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

ABSTRACT i
DECLARATION iii
ACKNOWLEDGEMENTS iv
PUBLICATIONS vi
LIST OF ABBREVIATIONS vii
CHAPTER ONE 1
INTRODUCTION 1
1.1 SHIGELLA 1
1.1.1 SHIGELLOSIS 1
1.1.2 EPIDEMIOLOGY 2
1.1.3 *S. FLEXNERI* INVASION 3
1.2 *S. FLEXNERI* VIRULENCE 4
1.2.1 LARGE VIRULENCE PLASMID 4
1.2.2 MXI-SPA TYPE III SECRETION SYSTEM 5
1.2.3 IPA, IPG P ROTEINS AND OTHER EF FECTOR PROTEINS SE CRETED B Y THE TTSS 6
1.2.4 ICSA AND ACTIN-BASED MOTILITY 8
1.3 LIPOPOLYSACCHARIDES 8
1.3.1 *SHIGELLA* LPS AND VIRULENCE 8
1.3.2 MODAL CHAIN LENGTH AND VIRULENCE 9
1.3.3 STRUCTURE 10
1.3.4 LIPID A AND CORE 11
1.3.5 O ANTIGEN GENETICS AND BIOSYNTHESIS 13
1.3.6 O ANTIGEN PROCESSING 15
1.3.7 LPS EXPORT 17
1.4 POLYSACCHARIDE CO-POLYMERASES (PCPS) 18
1.4.1 THE POLYSACCHARIDE CO-POLYMERASE (PCP) FAMILY 18
1.4.2 MODAL CHAIN LENGTH REGULATION BY PCPS 19
1.4.3 3D STRUCTURES OF PCP PROTEINS 22
1.5 PROPOSED MODELS OF WZZ MECHANISM 23
1.6 STUDY RATIONALE 25
CHAPTER TWO 26
MATERIALS AND METHODS 26
2.1 BACTERIAL STRAINS USED IN THIS STUDY 26
2.2 BACTERIAL GROWTH CONDITIONS 26
2.2.1 LIQUID MEDIA GROWTH CONDITIONS 26
2.2.2 SOLID MEDIA GROWTH CONDITIONS 26
2.3 CHEMICALS AND REAGENTS 27
2.4 DNA PREPARATION AND MANIPULATION 27
2.4.1 RESTRICTION ENZYME DIGESTS AND LIGATION REACTIONS 27
2.4.2 AGAROSE GEL ELECTROPHORESIS 28
2.4.3 DNA PREPARATION USING A KIT 28
2.4.4 PREPARATION OF BOILED CELL LYSATES FOR PCR 28
2.5 POLYMERASE CHAIN REACTION (PCR) 29
2.5.1 GENERAL PCR 29
2.5.2 DNA SEQUENCING 29
2.6 PREPARATION OF COMPETENT CELLS 30
2.6.1 CHEMICALLY COMPETENT CELLS 30
2.6.2 ELECTROCOMPETENT CELLS 30
2.7 BACTERIAL TRANSFORMATION 30
2.7.1 HEAT SHOCK TRANSFORMATION 30
2.7.2 ELECTROPORATION 31
2.7.3 CONJUGATION 31

2.8 IN-FRAME LINKER MUTAGENESIS 32
2.9 3D STRUCTURAL IMAGES 32
2.10 PROTEIN TECHNIQUES 33
2.10.1 SDS POLYACRYLAMIDE GEL ELECTROPHORESIS (SDS-PAGE) 33
2.10.2 WESTERN TRANSFER AND IMMUNOBLOTTING 33
2.10.3 OVER-EXPRESSION AND PURIFICATION OF HIS_6-TAGGED PROTEIN 34
2.10.3.1 IPTG INDUCTION OF HIS_6-TAGGED PROTEINS 34
2.10.3.2 HIS_6-TAGGED PROTEIN PURIFICATION WITH NI-NTA RESIN 34
2.10.4 OVER-EXPRESSION OF FLAG-TAGGED PROTEINS 35
2.10.5 OVEREXPRESSION OF STREPII AND GFP^+-TAGGED PROTEINS 35
2.10.6 FORMALDEHYDE CROSS-LINKING 36
2.11 LIPOPOLYSACCHARIDE TECHNIQUES 37
2.11.1 PREPARATION OF LPS SAMPLES 37
2.11.2 ANALYSIS OF LPS BY SILVER-STAINED SDS-PAGE 37
2.12 COLICIN SENSITIVITY ASSAY 38
2.13 FLUORESCENT MICROSCOPY TECHNIQUES 38
CHAPTER THREE 40
IN-FRAME LINKER MUTAGENESIS OF WZZ_{SF} 40
3.1 INTRODUCTION 40
3.2 LINKER MUTAGENESIS 41
3.3 LPS MUTANT PROFILES 42
3.4 EXPRESSION OF WZZ_i MUTANT PROTEINS 43
3.5 COLICIN SENSITIVITY ASSAY 44
3.6 CHEMICAL CROSS-LINKING OF WZZ_i MUTANT PROTEINS 45
3.7 STABILITY OF WZZ_{SF} WILD-TYPE AND WZZ_i DIMERS 47
3.8 MAPPING OF WZZ_i INSERTIONS ONTO PCP 3D STRUCTURES 47
3.8.1 LOCATION OF WZZ_i INSERTIONS MAPPED ONTO MONOMERIC PCP STRUCTURES 48
3.8.2 LOCATION OF INSERTIONS MAPPED TO OLIGOMERIC STRUCTURES 49
3.9 SUMMARY 51
CHAPTER FOUR 52
WZZ:WZZ PROTEIN INTERACTION STUDY 52
4.1 INTRODUCTION 52
4.2 THE ASSESSMENT OF MUTANT WZZ GENETIC DOMINANCE 54
4.2.1 MUTATIONS K267N, M32T/I35C AND MUTATIONS IN TM2 54
4.2.2 COMPLEMENTATION WITH WZZ_{ST} AND WZZ_{ST}/WZZ_{SF} HYBRID PROTEINS, AND WZZ_{K31A} 55
4.3 CONSTRUCTION PLASMIDS ENCODING FLAG-TAGGED WZZ 57
4.4 COMPLEMENTATION OF LPS OAG MODAL LENGTH BY FLAG-WZZ_{SF} 58
4.5 CO-EXPRESSION OF pBAD-BASED WZZ PROTEINS WITH WILD-TYPE WZZ_{SF} 59
4.6 DETECTION OF FLAG-TAGGED WZZ PROTEINS 60
4.7 CO-PURIFICATION ASSAY 62
4.7.1 PURIFICATION OF HIS_6-WZZ_{SF} 62
4.7.2 CO-PURIFICATION OF HIS_6-WZZ_{SF} AND FLAG-WZZ_{SF} 63
4.8 ANALYSIS OF LPS PRODUCED BY STRAINS CO-EXPRESSING WZZi CLASS II MUTANTS AND WZZ_{SF} WILD-TYPE 64
4.9 ANALYSIS OF LPS PRODUCED BY STRAINS CO-EXPRESSING WZZi CLASS V MUTANTS AND WZZ_{SF} WILD-TYPE 65
4.10 SUMMARY 66
CHAPTER FIVE 68
STUDIES ON THE WZZ FLUORESCENT FUSION PROTEIN 68
5.1 INTRODUCTION 68

5.2 CONSTRUCTION OF mCHERRY-WZZ_{SF} 68
5.3 COMPLEMENTATION BY mCHERRY-WZZ_{SF} 69
5.4 mCHERRY-WZZ LOCALISATION 70
5.5 SUMMARY 71
CHAPTER SIX 72
STUDIES ON WZY FUSION PROTEINS 72
6.1 INTRODUCTION 72
6.2 CONSTRUCTION OF pSTREPII-WZY_{SF} 72
6.3 ANALYSIS OF LPS COMPLEMENTATION BY STREPII-WZY_{SF} 73
6.4 DETECTION OF STREP-WZY_{SF} BY WESTERN IMMUNOBLOTTING 74
6.5 CONSTRUCTION OF GFP^+-TAGGED WZY PROTEINS 75
6.5.1 CONSTRUCTION OF pGFP^+-STREPII-WZY 75
6.5.2 CONSTRUCTION OF pGFP^+-WZY 76
6.6 LPS PHENOTYPES OF GFP^+-TAGGED WZY PROTEINS 76
6.7 COLICIN SENSITIVITY OF STRAINS EXPRESSING TAGGED WZY PROTEINS 77
6.8 DETECTION OF GFP^+-WZY PROTEINS 78
6.9 LOCALISATION OF GFP^+-TAGGED WZY PROTEINS 79
6.10 SUMMARY 80
CHAPTER SEVEN 81
DISCUSSION 81
7.1 INTRODUCTION 81
7.2 RESULTS 82
7.3 CORRELATION OF WZZi LOCATION AND FUNCTION 84
7.4 WZZi INSERTIONS MAPPED TO MONOMERIC AND OLIGOMERIC STRUCTURES 85
7.5 PROTEIN DETECTION 88
7.6 SUSCEPTIBILITY TO COLICIN E2 89
7.7 CROSSLINKING ANALYSES 90
7.8 CO-EXPRESSION STUDIES OF FLAG-WZZ AND HIS_6-WZZ 91
7.9 COPURIFICATION OF WZZ PROTEINS AND WZZ::WZZ INTERACTION 93
7.10 VS-CONFERRING WZZi CO-EXPRESSION WITH WILD-TYPE WZZ, AND THEIR FUNCTIONAL ACTIVITY 96
7.11 SUBCELLULAR LOCALISATION OF WZZ AND WZY 98
7.12 PHENOTYPIC VARIATION BETWEEN TAGGED WZY PROTEINS 99
7.13 CONCLUSION 101
REFERENCES 102

ABSTRACT

In *Shigella flexneri*, Wzz_{SF} determines the lipopolysaccharides (LPS) O antigen (Oag) modal chain length of 11-17 repeat units (RUs). Wzz_{SF} has two transmembrane regions, the N-terminal TM1 and the C-terminal TM2, and previous studies have shown that the TM2 region is important in Wzz function. The mechanism of Oag modal chain length regulation has not been revealed. Previous studies have probed the structure-function relationship of Wzz_{SF} and have suggested that function is a result of the overall structure and not one particular region. Genetic evidence suggests that Wzz may form a complex with other Oag processing proteins, possibly with the Wzy polymerase.

This thesis describes in-frame linker mutagenesis of Wzz_{SF}, and five classes of Wzz insertion (Wzz_i) mutants with 5 amino acid insertions were identified: Class I (non-functional), Class II (conferred very short (VS) Oag chain length, 2-10 RUs), Class III (8-14 RUs), Class IV (11-19 RUs), and Class V (16-25 RUs). The susceptibility of strains expressing Wzz_i mutants to colicin E2 was investigated, and a correlation between random/VS LPS Oag modal chain length, and susceptibility to colicin E2 was found. Chemical cross-linking analyses were conducted to assess Wzz oligomerisation and showed that high molecular weight proteins were easily detected in wild-type Wzz_{SF} and mutants from Classes V, but not easily detected in Classes II and III, and only Wzz_{SF}, Wzz_i Class IV and Class V dimers were detected after heating to 100°C and in the presence of SDS, suggesting wild-type/longer LPS Oag modal chain length may be dependent on dimer stability.

The involvement of the TM2 region in Wzz:Wzz interactions was also investigated. *S. typhimurium* Wzz (Wzz_{ST}) confers an LPS Oag modal chain length of Long-type (L-type) 19-30 RUs, and Wzz_{ST} and Wzz_{SF} have identical residues in the conserved residues of their TM2 regions. Wzz_{ST} expression in Wzz deficient *S. flexneri* strain confers an L-type LPS Oag

modal chain length, however co-expression of Wzz_{ST} with Wzz_{SF} resulted in mono-modal LPS Oag modal chain length. A previously constructed Wzz mutant which had two G305A/G311A substitutions in the TM2 region confers a VS LPS Oag modal chain length (3-8 RUs), however co-expression of $Wzz_{G305A/G311A}$ with wild-type Wzz_{SF} resulted in LPS with a bimodal Oag chain length distribution (both wild-type 11-17 RUs, and VS-type 3-8 RUs). Strains expressing His_6-Wzz_{SF} and FLAG-tagged version of Wzz_{ST}, Wzz_{SF} and $Wzz_{G305A/G311A}$ were used in co-purification assays. Purified His_6-Wzz_{SF} and had FLAG-Wzz_{ST} was in the eluted fraction, demonstrating that FLAG-Wzz_{ST} interacted with His_6-Wzz_{SF}. However when His_6-Wzz_{SF} was purified from a strain co-expressing FLAG-$Wzz_{G305A/G311A}$, very low amounts of the latter were detected in the elution fraction, indicating that His_6-Wzz_{SF} interacted very poorly with FLAG-$Wzz_{G305A/G311A}$. These data implicate residues G305 and G311 in Wzz:Wzz interactions.

A fusion of the red fluorescent protein mCherry to Wzz_{SF} was performed and showed the localisation of this protein was in the periphery regions of the cell, as determined by epi-fluorescence microscopy. Wzy was tagged with StrepTag-II at its amino terminal end; the resulting StrepII-Wzy protein was able to complement a *wzy* mutation, however the smooth LPS produced lacked any Oag modal chain length control. Two GFP^+-tagged Wzy fusion proteins were constructed, pGFP^+-StrepII-Wzy, (with the *strepII* linker region between gfp^+ and *wzy*), and pGFP^+-Wzy (*strepII* excised). GFP^+-StrepII-Wzy had near wild-type Wzy functionality, and the complemented *wzy* mutant exhibited the wild-type trait of resistance to the lethal action of colicin E2, but GFP^+-StrepII-Wzy could not be detected by epi-fluorescence microscopy, and was poorly detectable by Western immunoblotting. GFP^+-Wzy was partially functional and the complemented *wzy* mutant was susceptible to colicin E2, like the StrepII-Wzy complemented *wzy* mutant. GFP^+-Wzy could be detected by both Western immunoblotting and epi-fluorescence microscopy. These results implicate the N-terminal region of Wzy in Oag modal chain length determination.

DECLARATION

This work contains no material which has been accepted for the award of any other degree or diploma in any university or other tertiary institution and, to the best of my knowledge and belief, contains no material previously published or written by another person, except where due reference has been made in the text.

I give consent to this copy of my thesis, when deposited in the University library, being made available in all forms of media, subject to the provisions of the Copyright Act 1968.

Magdalene Papadopoulos

ACKNOWLEDGEMENTS

Firstly, I'd like to thank my supervisor, Associate Professor Renato Morona. Thank you for giving me the opportunity to learn from you, to grow and develop - and not just as a scientist. I am so grateful for your patience and understanding, and you will always have my sincerest affection, respect and admiration. Also, I would like to thank Luisa Van Den Bosch for helping me with techniques in the lab and for always cheering me up with that infectious smile of hers.

To other members of the lab including Gerald Murray, Leanne Purins, Alistair Standish, Kerrie Grabowicz, Marcin Grabowicz and Tony Focareta, thank you for the advice and expertise you provided, I sincerely appreciate it. To all past Honours students, you made the lab a joy to be a part of, and I thank you for that. I would also like to thank Chris Wong, Jamie Botten and the general staff, particularly Ros Hammond, Genny Drexel, Shirley Coad, Sergei Volgin, Shelley Pezy, Garry Penney and John Mackrill. Getting to know you has been a highlight of my time here. I would especially like to thank Martin Lennon for being one of the best friends I have encountered, and for introducing me to the very awesome hobby of running. We'll make it to another City to Bay, I'm sure. Special thanks also go to Min Teh, Damien Chong, Mabel Lum, Kim LeMessurier, Jai Denton, Sophia Tan, Campbell Strong, James 'Jiminion' Byrne, Georget Reaiche, Thomas Tu and particularly Elizabeth Tran, who has been more like a sister rather than a fellow student. I cannot express how valuable your friendship is to me. I will never forget any of you, and my apologies for the Saddle Club moment, but I think the bonds made by our experiences at MLS will keep us friends forever. Thanks also to the Paton lab for being excellent neighbours and fantastic company during morning tea. Thanks to Uwe Stroeher in particular, for being the go-to-guy. I think I'll put you on speed dial once I'm gone, just to be sure I can still ask questions from afar!

Very importantly, I would like to thank Damon Tumes. Thank you for your love, your support, and for being there when I needed you the most.

And finally, I would like to thank my amazing family. I am grateful to you for all your support and unwavering faith in me, not just as a student, but as a person. I love you all very much and could not have done this without you – of that, I am sure. Σας αγαπώ.

PUBLICATIONS

Tocilj, A ., Munger, C ., P roteau, A., M orona, R ., P urins, L ., A jamian, E ., Wagner, J ., **Papadopoulos, M**., Van Den Bosch, L., Rubinstein, J.L., Fethiere, J., Matte, A. and Cygler, M. (2008) Bacterial polysaccharide co-polymerases share a common framework for control of polymer length. *Nat Struct Mol Biol* **15**, 130-138.

Papadopoulos, M. and Morona, R. (2010) Mutagenesis and chemical cross-linking suggest that Wzz dimer stability and oligomerization affect lipopolysaccharide O-antigen modal chain length control. *J Bacteriol* **192**, 3385-3393.

LIST OF ABBREVIATIONS

~	approximately
°C	degree
%	percentage
#	number
α	alpha
β	beta
γ	gamma
λ	lambda
µg; µl; µm	microgram (s); microliter (s); micrometre (s)
aa	amino acid
3D	3-dimensional
Å	Amstrong
ACP	acyl carrier protein
ABM	actin based motility
Amp	ampicillin
Arp	actin-related protein
ATP	adenosine triphosphate
bp	base pairs
C-terminal	carboxyl-terminal
Cat. number	catalogue number
Cml	chloramphenicol
CPS	capsular polysaccharides
DDM	D-dodecyl β-D maltoside
DMF	dimethylformamide

DNA	deoxyribonucleic acid
dNTP	deoxynucleoside triphosphate
ECA	enterobacterial common antigen
EDTA	ethylene diamine tetra-acetic acid
EPS	exopolysaccharides
FAE	follicular associated epithelium
FITC	fluorescein isothiocyanate
FP	fluorescent protein
GlcNAc	N-acetylglucosamine
h; min; sec	hour (s); minutes (s); seconds (s)
Hep	L-*glycero*-D-*manno*-heptose
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
His_6	6x histidine
IM	inner membrane
Ipa	invasion plasmid antigens
Ipg	invasion plasmid gene
IPTG	isopropyl-β-D-thiogalactopyranoside
IS	insertion sequences
I-type	intermediate type
Kan	kanamycin
kb	kilobase pairs
kDa	kilodaltons
Kdo	3-deoxy-D-*manno*-oct-2-ulosonic acid
L	litres
Lab	laboratory
LB	Luria-bertani
LPS	lipopolysaccharides

Lpt	lipopolysaccharides transport
L-type	long type
Lys	lysine
M; mM	molar; millimolar (s)
mA	milli-amps
mAb	monoclonal antibody
M-cells	Membraneous epithelial cells
mg; ml	milligram (s); millilitre (s)
MOPS	3-(N-Morpholino)-propanesulfonic acid
MQ	MilliQ water
mxi-spa	membrane expression of Ipas-surface presentation of antigens
N-terminal	amino terminal
N-WASP	neural Wiskott-Aldrich syndrome protein
NEB	New England Biolabs
Ni-NTA	nickel-charged agarose
nm	nanometre
nt	nucleotide
Oag	O antigen
OD600	optical density of 600 nm
OM	outer membrane
Omp	outer membrane protease
OSP	outer surface protein
PBS	phosphate buffered saline
PCP	polysaccharide co-polymerase
PCR	polymerase chain reaction
PFO	perfluoro-octanoic acid
PMN	polymorphonuclear cells

PTK	protein tyrosine kinase
R	resistance
Rha	rhamnose
Rif	rifampicin
R-LPS	rough LPS
rpm	revolutions per minute
RT	room temperature
RU(s)	repeat unit(s)
SAP	shrimp alkaline phosphatase
SDS	sodium dodecyl sulphate
SDS-PAGE	SDS polyacrylamide gel electrophoresis
S-LPS	smooth LPS
S-type	short type
SR-LPS	semi rough LPS
Sm	streptomycin
SU-LPS	semi unregulated LPS
TBS	tris buffered saline
Tet	tetracycline
Tp	trimethoprim
TM	transmembrane
Tris	tris (hydroxymethyl) aminomethane
TTBS	tween tris buffered saline
TTSS	Type III secretion system
U	units
UV	ultraviolet
v/v	volume per volume
V	volts

VL-type	very long type
VS-type	very short type
w/v	weight per volume
WM	whole membrane
Wzz_{SF}	*Shigella flexneri* Wzz
Wzz_{ST}	*Salmonella typhimurium* Wzz
X-Gal	5’-bromo-4-chloro-3-indolyl-β-D-galactopyranoside