR (1998) Hydrocarbon rulers in UDP-N-acetylglucosamine acyltransferases. *J Biol Chem* **273:** 32369-32372

Yao Z, Valvano MA (1994) Genetic analysis of the O-specific lipopolysaccharide biosynthesis region (*rfb*) of *Escherichia coli* K-12 W3110: identification of genes that confer group 6 specificity to *Shigella flexneri* serotypes Y and 4a. *J Bacteriol* **176:** 4133-4143

Yethon JA, Heinrichs DE, Monteiro MA, Perry MB, Whitfield C (1998) Involvement of *waaY*, *waaQ*, and *waaP* in the modification of *Escherichia coli* lipopolysaccharide and their role in the formation of a stable outer membrane. *J Biol Chem* **273:** 26310-26316

Yethon JA, Whitfield C (2001) Lipopolysaccharide as a target for the development of novel therapeutics in gram-negative bacteria. *Curr Drug Targets Infect Disord* **1:** 91-106

Yoshida S, Handa Y, Suzuki T, Ogawa M, Suzuki M, Tamai A, Abe A, Katayama E, Sasakawa C (2006) Microtubule-severing activity of *Shigella* is pivotal for intercellular spreading. *Science* **314:** 985-989

Yoshida S, Katayama E, Kuwae A, Mimuro H, Suzuki T, Sasakawa C (2002) *Shigella* deliver an effector protein to trigger host microtubule destabilization, which promotes Rac1 activity and efficient bacterial internalization. *Embo J* **21:** 2923-2935

Young K, Silver LL, Bramhill D, Cameron P, Eveland SS, Raetz CR, Hyland SA, Anderson MS (1995) The *envA* permeability/cell division gene of *Escherichia coli* encodes the second enzyme of lipid A biosynthesis. UDP-3-O-(R-3-hydroxymyristoyl)-N-acetylglucosamine deacetylase. *J Biol Chem* **270:** 30384-30391

Zenk SF, Stabat D, Hodgkinson JL, Veenendaal AK, Johnson S, Blocker AJ (2007) Identification of minor inner-membrane components of the *Shigella* type III secretion system 'needle complex'. *Microbiology* **153:** 2405-2415

Zhang J, Liu H, Zhang X, Yang J, Yang F, Yang G, Shen Y, Hou Y, Jin Q (2003) Complete DNA sequence and gene analysis of the virulence plasmid pCP301 of *Shigella flexneri* 2a. *Sci China C Life Sci* **46:** 513-521