Polarity and Secretion of *Shigella flexneri* IcsA: A Classical Autotransporter

MATTHEW THOMAS DOYLE,
B. SC. (BIOTECHNOLOGY)

THE UNIVERSITY of ADELAIDE

Submitted for the degree of Doctor of Philosophy

Department of Molecular and Cellular Biology

School of Biological Sciences

The University of Adelaide

Adelaide, South Australia, Australia

July 2015

Declaration

I certify that this Thesis contains no material which has been accepted for the award of any other degree or diploma in my name, 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 certify that no part of this work will, in the future, be used in a submission in my name for any other degree or diploma in any university or other tertiary institution without the prior approval of the University of Adelaide and where applicable, any partner institution responsible for the joint-award of this degree.

I give consent to this copy of my thesis when deposited in the University Library, being made available for loan and photocopying, subject to the provisions of the Copyright Act 1968.

I acknowledge that copyright of published works contained within this thesis resides with the copyright holders of those works.

I give permission for the digital version of my thesis to be made available on the web, via the University's digital research repository, the Library Search and also through web search engines, unless permission has been granted by the University to restrict access for a period of time.

Matthew Thomas Doyle

July 2015

Abstract

The classical autotransporter IcsA is an essential virulence factor for the enteropathogen *Shigella flexneri* as it provides adherence properties and allows intra- and intercellular spreading in the colonic mucosa. IcsA is an outer membrane surface protein that specifically hijacks host-cell actin recruiting and polymerising complexes allowing actin polymerisation as a form of actin-based motility (ABM). Importantly, since IcsA is localised specifically at one end of the bacterium (the pole), the resulting ABM is unidirectional which is a requirement for efficient *S. flexneri* dissemination. However, the molecular mechanisms that generate IcsA polarity remain poorly understood. Furthermore, IcsA is a member of the autotransporter family of secreted virulence factors (Type Va). Although many steps in the autotransporter pathway have been elucidated, it is still poorly understood how these diverse proteins are efficiently translocated to the bacterial cell surface. As such, this thesis investigates the two arms of IcsA biogenesis: (1) polar targeting and (2) autotransport.

Regarding IcsA polarity, it was identified here that the IcsA-specific outer membrane protease IcsP localises to the septa of dividing *S. flexneri* and to the opposing pole relative IcsA. The concentration of IcsP was higher at the septum than the pole showing a life cycle dependent distribution of IcsP. This provides the basis of a model where IcsP is important during division of *S. flexneri* for setting up (and the continued maintenance of) IcsA polarity by the proteolysis of misdirected IcsA. Further, multiple previous reports have suggested that the *S. flexneri* lipopolysaccharide O-antigen surface structure can influence IcsA polarity by augmenting membrane fluidity or by asymmetric masking of IcsA surface exposure. These notions were tested here resulting in data that clearly refutes these models and agues simply that IcsA exposure is masked symmetrically over the bacterial cell surface. Finally, IcsA itself contains polar targeting (PT) regions that direct it to the pole by an, as yet, unclear mechanism. Examination of these regions revealed that the central PT (cPT) is most important in polarity augmentation and contains critical polarity targeting function residues.

Regarding IcsA autotransport, a highly conserved but uncharacterised autotransporter motif was scrutinized for potential biogenesis functions and was designated in this work as the passenger-associated transport repeat (PATR). It was found that the PATR plays a critical role in the efficient secretion of IcsA to the cell surface. Strikingly, bioinformatics analyses revealed that the PATR delineates an important separate autotransporter sub-type with unique functions, composition, and architecture.

Acknowledgements

First and foremost, I would like to thank my supervisor Associate Professor Renato Morona. Thank you for all our long conversations concerning the biogenesis of this enigmatic protein called IcsA. Thank you for always being there to bounce around ideas. Thank you for your help in times of great need. Thank you for believing in me and allowing me to take control of the development of my projects which has allowed me to form my own identity as a bacteriologist. You have guided me in my transition from a wide-eyed graduate and into an independently thinking researcher. You have imparted within me a real love for science that I just cannot shake, even if I wanted to. For all of these things, I am forever thankful and appreciative.

I would like to acknowledge The University of Adelaide for providing me with scholarship support during my candidature, and for facilitating many positive experiences throughout both my undergraduate and postgraduate years.

I would also like to specially thank Dr Alistair Standish and Dr Elizabeth Tran for teaching me numerous laboratory skills, tips, ticks, and experiments. Thank you both for critically analysing my works and providing me your opinions. This has been invaluable to me. Over the years you have both become very important people to me. Specifically Ali, thank you for all your encouragement to push myself and present my work at conferences, to write papers, and for teaching me all the extra things that come with a career in science. It was always a pleasure to go for a run with you, or have a beer, or continuously swap sexual innuendos. And Lizzy, it was also a pleasure to co-author with you and I am so appreciative for all our talks and for listening to my problems. Thank you for always making me feel better when I was down.

I also thank all the other past and present members of the Morona Laboratory including Jon, Min, Mabel, Pratiti, Paul, Zuleeza, and Jilong. You have all contributed to the atmosphere and dynamic of the lab and provided an interesting workplace. I would also like to thank past members Dr Marcin Grabowicz and Dr Kerrie May for your contributions and preliminary investigations that have led to co-authored publications. I would also like to thank all those that made coffee time so enjoyable including Dr Connor Thomas, Dr Lindsay Dent, and Dr Antonio Focareta. Tony, thank you for always asking how I'm going with my research, and just in general. I have appreciated this interest. To Dr Chris Wong and Dr Jamie Botten, thank you for making undergraduate laboratory demonstrating both a fun and enriching experience.

I also thank all the past and present members of the Paton Laboratory including Brock, Richard, and Charlie, the McDevitt Laboratory including Miranda, Bart, Jackie, Victoria, and Stef, and also Ash from the Booker Laboratory and Dr Stephen Kidd; you have all supported me, lent an ear, or provided social relief at different stages during my PhD. In particular, I would like to doubly thank Miranda, Bart, and Chris for providing various research and career advice. Stef, I would also like to especially thank you for your friendship over my undergraduate and postgraduate years. You have provided me with great support and encouragement throughout this time, and have become my 'yard-stick' from which I measure my own achievements. I really believe that you will go very far in science.

To Lucy and Rob, thank you for providing me with a second home where I always feel much love and warmth. Walking through your garden I always found myself de-stressed and visiting you both always left me refreshed and ready for the next week at work. I thank you both for your continued love and I hope that I can make you proud.

I also thank Callan. You mean a lot to Bianca and I, and I consider you the brother that I never had. Believe it or not, your friendship has enriched me and made me a better person at home and at work.

To my little sister Rebecca. Thank you for your support and always asking me about my studies. Thank you for always encouraging me and being proud of my achievements.

To my beautiful partner Bianca. You have been critical in my success so far, and like in all aspects of my life, you have been the biggest supporter of me and my scientific passions. You have given me life-balance; without you, I think I would sleep in the lab and become a lonely disconnected person. You have also foregone so much to help me during my PhD, without which, I think I would be finishing at a much later stage. You were always there to listen to my deepest fears and worries, and then build me back up. For all of these things, I am eternally grateful. I love you.

And to my dear Mum, you know exactly what it has taken for me to get to this point. I would never have got here without your protection and love during my early life. You have made such significant sacrifices for me and I don't think that I could ever do enough to properly repay you. I dedicate this thesis to you.

'For Mum, who protected me'

Thesis Style and Layout

This Thesis is submitted in the style of a 'Thesis by Publication'. As such, the results chapters are replaced by four Research Articles (In chronological publishing order Chapters 3 to 6). In Chapter 1: 'Introduction', the research is framed and the Aims of my Doctoral studies are outlined. Chapter 2 outlines the base 'Materials and Methods' used throughout candidature, with nuances and changes specific to each article described in the methods subsection of the respective publication.

Author contributions for each publication are stated in preceding 'Statements of Authorship', and the objectives of the article with respect to my thesis are outlined in 'Purpose of Article' subheadings. Article chapters retain the section order layout style of the journal publisher. Subheadings starting with 'Article …' indicate subdivisions that were included in the publication. Article supplementary data is included in this thesis (and [S] after the figure or table number). Please note that formatting may be slightly different between articles (i.e. US versus Australian spelling). This is a result of complying with the specific journal requirements. Important data that was collected in relation to a publication (but was not included in the article due to its peripheral nature) are marked with 'Additional Results' subheadings (and [A] after the figure number).

Chapter 7: 'Conclusions' draws together the outcomes of the presented publications, argues how the data aligns with the current mechanisms and models, and describes how this research leads to intriguing questions for future investigations.

Publications

Peer Reviewed Research Articles

Tran ENH*, **Doyle MT***, Morona R (2013) LPS unmasking of *Shigella flexneri* reveals preferential localisation of tagged outer membrane protease IcsP to septa and new poles. *PLoS ONE* 8(7): e70508. (* equal authorship)

Doyle MT, Grabowicz M, May KL, Morona R (2015) Lipopolysaccharide surface structure does not influence IcsA polarity. *FEMS Microbiology Letters* 362 (8), fnv042

Doyle MT, Tran ENH, Morona R (2015) The passenger-associated transport repeat promotes virulence factor secretion efficiency and delineates a distinct autotransporter subtype. *Molecular Microbiology* 97(2): 315-329.

Doyle MT, Grabowicz M, Morona R (2015) A small conserved motif supports polarity augmentation of *Shigella flexneri* IcsA. *Microbiology-SGM* doi: 10.1099/mic.0.000165.

Abbreviations

a.a.	amino acid
ABM	actin-based motility
Ap^R	ampicillin resistance cassette
AT	autotransporter
BAM	β-Barrel Assembly Machinery
BSG	buffered saline gelatin
CM	cytoplasmic membrane
Cml^R	chloramphenicol resistance cassette
cPT	central polar targeting region
DMEM	Dulbecco's modified eagles medium
DNA	deoxyribonucleic acid
dNTP	deoxyribonucleotide triphosphates
D-PBS	Dulbecco's phosphate buffered saline
DSP	dithiobis(succinimidyl propionate)
DTRM	domain targeted random mutagenesis
ESS	extended signal sequence
F-actin	filamentous actin
FAE	follicle-associated epithelia
FCS	foetal calf serum
g	force of gravity
GA	Gibson assembly
GFP	green fluorescent protein
GL	grey level
GRR	glycine rich repeats
GST	glutathione S-transferase
HRP	horse radish peroxidase
HS	hot-spot
IF	immunofluorescence
IM	inner membrane
IPTG	isopropyl-β-D-thiogalactopyranoside
kb	kilobase
kDa	kilodalton
Km^R	kanamycin resistance cassette
LB	Luria Bertani broth / Lysogeny broth
LPA	large polar aggregate
LPS	lipopolysaccharide
MEM	modified eagles medium
MQ	MilliQ H_2O
mRNA	messenger RNA
Mw	molecular weight
NC	non-cleavable
nPT	N-terminal polar targeting region
ns	not significant
N-WASP	Neural Wiskott Aldrich Syndrome Protein
OA	oligo annealing
Oag	O-antigen
OD_{600}	optical absorbance at 600 nm
OM	outer membrane
OMP	outer membrane protein
PAGE	polyacrylamide gel electrophoresis
PAP2	type 2 phosphatidic acid phosphatase
PATR	passenger-associated transport repeat
PbH1	parallel β-helix type 1
PBS	phosphate buffered saline
PCR	polymerase chain reaction
PGN	peptidoglycan
PK	proteinase K
PL	pertactin-like
PLL	pectin lyase like
PMBN	polymixin B nonapeptide
PMN	polymorphonuclear leukocytes
PMSF	phenylmethanesulfonyl fluoride
POMP	polymorphic outer membrane protein
POTRA	polypeptide transport associated
PT	polar targeting region
QC	QuikChange mutagenesis
QD	quantum dot
RBS	ribosome binding site
RDA	randomly distributed aggregate
R-LPS	rough lipopolysaccharide
RNA	ribonucleic acid
SD	standard deviation
SDS	sodium dodecyl sulfate
SEM	standard error of the mean
ShET	*Shigella* enterotoxin
^{S}LPS	short type lipopolysaccharide
S-LPS	smooth lipopolysaccharide
SOE	splicing by overlap extension
SPATE	serine protease AT of the enterobacteriae
SRL	*Shigella* resistance locus
SRP	signal recognition particle
SS	signal sequence
T3SS	Type III secretion system
TBS	tris buffered saline
TCA	trichloroacetic acid
Tc^R	tetracycline resistance cassette
TS	tryptic soy
TSP	tail-spike protein
TTBS	tris buffered saline plus tween
v/v	volume per volume
VC	vacuolating cytotoxin
^{VL}LPS	verly long type lipopolysaccharide
VP	virulence plasmid
w/v	weight per volume
WIP	WASP inhibiting protein
WM	whole membrane
WT	wild-type
X-gal	5-bromo-5-chloro-3-indolyl-β-D-galactoside
β-ME	β-mercaptoethanol

Contents

Chapter One

INTRODUCTION

Chapter 1: Introduction

Proteins perform an innumerable array of functions in living cells. Yet, the specific functions of proteins are commonly dependent on their correct spatial organisation within and around the cell. This is true for most bacterial virulence proteins, and it is becoming increasingly clear that bacterial pathogens exert considerable efforts in order to precisely place their virulence weapons. Indeed, virulence proteins are often secreted to the bacterial cell surface such that they can interact with the host-pathogen interface, assist colonisation, and augment pathogenic fitness. Secretion pathways must overcome physical barriers such as the cell membrane during nascent protein transit. This is even more of a challenge for Gram negative bacteria that have two cellular compartments, dual membrane cell walls, and multiple possible membrane topologies. Nascent protein navigation of these cellular compartments must be an efficient, timely, and structured process to prevent falling into low energy traps and degradative pathways. But while bacteria need to correctly distribute proteins between cellular compartments, they also have to distribute them correctly relative to the macrostructure of the cell. For appropriate functioning, rod shaped bacterial pathogens typically require the asymmetric placement of certain virulence factors at the cell tips or 'poles', as opposed to the lateral edge of the cell. Classic examples of this include the flagellum motility apparatus (Kazmierczak & Hendrixson, 2013) and chemotactic sensors(Collins *et al*, 2014).

Therefore, it is of critical importance to understand how bacterial pathogens precisely place and organise their virulence factors such that these mechanisms can be targeted with novel anti-virulence therapeutics. Accordingly, the theme of the research herein was to explore spatial distribution mechanisms of bacterial virulence factors in the contexts of both targeted placement around the macrostructure of the cell, and correct export to the bacterial cell surface. This Thesis examines the essential virulence factor IcsA (VirG) from *Shigella* species. IcsA is both: (i) polarly localised and (ii) secreted to the bacterial cell surface – although the localisation mechanisms for these spatial organisations are yet to be completely understood. This chapter reviews the role of IcsA in *Shigella* pathogenesis, IcsA polarity maintenance factors, IcsA autotransport to the bacterial surface, and establishes the Research Questions and Aims.

1.1. *Shigella flexneri*: classification, disease burden, and challenges

Shigella species are the causative agents of the disease Shigellosis which can manifest itself subtly as watery diarrhoea, or dramatically as painful bloody dysentery that includes symptoms such as cramps, vomiting, and seizures (Niyogi, 2005, von Seidlein *et al*, 2006). These human pathogens proficiently and specifically invade the mucosal tissues of the colon and rectum causing the formation of lesions (Speelman *et al*, 1984). The prevalence of this disease is highest in developing countries with an estimated 80 million cases per year resulting in 700,000 deaths (WHO, 2005). The incidences of *Shigella* related diseases are highest in the vulnerable age groups of the very young and the very old (Kotloff *et al*, 1999, von Seidlein *et al*, 2006, Bardhan *et al*, 2010). Although the mortality rate has dropped in industrialised countries (Wang *et al*, 2006), deaths frequently occur in developing countries where illness is coupled with immunocompromisation in instances such as HIV co-infection (Kownhar *et al*, 2007, Musiime *et al*, 2009, Simpore *et al*, 2009, Keddy *et al*, 2012). Also emerging is its classification as an atypical sexually transmitted disease with a significant disease burden in men who have sex with men (Baker *et al*, 2015).

The *Shigella* genus comprises four recognized Gram negative bacilli species named *S. dysenteriae*, *S. flexneri*, *S. boydii*, and *S. sonnei* (Niyogi, 2005). These can be further subdivided into serotypes and sub-serotypes according to their lipopolysaccharide (LPS) O-antigen structure (Levine *et al*, 2007). However, it has been shown that the classical *Shigella* genus is false. *Shigellae* actually belong to the *Escherichia coli* species in a single pathovar with enteroinvasive *E. coli* (Lan *et al*, 2004). A clear picture has now emerged revealing that the *Shigella* serogroups have arisen multiple separate times from ancient *E. coli* and have converged phenotypically (Jin *et al*, 2002, Yang *et al*, 2007). The attainment of a 220 kb virulence plasmid (termed pINV), the subsequent loss of its transferability, and loss of chromosomal anti-virulence genes, were likely the key steps in *Shigella* species emergence (Hale, 1991, Sansonetti, 2001, Yang *et al*, 2007, Schroeder & Hilbi, 2008). The virulence plasmid encodes multiple virulence factors required for establishing colonisation and disease, without which it is avirulent (Sansonetti, 2001, Schroeder & Hilbi, 2008).

S. flexneri (specifically serotype 2a) is the most commonly isolated *Shigellae* in developing countries (Kotloff *et al*, 1999, von Seidlein *et al*, 2006, Levine *et al*, 2007, Tiruneh, 2009). It has been reported that *S. flexneri* cause some of the more severe forms of the disease (von Seidlein *et al*, 2006) and is a focus of host-pathogen study around the world. Troublingly, there is an alarming rise in confirmed cases of multiply resistant *S. flexneri* (Kownhar *et al*,

2007, Tiruneh, 2009, Taneja *et al*, 2012). It has recently been identified that azithromycin-resistant *S. flexneri* 3a strains have efficiently dispersed across the continents of the world (Baker *et al*, 2015). Further, in *S. flexneri* 2a, an antibiotic resistance region has been identified (termed the *Shigella* Resistance Locus; SRL) that encodes resistances to streptomycin, ampicillin, chloramphenicol, and tetracycline (Luck *et al*, 2001, Turner *et al*, 2003). The SRL is set within a pathogenicity island, itself present in all *Shigellae*, and thus has the potential for horizontal transfer (Turner *et al*, 2003).

In parallel with the complications of treating antibiotic resistant strains of *S. flexneri*, there are also abundant challenges in the development of an effective vaccine. As yet, there is no vaccine in use against *S. flexneri* (Levine *et al*, 2007). To be effective a vaccine must be cross-protective for major *S. flexneri* serotypes (2a, 3a, and 6) and *S. sonnei* since it is a common replacement strain (Levine *et al*, 2007, Livio *et al*, 2014). Taken together, these challenges emphasize the importance of further work to elucidate the cellular and molecular pathogenesis of *S. flexneri*. Knowledge of the interactions between *Shigella* virulence factors and host cell molecules, and the biogenesis pathways of these virulence factors, is vital for the development of new strategies to deal with Shigellosis.

1.2. Invasive pathogenesis of *S. flexneri*

Shigella transmission occurs via the faecal-oral route. Accordingly, poor hygiene, sanitation, or handling of food can attribute to local spread of Shigellosis in the community (Niyogi, 2005). It has also been suggested that the common fly can act as a vector for *Shigella* spread via landing on infected human faeces, acquiring the organism, and then depositing it where it lands (Levine & Levine, 1991). In any case, once ingested *Shigellae* survive the passage to the colon as evidenced by the extremely low infective dose of less than 100 organisms of *S. flexneri* (DuPont *et al*, 1989). Indeed, *Shigellae* can survive in low pH for several hours allowing it to pass the acidic challenge of the stomach and reach the small intestine (Gorden & Small, 1993). In the small intestine *S. flexneri* releases ShET1/2 (*Shigella* enterotoxin) causing watery diarrhoea (Fasano *et al*, 1995, Fasano *et al*, 1997). *S. flexneri* then enters the colon in which the invasive stage of pathogenesis occurs.

1.2.1. Circumventing the epithelial barrier and invasion

Primary *S. flexneri* colonisation is targeted to follicle-associated epithelia (FAE) and Peyer's patches (Jensen *et al*, 1998). To establish infection in these sites *S. flexneri* must cross the epithelial barrier. However, it is unable to penetrate the apical side of epithelia directly (Mounier *et al*, 1992). Consequently, *S. flexneri* obtains access to the basolateral side of the epithelium by hijacking the luminal antigen sampling by membranous epithelial cells (M cells) which are present in FAE (Wassef *et al*, 1989, Jensen *et al*, 1998, Sansonetti & Phalipon, 1999) (Figure 1.1A). *S. flexneri* allows itself to be taken up by M cells which transcytose the bacterium across the barrier (Wassef *et al*, 1989) (Figure 1.1B). This results in the delivery of *S. flexneri* to a basolateral pocket containing resident macrophages which subsequently endocytose the bacteria.

In the macrophage endocytic vesicle, the bacterium quickly escapes before the endosome matures to a endolysosome capable of killing the bacterium. The virulence plasmid promotes this escape as it contains the *ipa-mxi-spa* locus encoding a Type III secretion system (T3SS) which is analogous to a molecular needle (Blocker *et al*, 1999, Moraes *et al*, 2008). The needle tip contains a complex of proteins IpaB, IpaC, and IpaD that is able to destabilise the vesicle lipid membrane in order for the *Shigellae* to break out into the host cytoplasm (High *et al*, 1992, Barzu *et al*, 1997, Blocker *et al*, 1999, Schuch *et al*, 1999) (Figure 1.1C).

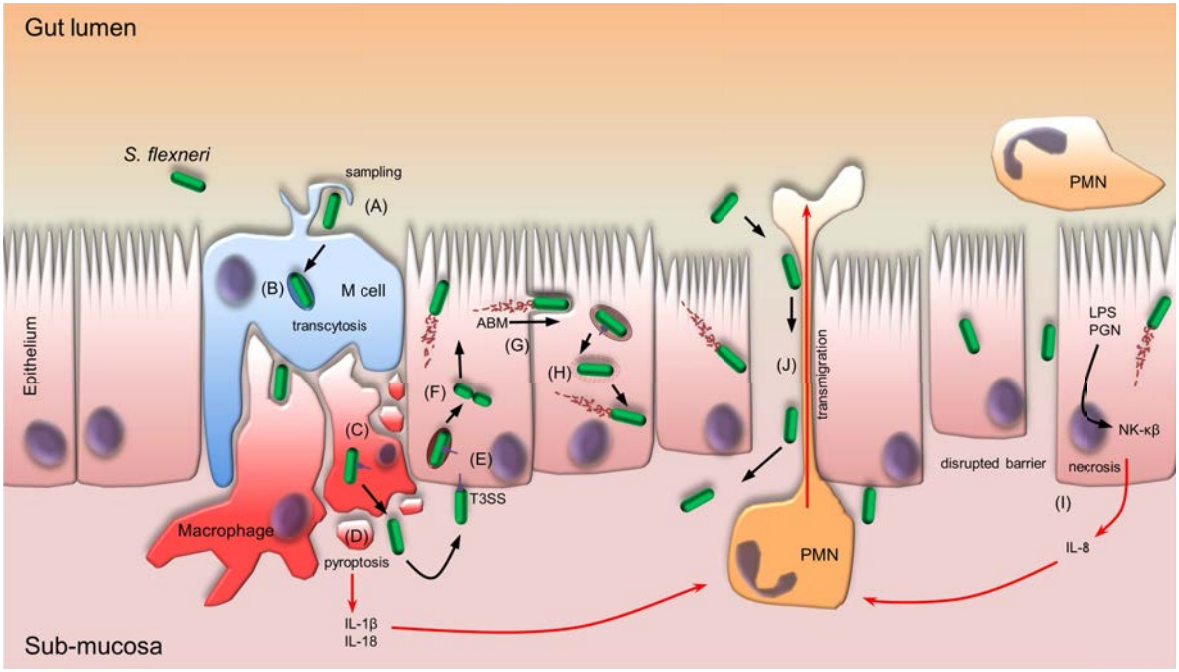


Figure 1.1: Summary of *S. flexneri* invasion and colonisation of the gut mucosa.

Shigella that enter the colon subvert the epithelial barrier via M cell uptake (**A**) and subsequent transcytosis (**B**). (**C**) Bacteria are immediately endocytosed by macrophages but are able to avoid killing by endosome escape. Induced macrophage pyroptosis (**D**) deposits *S. flexneri* into the sub-mucosa where it can trigger its uptake into epithelial cells (**E**). Escape from the endocytic vesicle (**F**) results in free *S. flexneri* in the cytosol were it can replicate. (**G**) IcsA induces actin-based motility (ABM) allowing intra- and intercellular spread between epithelial cells. Uptake into adjacent epithelial cells results in a double membrane vesicle bound bacterium, but this is destabilized (**H**) allowing escape of *S. flexneri* and enabling it to repeat **F** and **G**. Slow necrosis of infected epithelial cells produces inflammatory cytokines (**I**), and coupled with inflammatory cytokines produced from **D**, results in influx of polymorphonuclear leukocytes (PMN) which destabilize the mucosal barrier further causing *Shigella* entry via a paracellular route (**J**). LPS = lipopolysaccharide, PGN = peptidoglycan.

Cytoplasmic *S. flexneri* then induces rapid macrophage cell death (Zychlinsky *et al*, 1992). Induction of macrophage death is predominantly mediated by IpaB, although LPS and the lipid component of LPS, Lipid A, also contribute to this process (Zychlinsky *et al*, 1994, Suzuki *et al*, 2005, Senerovic *et al*, 2012). IpaB is able to activate IL-1β-converting enzyme (ICE) thus initiating cell death and inflammatory pathways (Zychlinsky *et al*, 1992, Chen *et al*, 1996, Hilbi *et al*, 1998). This releases both the bacterium and large amounts of inflammatory IL-1β and IL-18 into the sub-mucosa (Sansonetti *et al*, 2000, Schroeder & Hilbi, 2008) (Figure 1.1D). *S. flexneri* is now exposed to the basolateral surface of epithelial cells and it is here that its invasive capacity is realised.

S. flexneri loosely adheres to proteins within cholesterol/sphingolipid rafts of the epithelial cell membrane via its T3SS tip complex (Lafont *et al*, 2002). The T3SS tip is able to bind host surface CD44 as well as $\alpha_5\beta_1$ integrin (Watarai *et al*, 1996, Lafont *et al*, 2002). The T3SS secretes a range of effectors (such as IpaA, IpaC, IpgD, and VirA) into the host cell

cytoplasm which subsequently triggers cytoskeletal rearrangements required for *S. flexneri* uptake (Schroeder & Hilbi, 2008). IpgD disconnects the cell membrane with the underlying cytoskeleton allowing IpaC to induce actin polymerisation which forms filopodial-like membrane structures (Tran Van Nhieu *et al*, 1999, Niebuhr *et al*, 2002). At a later stage, IpaA causes actin depolymerisation which may allow closure of the filopod-like extensions over the bacterium and subsequent internalisation into an endocytic vesicle (Demali *et al*, 2006, Schroeder & Hilbi, 2008) (Figure 1.1E). The T3SS again secretes a range of effectors to allow *S. flexneri* to escape the endocytic vesicle and be released into the cell cytoplasm (Ashida *et al*, 2015) (Figure 1.1F).

1.2.2. IcsA initiates actin-based motility and lateral spread

In the epithelial cell cytosol, *S. flexneri* proliferates in its key pathological niche. At this point, it again hijacks host cell machinery to initiate an intriguing form of motility in order to spread laterally and cause severe destruction of the epithelium.

IcsA (VirG), a polar-localised, surface-exposed, outer membrane autotransporter protein, is sufficient to direct the movement of intracellular *S. flexneri* via a process called Actin-Based Motility (ABM) (Goldberg & Theriot, 1995). IcsA attracts host cell components to one pole of the bacillus to form an actin-nucleating complex (Goldberg, 2001) (see further subsection 1.3.3 and Figure 1.4). At the cell pole, this complex catalyses the polymerisation of globular actin subunits into actin filaments (F-actin) (Egile *et al*, 1999). The continued construction of F-actin behind the bacterium provides propulsive force and pushes it throughout the host cytoplasm. The distinctive F-actin trail left in the wake of motile *Shigella* species is widely known as the "comet-tail".

Motile *S. flexneri* now move around the cell cytosol, and unsurprisingly, some encounter the host plasma membrane. IcsA-induced ABM has sufficient force of propulsion as to allow *S. flexneri*-containing membrane protrusions to form which commonly occurs at tri-cellular junctions (Prevost *et al*, 1992, Fukumatsu *et al*, 2012). These protrude into the adjacent epithelial cells and are subsequently endocytosed by a clathrin-mediated pathway forming a double membrane vacuole containing the bacterium (Prevost *et al*, 1992, Monack & Theriot, 2001, Fukumatsu *et al*, 2012) (Figure 1.1G). *S. flexneri* can again escape from this compartment using T3SS effectors, such as IpaBCD, IcsB, and VirA which protect the bacterium from the autophagic system and allows release into a new host cell in which it can proliferate (Page *et al*,

1999, Campbell-Valois *et al*, 2015) (Figure 1.1H). In this way, *S. flexneri* is able to disseminate laterally throughout the epithelial barrier, spreading the focus of infection outwards.

1.2.3. Consequences of efficient lateral spread

S. flexneri components, such as LPS and peptidoglycan subunits, are detected by surveillance mechanisms within the cytosol of host epithelial cells (Philpott *et al*, 2000, Chamaillard *et al*, 2003). It has been shown that host NOD1 protein can detect bacterial peptidoglycan which subsequently causes an activation pathway of transcription factor NF_KB (Philpott *et al*, 2000, Girardin *et al*, 2003). This causes epithelial cells to up-regulate IL-8 production which is a major attractor of polymorphonuclear leukocytes (PMN) (Sansonetti *et al*, 1999, Philpott *et al*, 2000). There is also an array of T3SS effectors secreted such as Ipa's and Osp's that delay epithelial apoptosis, but promotes slow necrosis (Ashida *et al*, 2015) (Figure 1.1I). The IL-8 production and epithelial necrosis, coupled with the mass IL-1β and IL-18 release from *S. flexneri*-induced macrophage pyroptosis, causes an enormous influx of PMN to the site of infection (Sansonetti *et al*, 1995, Sansonetti *et al*, 1999, Schroeder & Hilbi, 2008). This influx of PMN is responsible for the majority of mucosal tissue destruction seen in Shigellosis. Indeed, the disruption of the epithelial barrier by PMN opens up an additional pathway for luminal *S. flexneri* to enter the sub-mucosa (Perdomo *et al*, 1994, Sansonetti *et al*, 1999) (Figure 1.1J).

Thus, a picture emerges where lateral spread of *Shigellae* throughout the epithelium increases the number of IL-8 secreting epithelial cells which, in turn, increases the severity of PMN-induced mucosal destruction and necrotic defacement. Resulting lesions occur throughout the colon, from the ascending colon to the rectum (Speelman *et al*, 1984). This exemplifies the importance of ABM-based, IcsA-dependant, *S. flexneri* lateral spreading in causing the major symptoms of bacillary dysentery.

1.3. Intra- / inter-cellular spread protein A (IcsA)

In the mid 1980's, a SalI-EcoRI fragment within the *S. flexneri* virulence plasmid was linked to the bacteria's ability to spread between epithelial cells. This was done by making transposon insertions in this region and observing that mutant strains were deficient in lateral spread between Rhesus monkey epithelial cells (Makino *et al*, 1986). This region, initially termed *virG* by Makino *et al* (1986), was later sequenced and found to contain a 3.3 kbp open reading frame that encodes a surface exposed 1102 amino acid VirG protein (Lett *et al*, 1989). At the same time, another group identified a region termed *intracellular spread A* (*icsA*) via Tn*phoA* mutagenesis and established that it is one and the same as the *virG* locus (Bernardini *et al*, 1989). They found that the encoded IcsA protein was an outer membrane protein of approximately 115 kDa and was required for intracellular motility and protrusion formation by means of F-actin accretion at one pole of *S. flexneri* (Bernardini *et al*, 1989). ABM and IcsA has been the focus of extensive study via many modes of experimentation. The Sereny test which is a test for the ability of bacteria to cause keratoconjuntavitis in mouse or guinea pig eyes, has been useful since virulence is directly associated with ABM (Murayama *et al*, 1986). Furthermore, other *in vitro* methods such as: tissue culture plaque formation assays (Oaks *et al*, 1985), comet tail formation assays in *Xenopus* egg cytoplasmic extracts (Kocks *et al*, 1995, Magdalena & Goldberg, 2002), and immunofluorescence microscopy of infected tissue culture cells coupled with F-actin staining (Kocks *et al*, 1995, Egile *et al*, 1999, May & Morona, 2008), have all been very valuable for elucidating the mechanisms of ABM and lateral spreading.

Indeed, immunofluorescence microscopy and comet tail formation assays have been used to characterise the heterologous expression of IcsA in *E. coli*. Two groups undertook these experiments concomitantly and found that IcsA expressed in *E. coli* can form actin based comet tails and that these tails exert force minimally sufficient for motility (Goldberg & Theriot, 1995, Kocks *et al*, 1995). Subsequently, it was found that IcsA expressing *E. coli* are able to conduct rudimentary ABM in infected HeLa cells, are able to form protrusions, and that these protrusions can be pinched off into adjacent cells (Monack & Theriot, 2001). However, this occurred with very low efficiency since IcsA surface localisation was both polar and lateral. But when IcsA unipolarity was increased by co-expressing either the IcsP outer membrane protease (specific for IcsA) or LPS O-antigen synthesis genes, the subsequent tail formation efficiency, unidirectional movement, and protrusion formation all increased (Monack & Theriot, 2001). Therefore, the asymmetric maintenance and refinement of IcsA at one bacterial pole is a key requirement for efficient *S. flexneri* spread, and

consequently, virulence. IcsA polarity, and other factors involved in polarity (IcsP and LPS), are discussed in section 1.4.

1.3.1. IcsA quantification, expression, and regulation

The quantity of IcsA molecules on the bacterial surface affects the ability of *S. flexneri* to conduct ABM. Indeed, the threshold for actin tail formation in *Xenopus* egg cytoplasmic extracts is approximately 4000 IcsA molecules (Magdalena & Goldberg, 2002). Furthermore, when compared to growth *in vitro*, intracellular *S. flexneri* increase their IcsA surface expression 3.2 fold enabling a portion of the infecting bacterial population to be comet tail formation competent (Magdalena & Goldberg, 2002). The expression of IcsA is tightly controlled by a multi-tiered hierarchy of regulatory mechanisms that respond to physical environmental stimuli. *S. flexneri,* like many other bacteria, has a histone-like protein called H-NS (VirR) (Hromockyj *et al*, 1992) that is shown to bind widely to promoter regions of virulence factors and regulatory genes in a manner which represses transcription (Prosseda *et al*, 2002, Landick *et al*, 2015). The promoters of *virB, virF, icsA, rnaG* and *icsP* are all repressed by H-NS at low temperatures (Prosseda *et al*, 1998, Wing *et al*, 2004, Tran *et al*, 2011) and are the relevant genes for IcsA expression regulation. VirB and VirF are virulence plasmid encoded transcription factors, IcsP is an outer membrane protease, and RnaG is a small RNA antisense to *icsA* mRNA.

Upon *S. flexneri* entering the host, it encounters environmental signals such as a temperature increase to 37 °C and pH changes. These two factors are sensed by H-NS which cause it to relax its repression of *virF* and *virB* (Prosseda *et al*, 1998, Tran *et al*, 2011). VirF is a major activator transcription factor which contributes to derepression of *virB* and *icsA* by binding their promoter regions (Sakai *et al*, 1988, Adler *et al*, 1989, Tran *et al*, 2011). Therefore, VirF is a direct activator of *icsA* transcription (Tran *et al*, 2011). During *icsA* transcriptional activation by VirF, VirB is also being expressed and is an activator of a range of virulence plasmid genes including *ipa* invasion genes (Adler *et al*, 1989) allowing invasion and pathogenesis. The level of *icsA* transcription control is further modulated by RnaG since its promoter overlaps that of *icsA* (Giangrossi *et al*, 2010). RnaG negatively regulates *icsA* transcription via two ways; (i) the promoter of *rnaG* is more 'aggressive' and its expression interferes with the expression of *icsA* from its promoter, and (ii) antisense RnaG binds to *icsA* mRNA during transcription promoting early termination (Giangrossi *et al*, 2010). However, during permissive temperatures VirF which is an activator of *icsA* is also a repressor of *rnaG* (Tran *et al*, 2011).

It should be noted here that post-translational IcsA spatial regulation is also initiated by this regulatory network. As mentioned, VirB activates a range of virulence factor genes and one of these is via the derepression of *icsP* transcription (Wing *et al*, 2004). This results in increased production of the outer membrane protease IcsP which can subsequently refine the distribution of IcsA on the cell surface by IcsA-specific proteolysis (Wing *et al*, 2004) (see sub-section 1.4.3).

1.3.2. Outline of IcsA primary structure and important sub-regions

IcsA has multiple primary structural regions that allow its proper functioning, localisation, and secretion to the cell surface (Figure 1.2). These regions will be outlined from N- to C-terminus in preview of a more complete examination in later sections.

IcsA contains an N-terminal signal sequence ($IcsA_{1-52}$) for inner membrane passage via the Sec pathway (Goldberg *et al*, 1993, Suzuki *et al*, 1995). This signal sequence is longer than usual but is a feature found exclusively in autotransporters (Leyton *et al*, 2012, Grijpstra *et al*, 2013). Since IcsA is a Type V secreted autotransporter (see section 1.5), the region $IcsA_{53-740}$ comprises a passenger domain that is transported to the bacterial cell surface and contains the information necessary for its virulence functions (Goldberg & Theriot, 1995, Suzuki *et al*, 1995). The signal sequence is cleaved from the passenger during inner membrane passage. IcsA has a β-barrel putatively spanning $IcsA_{813-1102}$ that anchors the protein into the bacterial outer membrane and plays a key role in passenger translocation across the outer membrane (Fukuda *et al*, 1995, Suzuki *et al*, 1995). The passenger and the barrel are connected by a linker region that putatively spans $IcsA_{741-812}$. The passenger can be specifically cleaved (with low efficiency) between R^{758} and R^{759} by IcsP (Goldberg *et al*, 1993, Fukuda *et al*, 1995, d'Hauteville *et al*, 1996, Egile *et al*, 1997). A protease stable portion of the passenger called the stable core has been structurally solved and exhibits a right-handed β-helical fold ($IcsA_{591-740}$; see Figure 1.10A sub-section 1.5.2) (Kuhnel & Diezmann, 2011). The region $IcsA_{320-433}$ is known to be targeted by the host cell autophagic response system protein Atg5, but this region is masked by the virulence plasmid encoded T3SS effector IcsB (see sub-section 1.3.4) (Ogawa *et al*, 2005, Kayath *et al*, 2010). An N-terminal polar targeting (nPT) region and a central polar targeting (cPT) region ($IcsA_{1-104}$ and $IcsA_{506-620}$ respectively) have been attributed roles in IcsA-dependent polar localisation (see sub-section 1.4.1) (Charles *et al*, 2001). The passenger shares sequence homology with pectin lyases and P22 phage tail-spike protein ($IcsA_{154-737}$). There is also a conserved Pertactin-like region at the C-terminus of the passenger

($IcsA_{650-736}$). A highly conserved 32-aa repeat motif of unknown function (autotransporter repeat) is found over $IcsA_{526-557}$ (see sub-section 1.5.3).

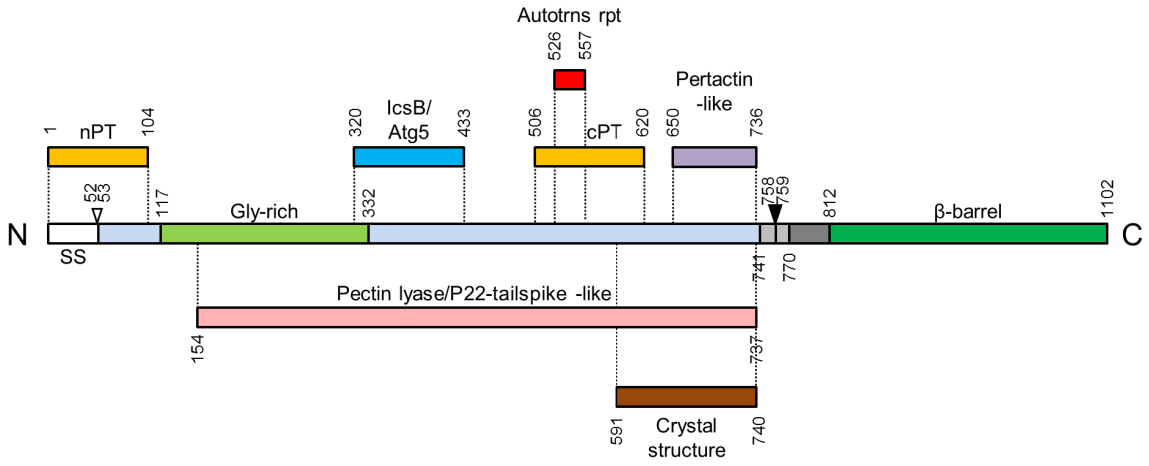


Figure 1.2: Structural, functional and conserved regions of IcsA.

A scaled schematic of IcsA primary structural features. As an autotransporter protein, IcsA contains an N-terminal signal sequence (SS; white) comprising $IcsA_{1-52}$, a passenger domain ($IcsA_{53-740}$; pale blue) that contains N-WASP binding regions including the glycine rich region $IcsA_{117-332}$ (light green), and a β-barrel domain ($IcsA_{813-1102}$; dark green) for outer membrane insertion and passenger translocation. The passenger and β-barrel are separated by a linker region ($IcsA_{741-812}$) that contains a putative unstructured region ($IcsA_{741-770}$, light grey) and a putative α-helical region ($IcsA_{771-812}$, dark grey). The SS is cleaved between $A^{52}T^{53}$ (white arrow) during inner membrane transit. There is also a double arginine site ($R^{758}R^{759}$; black arrow) that can be hydrolysed by outer membrane protein IcsP. A portion called the stable-core ($IcsA_{591-740}$; brown) has been solved by X-ray crystallography. $IcsA_{320-433}$ (blue) is targeted by host cell autophagy protein Atg5, but masked by T3SS effector IcsB. Polar targeting (PT; yellow) regions $IcsA_{1-104}$ (N-terminal PT; nPT) and $IcsA_{506-620}$ (Central PT; cPT) promote IcsA-dependent polar targeting. $IcsA_{154-737}$ (pink) has sequence conservation with pectin lyases (IPR011050) and P22 phage tail-spike proteins (IPR012332). $IcsA_{650-736}$ (purple) has sequence conservation with the C-terminal portion of the autotransporter Pertactin (IPR004899). The region $IcsA_{526-557}$ (red) is a conserved repeat motif that remains completely uncharacterized (IPR013425). Codes are InterPro identifiers (ebi.ac.uk/interpro).

1.3.3. IcsA passenger functions: N-WASP binding and adhesion

As mentioned previously, IcsA is sufficient for ABM (Goldberg & Theriot, 1995, Kocks *et al*, 1995, Monack & Theriot, 2001). It is the surface exposed portion of IcsA (the passenger domain) that drives the molecular processes that result in actin comet tail formation. IcsA does this by recruiting host cell actin polymerisation machinery to the *S. flexneri* polar surface initiating the nucleation and formation of F-actin. Many lines of evidence show that the IcsA passenger domain primarily attracts a host protein called Neural Wiskott-Aldrich Syndrome protein (N-WASP) from the cytosol to form an F-actin polymerising complex with Arp2/3.

WASP family proteins are involved in the maintenance of host cell cytoskeleton and actin polymerisation (Kim *et al*, 2000, Anton *et al*, 2007). They are held in a stable state by WASP interacting protein (WIP) and auto-inhibited by intramolecular interactions until Cdc42 GTPase binding mediates WASP activation (Kim *et al*, 2000, Anton *et al*, 2007). In relation to *S. flexneri* ABM, N-WASP has been repeatedly observed via immunofluorescence microscopy to localise to the bacterial pole end of the actin comet tail (Suzuki *et al*, 1998, Egile *et al*, 1999, Suzuki *et al*, 2002) (see Figure 1.3). In addition, N-WASP depleted *Xenopus* egg extracts are unable to support *S. flexneri* F-actin tail formation, and back-adding of N-WASP restores ABM (Suzuki *et al*, 1998). The dependence on N-WASP for ABM has been further shown in N-WASP deficient fibroblast-like cells (Lommel *et al*, 2001, Snapper *et al*, 2001). Importantly, other WASP family proteins such as WASP and WAVE1 are unable to interact with IcsA (Suzuki *et al*, 2002). Indeed, macrophages which express WASP preferably to N-WASP are unable to support *S. flexneri* ABM, and macrophages ectopically expressing N-WASP allow ABM (Suzuki *et al*, 2002). Hence, the intracellular motility of *Shigella* is restricted to cells that express N-WASP.

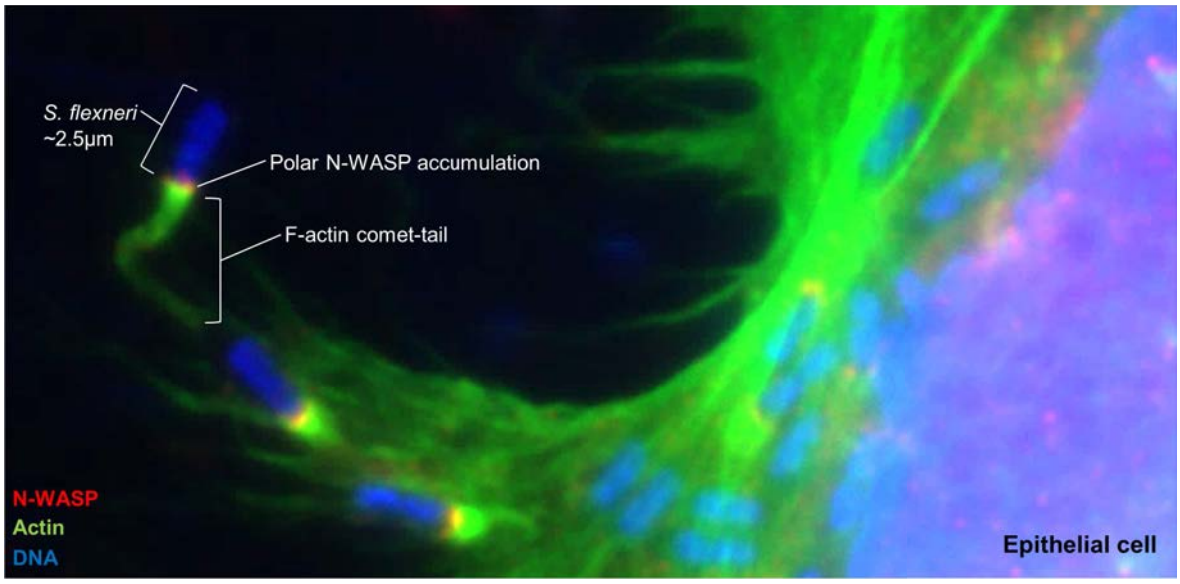


Figure 1.3: *S. flexneri* induces F-actin comet-tail formation in epithelial cells.

Example immunofluorescence micrograph of a HeLa cell infected with *S. flexneri* (blue). An F-actin comet-tail (green) is clearly visible and forms a *Shigella*-containing protrusion that extends outwards from the cell. The tail is unidirectionally formed at the pole that has accumulated N-WASP (red) in an IcsA dependent manner. Nucleus and bacterial nucleoids stained with DAPI, F-actin stained with phalloidin-488, and N-WASP stained with rabbit anti-N-WASP antibody and donkey anti-rabbit-594 antibody.

Although, the exact IcsA-to-N-WASP interaction events have not yet been completely elucidated, the use of *in vitro* pull-down experiments using several fusions and truncations of N-WASP and IcsA have attributed significantly to understanding which regions

are important for binding. By using various lengths of the IcsA passenger domain fused to GST (glutathione S-transferase), the minimum region required for N-WASP binding was found to be $IcsA_{106-433}$ (Suzuki *et al*, 1998). This region containing glycine-rich repeats (GRR) was thought to be the lone N-WASP binding region until a linker insertion mutagenesis study revealed that other parts of the passenger domain may be required for N-WASP binding as well (May & Morona, 2008). Subsequent studies have indicated that N-WASP binding function may extend to C-terminal zones of IcsA passenger domain (Teh & Morona, 2013).

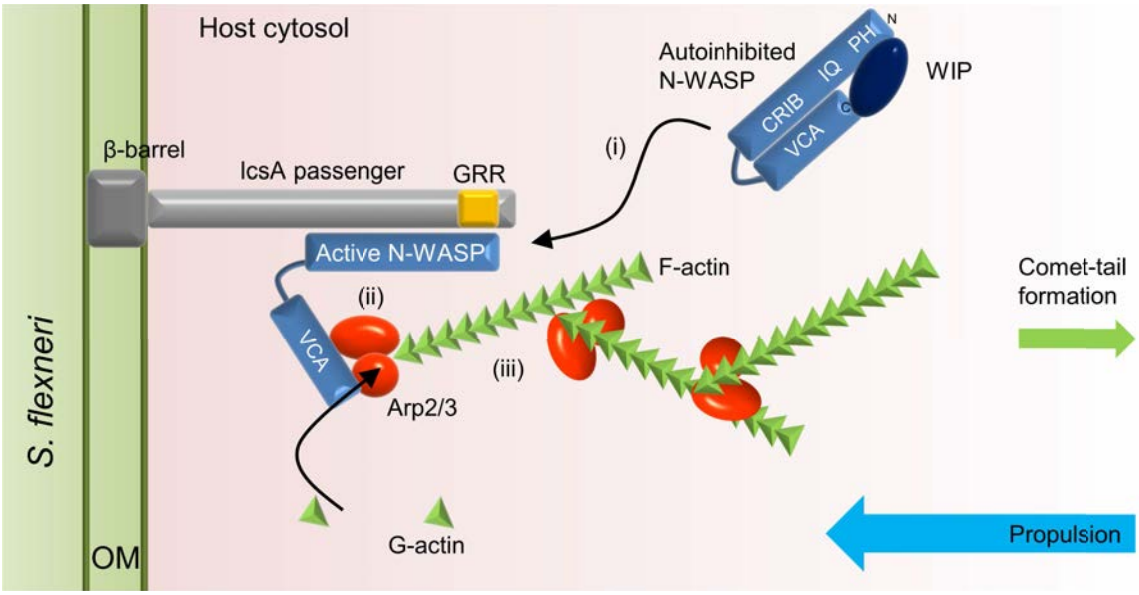


Figure 1.4: IcsA recruits host actin-polymerising proteins for actin based motility.

In the epithelial cell cytosol (pink background) *S. flexneri* (green) uses the polar localized surface protein IcsA to initiate ABM. Host protein N-WASP is normally auto-inhibited and inhibited with the help of WASP-interacting protein (WIP). IcsA recruits N-WASP (**i**) and activates it to free the VCA domain (**ii**). The VCA domain then recruits host cell Arp2/3 proteins forming a complex that polymerizes globular actin (G-actin) into filamentous actin (F-actin) (**iii**). Growth of F-actin comet tail provides a unidirectional force that propels *S. flexneri* intra- and intercellularly.

The N-terminal region of N-WASP is the prime target of IcsA. To determine the N-WASP domains that interact with IcsA, researchers have utilised fusions of fluorescent proteins to a series of N-WASP truncations and deletions. GFP fusions to N-WASP(ΔVCA) and the CRIB domain of N-WASP are still able to bind to the polar surface of *S. flexneri* but result in an inhibitory effect on ABM (Moreau *et al*, 2000). Subsequently, it was found that $N\text{-}WASP_{1-154}$ (which contains the PH and IQ domains) can also be recruited to IcsA independently of CRIB (Lommel *et al*, 2001) showing that IcsA may have redundant mechanisms of attracting N-WASP. This paper also showed that N-WASP fusion proteins must include the VCA domain to allow actin tail formation (Lommel *et al*, 2001). This finding links well with the fact that the VCA domain is responsible for binding host Arp2/3, a

complex that conducts actin polymerisation during normal cytoskeletal regulation (Miki & Takenawa, 2003), and that Arp2/3 has been shown to co-localise to IcsA-bound N-WASP (Egile *et al*, 1999).Thus, the following outlines the current model of IcsA-mediated ABM: (Figure 1.4; i) IcsA passenger domain, acting similarly to Cdc42, recruits N-WASP from the host cytosol by binding its N-terminal domains, (Figure 1.4; ii) this activates N-WASP from its auto-inhibited state and frees its VCA domain to recruit the Arp2/3 complex, (Figure 1.4; iii) the Arp2/3 complex then attracts globular actin subunits which initiates the barbed end growth of F-actin (Egile *et al*, 1999, Sansonetti, 2001, Schroeder & Hilbi, 2008). IcsA may act as a functional oligomer during this process (May *et al*, 2012, Teh & Morona, 2013).

It has also recently been found that IcsA can act as an adhesin. When *S. flexneri* are grown in the presence of bile salts such as deoxycholate, IcsA exhibits a conformational change that switches its adhesive activity on and subsequently enhances invasion of epithelial cells (Brotcke Zumsteg *et al*, 2014). Although, the host receptor is not yet identified, this is an extremely exciting finding considering the lack of knowledge of *Shigella* adhesion mechanisms.

1.3.4. IcsA is a target of autophagy

As previously stated, the IcsA passenger is targeted by host cell intracellular pathogen trapping mechanisms. However, a virulence plasmid encoded T3SS effector, IcsB, counteracts these mechanisms. IcsB was identified as a virulence factor important in the post-invasion stage of *S. flexneri* pathogenesis since ⊿*icsB* mutants form small plaques *in vitro* and do not give a positive Sereny test (Ogawa *et al*, 2003). Using fluorescence microscopy and electron microscopy of infected cell lines (such as canine epithelial cells or murine embryonic fibroblasts), ⊿*icsB* mutants were found to be covered in lamellar autophagic membranes more frequently than wild type *S. flexneri*, and were accordingly growth retarded (Ogawa *et al*, 2005). However, the growth rates of ⊿*icsB* mutants in cell lines that did not have the autophagic protein Atg5, were restored (Ogawa *et al*, 2005). Subsequent *in vitro* pull down experiments showed that $IcsA_{320-433}$ is a binding site for both Atg5 and IcsB (Ogawa *et al*, 2005). This led to the proposal that IcsA is targeted by Atg5 to initiate autophagy but must compete for binding to IcsA with IcsB which binds the same epitope (Ogawa *et al*, 2005). Thus, the passenger domain of IcsA is 'camouflaged' by IcsB. More recently it was found that IcsB can also bind cholesterol for destabilising autophagic membrane maturation, and that IcsB's role crosses over to enable escape of *S. flexneri* from vacuoles formed during cell-to-cell spread (Kayath *et al*, 2010, Campbell-Valois *et al*, 2015).

1.4. Polar localisation of IcsA

As previously outlined, the asymmetric maintenance of IcsA at the surface of one bacterial pole is critical for efficient and unidirectional *S. flexneri* spread and dissemination (Monack & Theriot, 2001). Yet, the question of what mechanism(s) drives IcsA polar placement, has remained enigmatic and proven difficult to answer. Significant conflicts in the literature, combined with hypotheses and models that have not yet been properly verified experimentally, have resulted in quite a poor understanding of the overall IcsA polarisation process. The following describes contributing factors in IcsA polarity determination and maintenance.

1.4.1. IcsA cytoplasmic polar targeting

Due to the macrostructure of rod shaped bacteria such as *Shigella* spp, there is a lateral edge with curvature about the main body of the bacterium, and two poles which have hemispherical curvature. It has been determined that bacterial poles have very inert peptidoglycan since its turnover rate is low (de Pedro *et al*, 1997). The poles can be further classified as an 'old pole' or a 'new pole'. The new pole arises from material originating from the site of septation of the mother cell and the old pole is that which is not derived from the septum of the mother cell (Dworkin, 2009, Laloux & Jacobs-Wagner, 2014). Peptidoglycan synthesis occurs during septum formation, and when division is complete it becomes as inert as the old pole (de Pedro *et al*, 1997, de Pedro *et al*, 2003). IcsA is well known for its localisation specifically to the old pole only (Robbins *et al*, 2001), therefore it is unlikely that peptidoglycan plays a role in IcsA targeting since it is equally inert at both poles (de Pedro *et al*, 1997). There is also a polar bias of certain lipids such as cardiolipin which can spontaneously localise to the inner leaflet of the polar membrane (Huang *et al*, 2006). Indeed, the polar localisation of *E. coli* osmosensory transporter ProP has been found to be reliant on cardiolipin placement at the poles (Romantsov *et al*, 2007). However, it is thought that few proteins would rely on lipid placement for localisation (Dworkin, 2009) and, again, this model is not compatible with old polar establishment of IcsA as cardiolipin has not been shown to favour a particular pole. It has also been found that some proteins can gradually accumulate at poles simply due to cell division and population expansion (Ursell *et al*, 2012, Laloux & Jacobs-Wagner, 2014). Although this likely contributes a very small driving force for IcsA polarisation, it cannot account for its highly refined polar placement. Nor can this account for the observation that IcsA can augment its own polar targeting (see further).

IcsA itself appears to contain polar targeting properties. Using immunofluorescence microscopy of *S. flexneri* expressing IcsA, the deletion $IcsA_{\Delta 509\text{-}729}$ was observed not to localise to the pole but was distributed circumferentially over the surface of the bacterium (Suzuki *et al*, 1996). This was the first indication that $IcsA_{509\text{-}729}$ provided polar targeting information or interacted with factors required in IcsA polar biogenesis. Subsequently, the use of various truncations of IcsA fused to green fluorescent protein (GFP) revealed that two regions exhibited polar targeting properties. Fusions $IcsA_{1\text{-}104}$-GFP and $IcsA_{507\text{-}620}$-GFP (nPT and cPT respectively; see Figure 1.2) were both able to form cytoplasmic foci of fluorescence at the bacterial old pole (Charles *et al*, 2001). Other experiments have been conducted where linker insertion mutagenesis was used to randomly mutagenize the length of the IcsA passenger domain; the mutants with linker insertions within the cPT caused lowered cell surface polarity comparatively to the wild-type IcsA (May & Morona, 2008). On the other hand, linker insertions within the nPT caused no measureable effect on resulting IcsA surface polarity, indicating that the role of the cPT is more significant (May & Morona, 2008). Regardless, based on these data the major initial set-up of IcsA polarity seems to occur in the cytoplasm at some point before secretion and involves the PT regions in some capacity. It has also been suggested that polar localisation may be a common theme in the family of autotransporters. For example, a range of autotransporters such as SepA, AIDA-1, and BrkA can be detected by IF microscopy at one pole when expressed in *E. coli* (Jain *et al*, 2006). Even the NalP autotransporter from *N. meningitidis* (which is spherical) is able to localise to poles during expression in *E. coli* (Jain *et al*, 2006). However, the idea that polar export is widely employed by autotransporters is not currently a generally accepted notion by the autotransporter research community (Henderson I, personal communication). IcsA has also been shown to competent in polar targeting when heterologously expressed in *E. coli, S. typhimurium*, *Y. pseudotuberculosis*, and *V. cholera*, using various lengths of IcsA passenger domain fused to GFP (Charles *et al*, 2001).

Further investigations have suggested that other factors including IcsA aggregation, and the DnaK heat-shock protein, may also be contributors to IcsA polar localisation. DnaK is a heat shock protein that works in conjunction with ClpB to form the 'bichaperone' system which is responsible for protein disaggregation and also the initial depository of aggregates to the cell pole (Carrio & Villaverde, 2005, Calloni *et al*, 2012, Castanie-Cornet *et al*, 2013, Saibil, 2013). Interestingly, in *E. coli ⊿dnaK* strains, $IcsA_{507\text{-}620}$-GFP forms irregularly shaped foci around the cytoplasm which can be rescued by complementation with *dnaK* on a plasmid (Janakiraman *et al*, 2009). Furthermore, DnaK depleted *S. flexneri*

secrete IcsA in small amounts around the cell periphery as detected by immunofluorescence microscopy (Janakiraman *et al*, 2009). In a concurrent study, the state of IcsA foci and other aggregating proteins were further elucidated. $IcsA_{507-620}$-GFP polar foci formation was studied along with a group of other fluorescent fusions to known thermo-aggregating proteins (Rokney *et al*, 2009). Here, aggregate formation of the thermo-aggregating proteins were observed to form first from small randomly placed aggregates which then coalesce into large polar aggregates (LPA) (Rokney *et al*, 2009). This too was observed for $IcsA_{507-620}$-GFP (Rokney *et al*, 2009). At the same time it was found that the formation of LPAs was abrogated in cells with disrupted *dnaK* (Rokney *et al*, 2009). Collectively, these data indicate that the IcsA PT region(s) act to aggregate IcsA into LPAs similarly to other thermo-aggregating proteins. Hence, the proposed model by this group was that IcsA self directs its own aggregation via its PT regions allowing it to be spatially positioned close to the pole (see Figure 1.5) (Rokney *et al*, 2009). At this point it is in the correct space for secretion and autotransport to be biased at the pole. Yet, for this idea to be plausible, the removal of IcsA from these aggregates must be an efficient process. Therefore the DnaK-ClpB bichaperone system may be the mediators of pulling IcsA out of these polar aggregates. Indeed, this links well with the knowledge that; (i) DnaK is required for proper IcsA biogenesis (Janakiraman *et al*, 2009), (ii) DnaK, its partner chaperone DnaJ, and ClpB move dynamically in and out of soluble aggregates (Winkler *et al*, 2010), and (iii) these chaperones are well known for their disaggregating roles (Weibezahn *et al*, 2005, Rokney *et al*, 2009, Winkler *et al*, 2010). Further, the fact that other proteins can be deposited, stored, and removed dynamically from polar aggregates for use, provides a precedent for this model (Michaelis & Gitai, 2010).

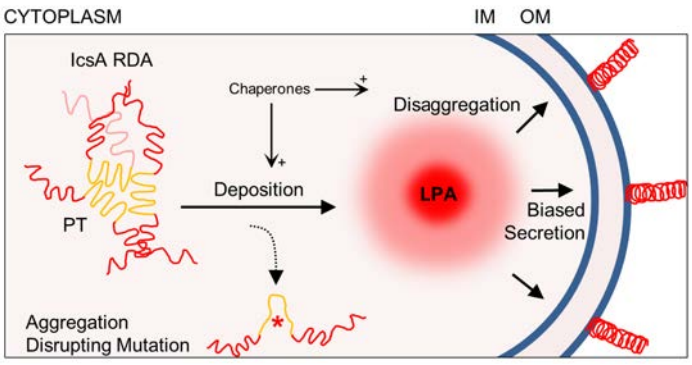


Figure 1.5: Model of biased IcsA polar secretion due to cytoplasmic polar aggregation.

In this model synthesized nascent IcsA polypeptide forms randomly distributed aggregates (RDA) which coalesce to form large polar aggregates (LPA). Deposition and disaggregation is dependent on cytoplasmic chaperones such as DnaK. Disaggregation from LPAs causes spatially biased autotransport at the old pole. The function of polar targeting (PT) regions may be to nucleate aggregation events for the formation of RDAs, and mutations in this region may disrupt this process.

Although this model is attractive and simple, the role of the PT regions in this is still uncertain. Indeed, what is the relative importance of the nPT relative the cPT? Further, what is the importance of specific PT region amino acid residues in the exertion of the polar targeting effect? These questions were therefore investigated in Chapter 6 of this thesis.

1.4.2. Other factors in IcsA polarity

Using a recently developed screening method that can screen for mutants that have aberrant subcellular localisation of IcsA, it was found that both FtsQ and RpoS are implicated in polar targeting (Fixen *et al*, 2012). IcsA, SepA, and BrkA are not polarly localised in *ftsQ* mutants (Fixen *et al*, 2012). FtsQ is a conserved cell division protein, and upon closer inspection, it was found that the periplasmic portion of FtsQ was sufficient to rescue IcsA polarity in FtsQ deficient cells (Fixen *et al*, 2012). This is an unusual finding as it indicates that: (i) FtsQ would act from its periplasmic topology to effect IcsA polar targeting in the cytoplasm or (ii) periplasmic polar targeting of IcsA may be more important than previously thought. RpoS (sigma factor S) was the other noted player in polar targeting because it's up-regulation caused a loss of IcsA polar targeting (Fixen *et al*, 2012). A model for how RpoS effects IcsA targeting was suggested whereby it may up-regulate the expression of polar targeting inhibiting proteins during its own up-regulation (Fixen *et al*, 2012). However, this speculation remains unsubstantiated thus far.

It has also been reported that when the *S. flexneri* gene *phoN2* is deleted, it causes the following phenotypes: slightly smaller plaque formation, impaired comet tail formation, and a somewhat lowered IcsA polarity on the bacterial surface (Santapaola *et al*, 2006). The later observation is questionable based on the quality of the presented immunofluorescence micrographs and the lack of any quantification. This gene encodes the protein Apyrase which is a periplasmic protein involved in hydrolysis of nucleoside di- and tri-phosphates (Santapaola *et al*, 2006). These authors went on to report that the major outer membrane protein OmpA can interact with Apyrase and that Apyrase is polarly positioned (Ambrosi *et al*, 2012, Scribano *et al*, 2014). However they were unable to show any specific interaction of OmpA with IcsA. Therefore, these findings are yet to be substantiated by other reports and the authors could not derive a mechanistic model for this phenotype besides suggesting that it may be an interacting partner of other proteins involved in IcsA polar targeting. This may be unlikely as IcsA is thought to already be polar targeted before secretion into the periplasm.

1.4.3. Proteolytic surface cleavage by IcsP

Once on the bacterial cell surface IcsA is exposed to factors that tune its polarity. Furthermore, these factors can affect the ability of IcsA to function in initiating ABM and its access to the extracellular environment. As previously mentioned, IcsA is proteolytically cleaved with low efficiency from the cell surface between R^{758} and R^{759} by the protease IcsP (also known as SopA) which affects the levels of IcsA on the bacterial surface (Fukuda et al. 1995; d'Hauteville et al. 1996). The cleavage event leaves the ~37 kDa β-barrel and linker embedded in the outer membrane, and releases the remaining ~90 kDa passenger portion of IcsA into the extracellular environment (Goldberg *et al*, 1993, Fukuda *et al*, 1995, d'Hauteville *et al*, 1996, Egile *et al*, 1997).

IcsP belongs to the Omptin family of proteases which includes the well-known OmpT and OmpP of *E. coli*. Indeed, heterogeneously expressed IcsA in *E. coli* is also cleaved by OmpT and OmpP (Goldberg *et al*, 1993). IcsP has been implicated in the maintenance and refinement of IcsA polarity as *⊿icsP* mutants are observed by immunofluorescence microscopy to have IcsA further down the lateral edge of the bacterium (although still polarly reinforced) (Egile *et al*, 1997). It is currently thought that both IcsP's localisation, and therefore its cleavage activity, is distributed evenly over the entire cell surface (Figure 1.6A) but the polar IcsA remains because IcsA is more concentrated at this site initially (Steinhauer et al. 1999). This model suggests that IcsP is important for cleavage of any IcsA that might be misdirected to the lateral edge instead of the old pole.

Unfortunately, while it is suggested that the Omptin IcsP is evenly distributed over the cell surface, there are no reports validating this hypothesis directly via experimentation. The only Omptin that has been directly analysed for surface localisation is OmpP of *E. coli* where OmpP was overexpressed and detected evenly over the surface of the cell by immuno-gold electron microscopy (Kaufmann *et al*, 1994). This is the only known distribution of an Omptin, and may be artifactual due to the over-expression. Information about the outer membrane distribution of IcsP would greatly enhance our understanding polar IcsA maintenance. The simple hypothesis would be that IcsP localises to the opposing pole of IcsA to ensure IcsA unipolarity (Figure 1.6B). Indeed, unipolarity has previously been shown to be a critical factor for virulence (Monack & Theriot, 2001). These IcsP distribution hypotheses are investigated in Chapter 3 of this thesis.

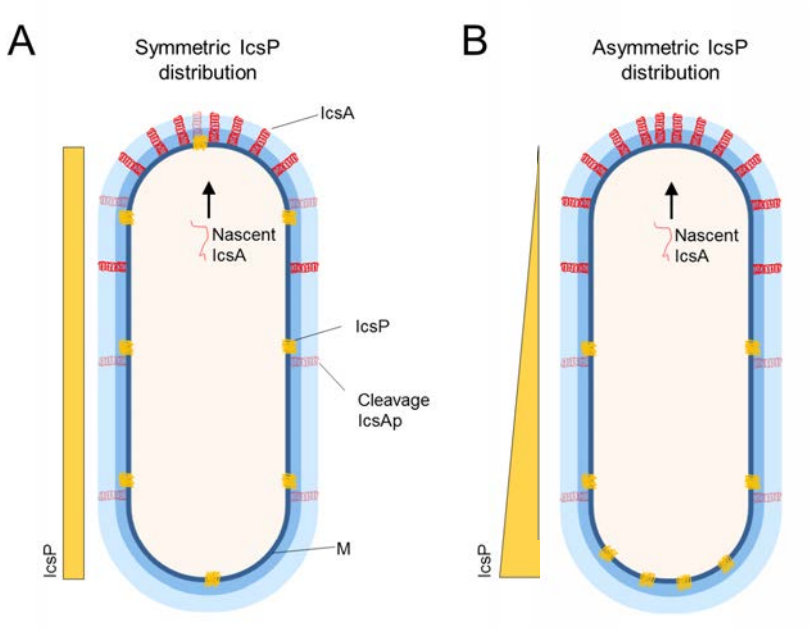


Figure 1.6: Models of IcsP cell surface distribution relative IcsA.

(**A**) The outer membrane protease IcsP may be distributed evenly over the cell surface resulting in symmetric circumferential cleavage events. (**B**) IcsP may accumulate at the pole opposing that which is occupied by IcsA resulting in asymmetrical cleavage. The IcsA pole is the 'old pole'. IcsAp is the cleaved IcsA passenger. Dark blue = outer membrane, yellow = IcsP, red = IcsA.

1.4.4. IcsA-LPS interplay

In conjunction with the action of outer membrane proteolysis, LPS is another influencing factor on IcsA in terms of its surface localisation, exposure to the extracellular milieu, and its activity. LPS is composed of three structural units: (1) a membrane embedded Lipid A, (2) a core oligosaccharide component, and (3) O-antigen that is a polysaccharide composed of polymerised oligosaccharide repeat units. Strains that have no O-antigen in their LPS structure are called rough (R)-LPS strains, those with one O-antigen unit are known as semi-rough, and smooth (S-LPS) strains have the polysaccharide O-antigen in their LPS structure (Kabanov & Prokhorenko, 2010) (Figure 1.7). Wild-type *S. flexneri* have a S-LPS phenotype which is present on the surface as a mixture of two modal lengths of O-antigen repeats: short type (11-17 repeats; ^{S}LPS) (Morona *et al*, 1995) and very long type (~90 repeats; ^{VL}LPS) (Hong & Payne, 1997). The ^{S}LPS and ^{VL}LPS phenotypes are regulated by the chromosomally encoded and plasmid encoded inner membrane $WzzB_{SF}$ and $WzzB_{pHS-2}$ co-polymerases respectively (Morona *et al*, 1995, Stevenson *et al*, 1995, Hong & Payne, 1997).

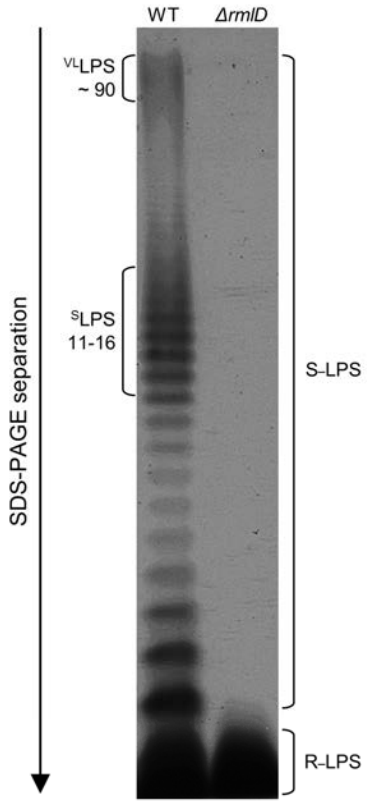


Figure 1.7: *S. flexneri* lipopolysaccharide O-antigen structure.

Silver-stained SDS-PAGE separated LPS samples. Samples are from wild-type *S. flexneri* 2a 2457T and from a ⊿*rmlD* mutant strain that is unable to synthesize TDP-rhamnose required for O-antigen biosynthesis. Band pattern is due to varied O-antigen chain lengths of the LPS. Wild-type smooth (S)-LPS has two 'modal lengths' – a short type (^{S}LPS) and a very long type (^{VL}LPS). When no O-antigen is produced, the resulting rough (R)-LPS only contains lipid A and core sugars. Numbers are O-antigen repeat units.

Immunofluorescence microscopy and immunogold electron microscopy detects IcsA along the bacterial lateral edges of R-LPS strains (but still with polar reinforcement) as opposed to the refined polar detection on wild-type (S-LPS) *S. flexneri* (Sandlin *et al*, 1996, Van Den Bosch *et al*, 1997, Robbins *et al*, 2001). One rationalisation for the perturbed polarity is that the membrane fluidity of rough strains may be higher resulting in fast diffusion of IcsA away from the pole and down the sides of the bacterium (Figure 1.8A). At least one study has indicated this where R-LPS strains apparently had their IcsA polarity restored upon treatment with chlorpromazine – a drug that can decrease membrane fluidity (Robbins *et al*, 2001). However this appears quite an unlikely scenario given there is little evidence that outer membrane proteins diffuse in the outer membrane. Rather, gram-negative bacteria grow via lateral insertion of new outer membrane material which results in a driving force that accumulates old material towards the old poles (Ursell *et al*, 2012). Therefore, any movement

of proteins in the outer membrane is like likely the result of outer membrane insertion events, not diffusion (Ursell *et al*, 2012).

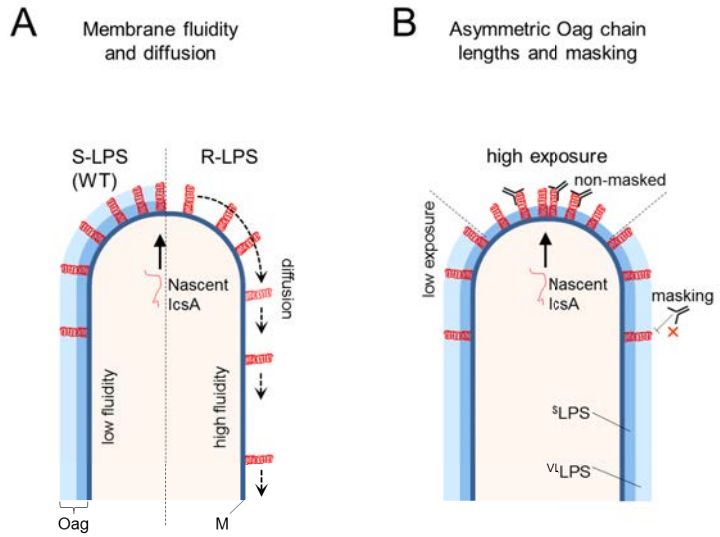


Figure 1.8: Models of LPS augmentation of IcsA surface polarity.

(A) Model of membrane fluidity and lateral diffusion. R-LPS strains may have higher membrane fluidity resulting in IcsA deposited at the pole to defuse away from the pole and down the lateral edge. **(B)** Model of asymmetrical masking due to O-antigen modal lengths distribution. It is possible that different LPS O-antigen modal chain lengths may be useful in optimizing IcsA exposure at the pole by the use of ^{S}LPS (medium blue) at the pole and ^{VL}LPS (light blue) on the lateral edges to restrict IcsA exposure by masking. M = outer membrane, red = IcsA on the bacterial surface, Oag = O-antigen, red cross = O-antigen chains can inhibit access of antibodies to IcsA via steric hindrance.

These models are further confounded by the realisation that IcsA is masked by LPS. Evidence of this was seen in a *S. flexneri* $wzzB_{SF}$*::Km* strain (only has ^{VL}LPS) where minimal IcsA could be detected by immunofluorescence microscopy (Van Den Bosch *et al*, 1997, Morona & Van Den Bosch, 2003a). This masking has been more thoroughly observed during *in situ* O-antigen hydrolysis studies where the O-antigen was removed from the surface of *S. flexneri* Y serotype using Sf6 phage tail-spike protein (TSP) (Morona & Van Den Bosch, 2003b). After removal of O-antigen from the bacterial surface by TSP, IcsA was again observed along lateral sites of bacteria as opposed to the refined polar staining of an untreated control (Morona & Van Den Bosch, 2003b). Therefore, the refined polar detection of IcsA observed on *S. flexneri* may not be the complete picture of its surface localisation. There remains the possibility that IcsA, though at high concentrations at the pole, is also present at low concentrations at lateral sites and that these sites are difficult for antibody to access due to LPS-mediated steric hindrance. It could be proposed that the reason for *S. flexneri* having two types of O-antigen lengths is to counteract the steric hindrance effect of LPS, whilst retaining protection from host defences (Figure 1.8B). In this model the very long O-antigen is required

for serum resistance and colicin resistance (Morona *et al*, 2003, Pugsley & Buddelmeijer, 2004, Scribano *et al*, 2014, Tran *et al*, 2014) whereas the short type O-antigen minimises IcsA masking such that it can access external molecules (i.e. N-WASP, or substrates for adhesion).

Given the confounding theories and models of the IcsA-LPS interplay, it is difficult to assess the actual role that LPS plays in IcsA polarity determination. Therefore, experimentation was conducted to dissect the relationship between these two molecules, and is presented in Chapter 4 of this thesis.

1.5. IcsA and Autotransport

As mentioned, once IcsA has been positioned at the pole, it is exported to the cell surface via the autotransporter pathway. This secretion of IcsA across the double-membrane cell wall must be an efficient process to allow pathogenic fitness. However, there still remain fundamental questions concerning the mechanism of autotransporter secretion – particularly concerning the stage of outer membrane translocation. Since the autotransporter family is large and has diverse functions, it is unknown how widely any one model of transport can be applied. This section reviews the knowledge and understanding regarding the biogenesis of autotransporters, and with a focus on how these findings apply to IcsA.

1.5.1. Type V secretion systems

Autotransporters were so named as they appeared to have the necessary machinery to direct their own transport onto the outer membrane surface in the absence of other participating factors (besides of course the Sec translocon). In reality, it is now known that autotransporters are reliant on many interacting partners for their biogenesis including periplasmic chaperones, and complex outer membrane translocation machinery (Leyton *et al*, 2012, Grijpstra *et al*, 2013, Bernstein, 2015). Autotransporters are sub-types of a larger Type V family of secreted virulence factors employed by Gram negative bacterial pathogens. The Type V secretion umbrella currently contains sub-types Va through to Ve (Henderson *et al*, 2004, Gawarzewski *et al*, 2014, Drobnak *et al*, 2015b). Type Va (generally just referred to as autotransporters) are known alternatively as 'classical autotransporters' or 'monomeric autotransporters'. These are polyproteins with an N-terminal signal sequence, a central passenger, and a C-terminal β-barrel connected to the passenger by a linker region (Bernstein, 2015, Drobnak *et al*, 2015b). Type Vb comprises the two-partner pathway with separately encoded passenger and translocator barrels, Type Vc are trimeric autotransporters, Type Vd appears to be a hybrid between both Va and Vb, and Type Ve are inverse autotransporters that have the β-barrel at the N-terminus and the passenger at the C-terminus (Bernstein, 2015, Leo *et al*, 2015).

As outlined in sub-section 1.3.2, IcsA has these classical autotransporter associated structures and follows this pathway for secretion. Amino acids 1 to 52 of IcsA comprise an extended signal sequence responsible for the direction of IcsA to the periplasm via the Sec secretion machinery (Goldberg et al., 1993, Suzuki et al., 1995). $IcsA_{53\text{-}740}$ comprises the functional passenger domain (Goldberg & Theriot, 1995, Suzuki *et al*, 1995), and an outer

membrane β-barrel putatively spans $IcsA_{813-1102}$. The passenger and the barrel are connected by a linker region that putatively spans $IcsA_{741-812}$.

1.5.2. Secondary and tertiary structures of autotransporters

The first autotransporter to have a portion of its structure solved by X-ray crystallography was the passenger domain of Pertactin (Prn) which is a *Bordetella pertusis* outer-membrane adhesin involved in adhesion to mammalian host cells (Emsley *et al*, 1996). There have been 14 additional reports of solved autotransporter structures since this time. Autotransporter passenger domains (other than Prn) that have been solved (chronologically listed) are: Haemoglobin protease (Hbp) of pathogenic *E. coli* (Otto *et al*, 2005), a *Helicobacter pylori* toxin called VacA (Gangwer *et al*, 2007), IgA1 protease (IgA1P) from *Haemophilus influenzae* (Johnson *et al*, 2009), an esterase (EstA) from *Pseudomonas aeruginosa* (Van den Berg, 2010), the protease stable portion of IcsA (Kuhnel & Diezmann, 2011), another adhesin called Hap from *H. influenzae* (Meng *et al*, 2011), an extracellular serine protease (EspP) of *E. coli* O157:H7 (Khan *et al*, 2011), the 'velcro-like' Ag43 of pathogenic *E. coli* (Heras *et al*, 2014), another toxin Pet from enteroaggregative *E. coli* (Domingo Meza-Aguilar *et al*, 2014), and finally the adhesin TibA from enterotoxigenic *E. coli* O78:H11 (Lu *et al*, 2014). These are shown in Figure 1.9A and B.

One commonality between the majority of autotransporters is the prevalence of a right-handed β-helical tertiary structure for all or part of the passenger domain. Every autotransporter structure solved to date has this β-helix except EstA (Van den Berg, 2010). Indeed, sequence analysis of over 500 putative autotransporter sequences predicted that more than 97% of these possess a right-handed β-helical structure (Junker *et al*, 2006). Hence, the sequence analysis fits in with the observed dominance of the β-helical feature in solved passenger structures. Modelling has also identified three major cross-sectional shapes of these β-helices; triangular, L-shaped, and oval (Kajava & Steven, 2006). Autotransporters such as Prn and TibA contain the β-helix as its entire passenger domain structure, whereas others (i.e.: Hbp, IgA1P, Hap, and EspP) use the helix as a base scaffold off-which globular domains can be projected. These globular domains usually contain active sites required for protein activity such as the proteolytic site of IgA1P (Johnson *et al*, 2009). EstA is seemingly an oddity as its passenger domain does not contain the common β-helix structure but is composed entirely of a globular structure with many α-helices (Van den Berg, 2010) (Figure 1.9B).

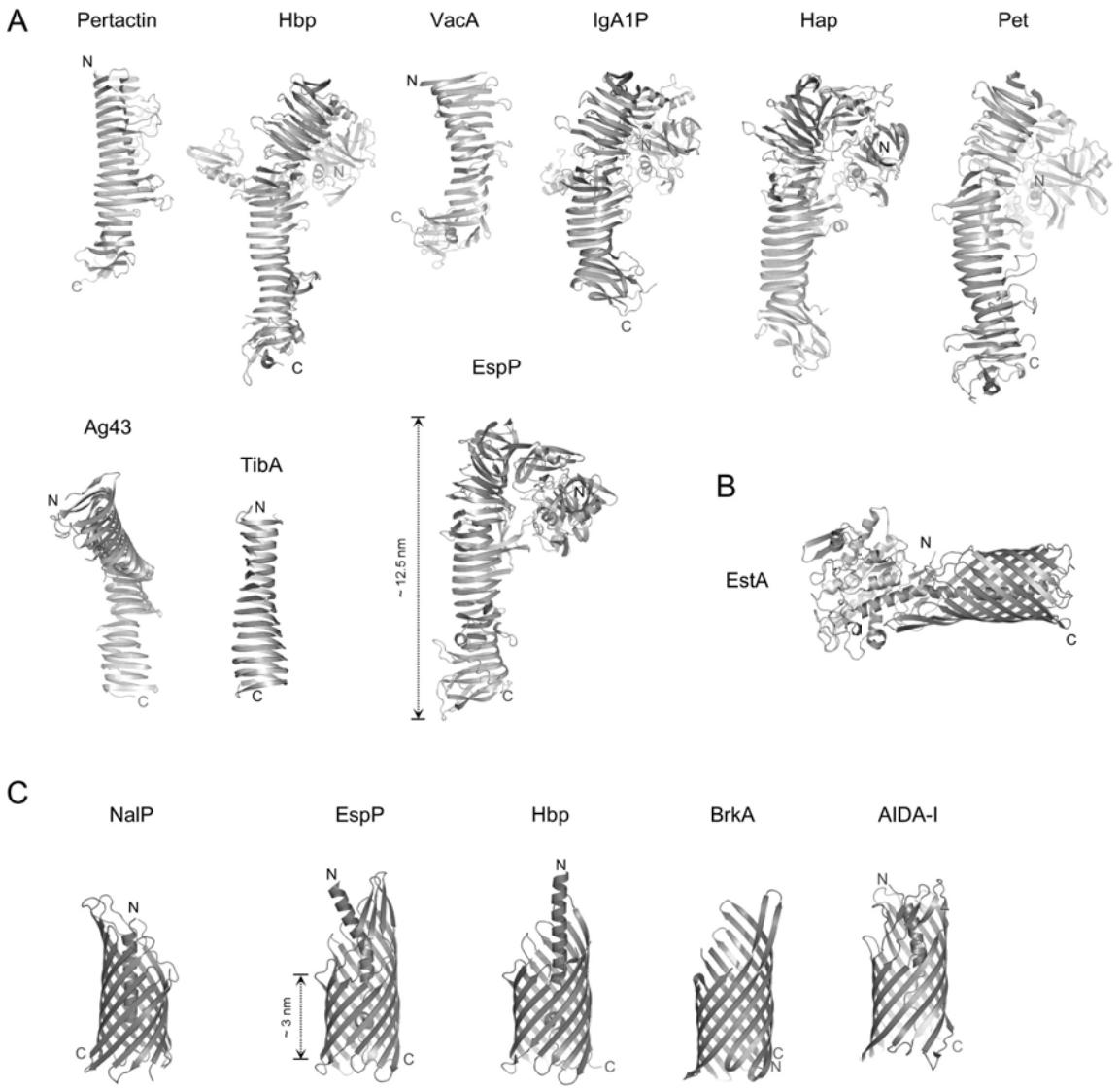


Figure 1.9: Autotransporter passenger and barrel crystal structures.

(**A**) Ribbon diagrams of solved autotransporter passengers: *B. pertussis* Pertactin (1DAB), *E. coli* Hbp (1WXR), *H. pylori* VacA (2QV3), *H. influenzae* IgA1 (3H09 and Hap (3SYJ), *E. coli* Pet (4OM9), Ag43 (4KH3), and TibA (4Q1Q). The base structure is commonly a β-helical stalk with loops occasionally harbouring globular functional domains (**B**) The full length *P. aeruginosa* autotransporter EstA has also been solved (3KVN). (**C**) β-barrels for five autotransporters have been solved: *N. meningitidis* NalP (1UYN), EspP (multiple structures, pre-cleavage 3SLJ), Hbp (3AEH), *B. pertussis* BrkA (3QQ2), and *E. coli* AIDA-I (4MEE). All autotransporter β-barrels are 12-stranded and nearly completely superimposable.

The structure of IcsA's passenger domain (IcsAp) has not yet been completely solved. Only a small portion, the stable core ($IcsA_{591-740}$), has been crystallised (Kuhnel & Diezmann, 2011). Even though this is an incomplete structure, it does show that at least this portion of IcsA adopts the common right-handed β-helical structure with a triangular cross-section (Figure 1.10A). Furthermore, $IcsA_{561-740}$ has a distinctive β-hairpin at the extreme C-terminus. This hairpin motif is common to passengers that are β-helix based such as Prn,

Hbp, IgA1P, Hap, and EspP (Emsley *et al*, 1996, Otto *et al*, 2005, Johnson *et al*, 2009, Khan *et al*, 2011, Meng *et al*, 2011). Even VacA's C-terminal passenger portion (p55) has this hairpin, despite the fact it has an unusually high amount of α-helices (Gangwer *et al*, 2007). Due to the incidence of the hairpin at the extreme C-terminus of many passengers, it has been proposed that this structure may contribute to secretion stages of biogenesis, passenger domain stability, and folding (Oliver *et al*, 2003, Renn & Clark, 2008, Soprova *et al*, 2010, Leyton *et al*, 2012).

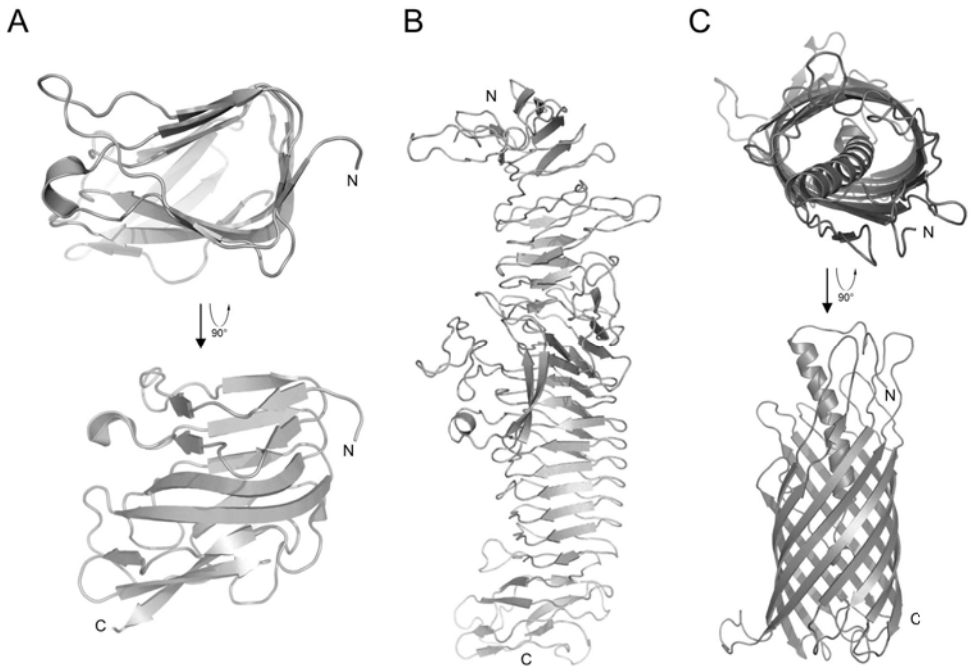


Figure 1.10: Solved structure of the IcsA stable core and putative passenger and barrel structures.

(**A**) Top-down and side views of the stable core portion from the IcsA passenger (3ML3) showing a right-handed β-helical structure with a triangular cross-section. (**B**) The IcsA region 53-735 was modeled by threading and *ab initio* modelling using I-TASSER (zhanglab.ccmb.med.umich.edu/I-TASSER) for Chapter 6: Research Article 4. The model contains primarily β-helical stalk structure similarly to solved passenger structures. (**C**) The β-barrel domain and linker ($IcsA_{759\text{-}1102}$) was also modelled using I-TASSER. Top-down and side views show putative 12-stranded barrel with embedded α-helix for a portion of the linker. C-score = -0.21, contributing templates: Hbp (3AEH), BrkA (3QQ2), EspP (3SLJ), and EstA (3KVN).

The outer membrane embedded β-barrel domains of autotransporters have common structural themes evidenced by crystallographic studies (Figure 1.9C). The β-barrel domain structures solved to date are: NalP from *Neisseria meningitidis* (Oomen *et al*, 2004), EspP (Barnard *et al*, 2007), EstA (Van den Berg, 2010), Hbp (Tajima *et al*, 2010), the *B. pertussis* serum resistance protein BrkA (Zhai *et al*, 2011), and *E.coli* diffuse adherence protein AIDA-I (Gawarzewski *et al*, 2014). All solved β-barrel domains have twelve antiparallel β-stands which form β-barrel cylindrical structures. These barrels possess hydrophilic loops and a

hydrophobic 'belt' which stabilises it in the outer membrane (Barnard *et al*, 2007). The interior of these barrels are hydrophilic and usually contain an α-helical 'plug' part of the linker (Barnard *et al*, 2007, Leyton *et al*, 2012). It should be noted here that the pore size of these barrels is ~1 nm which is too small to accommodate the appropriate passage of the nascent passenger which can sometimes include disulphide bonded loops (Brandon & Goldberg, 2001, Leyton *et al*, 2011). It has been estimated that the active pore size may be approximately twice this size (Sauri *et al*, 2012).

Although the IcsA β-barrel domain has not been solved, it is likely that it too adopts a similar barrel structure as the above mentioned autotransporters. Indeed, early computer based studies of IcsA also identified a putative antiparallel β-strand configuration with at least eleven strands as well as an α-helical linker (Suzuki *et al*, 1995). *In silico* modelling of the β-barrel and linker conducted during this work also shows the likelihood of a 12-stranded structure with a pore embedded α-helix (Figure 1.10C).

1.5.3. The autotransport pathway and IcsA

The first step in the secretion of IcsA is its translocation across the inner membrane to the periplasm. Autotransporters, like many secreted proteins, are targeted to the SecYEG translocon channel (inner membrane translocation machinery) via N-terminal signal sequences containing N, H, and C regions (Leyton *et al*, 2012). These are transported across the inner membrane upon ATP hydrolysis by SecA (Henderson *et al*, 2004, Leyton *et al*, 2012). Generally, proteins destined for the inner membrane are co-translationally targeted to the SecYEG translocon by signal recognition particle (SRP), and proteins destined to be periplasmic or outer membrane proteins are post-translationally targeted to SecYEG by SecB chaperone (van Ulsen, 2011). IcsA's requirement for SecYEG has been directly established (Brandon *et al*, 2003), and it is probable that IcsA is post-translationally translocated, presumably via SecB targeting.

As previously mentioned, IcsA has an extended signal sequence (ESS) which contains a lengthened N-terminal stretch in addition to the normal N, H, and C regions (Brandon *et al*, 2003). Only ~10% of known autotransporters contain an ESS (Dautin & Bernstein, 2007), and all proteins containing an ESS are found in Gram negative Beta- or Gamma-proteobacteria (Desvaux *et al*, 2006). The role of IcsA's ESS is unknown and the roles of ESS from other autotransporters are still debated. Upon investigation of the ESS of EspP it was found that deletion of the extension does not impair translocation across the inner

membrane (Szabady *et al*, 2005). However, it did cause the efficiency of EspP release from the SecYEG translocon to be compromised and the amount of terminally processed EspP to be reduced (Szabady *et al*, 2005). From this, the authors suggested that the ESS acts as an 'anchor' by increasing the time bound to the Sec machinery and possibly limiting access of signal peptidases (Szabady *et al*, 2005). Furthermore, they speculated that this increased time spent at the inner membrane allows binding of periplasmic chaperones which can prevent misfolding (Szabady *et al*, 2005). However, an investigation on Pet ESS noted that Pet with a deleted extension is secreted folded and functional (Leyton *et al*, 2010). This indicates that the stalling caused by native Pet ESS (Desvaux *et al*, 2007) does not contribute to its stability in the periplasm (Leyton *et al*, 2010). Therefore, the importance of the ESS remains to be definitively answered. Indeed, nothing is known about the importance of IcsA's signal sequence, and upon BLAST searching $IcsA_{1-52}$ as a query, no hits are observed within Type V secreted proteins besides IcsA itself (Desvaux *et al*, 2006). Thus, there remains the possibility that the ESS of IcsA possess a novel function. However, it is certain that once transported across the inner membrane, IcsA is cleaved (probably by signal peptidase) between amino acids A^{52} and T^{53} releasing it into the periplasm (Goldberg *et al*, 1993, Suzuki *et al*, 1995) (Figure 1.11(i)).

Once in the periplasm, IcsA is escorted by various periplasmic molecular chaperones (Figure 1.11(ii)). For instance, it was shown that *S. flexneri degP* mutants exhibit aberrant plaque formation (Purdy *et al*, 2002). Furthermore, these strains insert IcsA into the outer membrane but the passenger domain has abrogated ability to be translocated and surface exposed (Purdy *et al*, 2002, Purdy *et al*, 2007). Hence, periplasmic DegP (which has both chaperone and proteolytic functions) seems to have a role in keeping IcsA in a translocation competent state. Subsequently, DegP, Skp and SurA periplasmic chaperones have also been implicated in the prevention of IcsA misfolding in the periplasm allowing it to remain translocationally competent (Purdy *et al*, 2007, Wagner *et al*, 2009). Skp is an important heterotrimeric anti-aggregation chaperone, and SurA acts to catalyse isomerisation steps of periplasmic proteins (and autotransporters) and recognises the residue motif: aromatic-X-aromatic (Walton & Sousa, 2004, Ruiz-Perez *et al*, 2009). Using various interaction experiments such as yeast two-hybrids and surface plasmon resonance analysis, it has been found that autotransporters generally require many periplasmic chaperone interactions. Indeed, EspP, also interacts with DegP, Skp, and SurA similarly to IcsA, and in the absence of DegP, EspP is inserted in to the outer membrane in a misfolded manner (Ruiz-Perez *et al*, 2009). DegP and SurA interactions have also been shown for Hbp where DegP switches to a protease character to degrade stalled Hbp intermediates (Jong *et al*, 2007). Additionally, other

chaperones including FkpA and VirK have been implicated in the pathways of the EspP passenger and Pet respectively (Ruiz-Perez *et al*, 2010, Tapia-Pastrana *et al*, 2012). Therefore it is evident that periplasmic chaperones are widely required for maintaining nascent autotransporters in an export competent state.

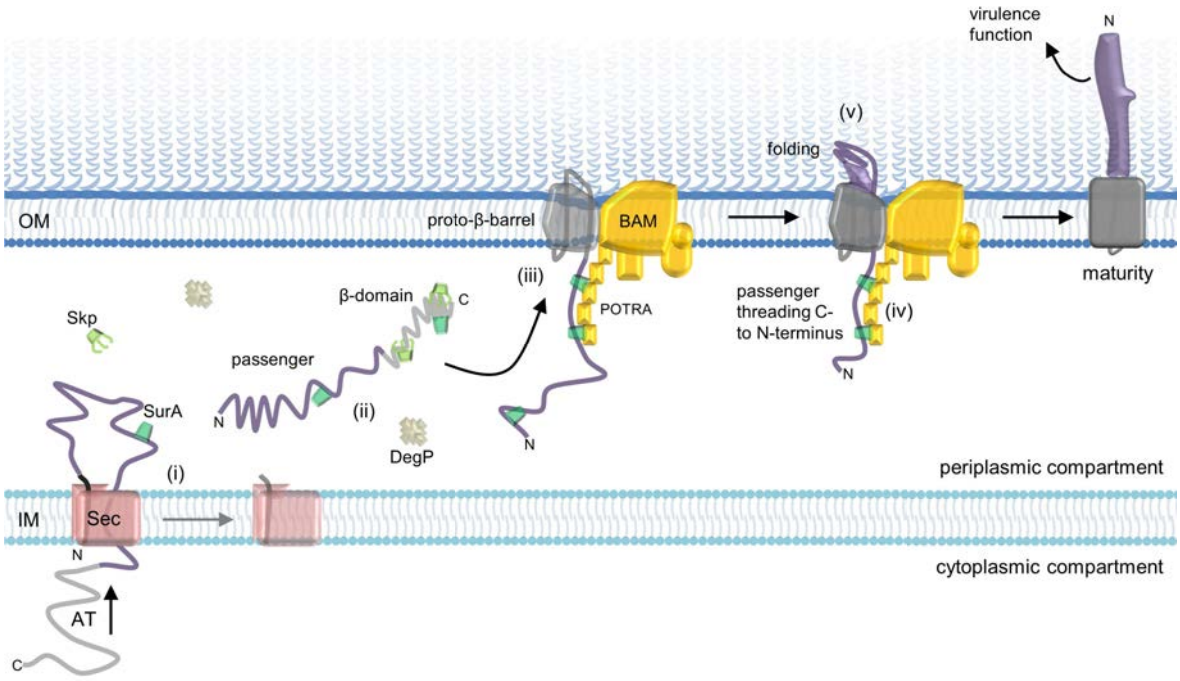


Figure 1.11: Model of secretion to the cell surface by the Autotransporter pathway.

The nascent autotransporter (AT) is targeted to the Sec translocon. Sec secretion passes the autotransporter across the inner membrane (IM) and into the periplasmic compartment (**i**). The signal sequence (black) is cleaved off. In the periplasm the autotransporter passenger (purple) and β-barrel (grey) is bound by various chaperones including Skp, DegP, and SurA to hold it in an export competent state (**ii**). The β-domain forms as a partly folded proto-β-barrel and requires the β-Barrel Assembly Machinery (BAM) for outer membrane (OM) insertion (**iii**). SurA and the passenger interact with the POTRA domains of BamA as the passenger is translocated from C- to N-terminus progressively (**iv**). The folding of the passenger portions already translocated may drive the rest of the translocation (**v**). Autotransporters frequently require a further step involving the cleavage of the passenger from the β-barrel which can occur intra - or inter-molecularly (not shown).

For outer membrane insertion and translocation, autotransporters (including IcsA) then require appropriate and timely interactions with the β-Barrel Assembly Machinery (BAM) (Voulhoux *et al*, 2003, Jain & Goldberg, 2007, Ruiz-Perez *et al*, 2009, Sauri *et al*, 2009, Peterson *et al*, 2010, Leyton *et al*, 2011, Rossiter *et al*, 2011). The BAM is a multi-protein complex consisting of an integral 16-stranded β-barrel called BamA (Omp85/YaeT) which has periplasmic polypeptide transport-associated (POTRA) domains, and associated lipoproteins BamB-E (Noinaj *et al*, 2013, Noinaj *et al*, 2015). This complex (particularly BamA and BamD) are essential for autotransporter outer membrane insertion and translocation (Rossiter *et al*, 2011, Roman-Hernandez *et al*, 2014). The autotransporter β-barrel domain, held by

interactions with Skp, is delivered to the BAM with first interactions possibly occurring with BamB and BamD (Ieva *et al*, 2011, Pavlova *et al*, 2013). At some point it is likely that BamA recognises an 'OMP-substrate' motif (Robert *et al*, 2006) found at the C-terminal extreme of the β-barrel domain of autotransporters and other outer membrane proteins (IcsA also has this; ^{1094}GILGVKYTF^{1102}) (Celik *et al*, 2012). Regardless, a proto-β-barrel structure (which has a non-final conformation) forms in conjunction with interactions with BamA, but includes the C-terminus of the passenger domain as a hair-pin (Ieva *et al*, 2008, Ieva & Bernstein, 2009, Peterson *et al*, 2010, Sauri *et al*, 2012, Pavlova *et al*, 2013) (Figure 1.11(iii)). Further, it has been realised that adjacent strands of the autotransporter β-barrel have conserved motifs that (for at least one pair) can link in a manner analogous to a mortise-tenon joint resulting in stabilisation the required positioning of a portion of the passenger within the barrel prior to translocation (Yen *et al*, 2010, Celik *et al*, 2012, Pavlova *et al*, 2013, Leyton *et al*, 2014). Intriguingly, recently solved structures of BamA suggest that the first and final strands may be able to open laterally and reduce the local outer membrane depth (Noinaj *et al*, 2013, Noinaj *et al*, 2014). From this, it has been speculated that the membrane disruption may provide an easy pathway for proto-β-barrel formation and integration, and that the lateral opening may allow the formation of a large autotransporter-BamA active translocation pore (Bernstein, 2015). This would also solve the issue of size requirements of the actively translocating pore versus the final mature plugged pore observed for the solved structures (Skillman *et al*, 2005, Jong *et al*, 2007, Leyton *et al*, 2011, Sauri *et al*, 2012). Besides the requirement for BAM, most of outer membrane insertion part of the pathway has not been directly assessed for IcsA. However, considering the conservation of the β-barrel and the accessory components, it is likely that IcsA follows much, if not all, of the above steps.

Whilst inserted, autotransporter passengers must be sleeved though the active β-barrel pore and correctly folded on the cell surface. This occurs in a C- to N-terminal manner (Ieva & Bernstein, 2009, Junker *et al*, 2009) and likely requires continued passenger-BamA/POTRA interaction events mediated by SurA (Ieva & Bernstein, 2009, Bennion *et al*, 2010) (Figure 1.11(iv)). Since this must occur in isolation of external energy inputs (such as from ATP hydrolysis), there must be alternative forces that drive this process and prevent sliding of the nascent passenger back into the periplasm. It is now quite well established that the C- to N-terminal β-helical 'vectorial folding' itself drives passenger translocation (Junker *et al*, 2006, Braselmann & Clark, 2012) (Figure 1.11(v)). In this model, the initially translocated C-terminal portion of the passenger domain usually has high stability and folds with high efficiency preventing back sliding (Junker *et al*, 2006, Renn & Clark, 2008, Peterson *et al*, 2010,

Renn *et al*, 2012, Kang'ethe & Bernstein, 2013a). Studies on autotransporters (including Pertactin, Pet, and IcsA) have reported a stable core in the C-terminus of the passenger that contributes to efficient and appropriate passenger folding (Junker *et al*, 2006, May & Morona, 2008, Renn & Clark, 2008, Junker *et al*, 2009, Kuhnel & Diezmann, 2011). Yet, due to the incredible diversity in the size and sequence of autotransporter passengers (Renn & Clark, 2008, Celik *et al*, 2012), it not known how widely a vectorial folding model is adhered to. Indeed, the autotransporter YapV has a slight alteration of the model as it lacks a C-terminal (Pertactin-like, PL) stable core region but uses an unstable region at the extreme N-terminus such that it does not fold prematurely in the periplasm (Besingi *et al*, 2013). Furthermore, it has been realised that the vast majority of ATs do not contain a PL region (Drobnak *et al*, 2015b), and it is uncertain how the model might apply to very large autotransporters, those that possess globular passengers, or those that have other unusual sequence features. It has also been suggested that very large autotransporters may require as yet unidentified secretion driving forces (Drobnak *et al*, 2015a).

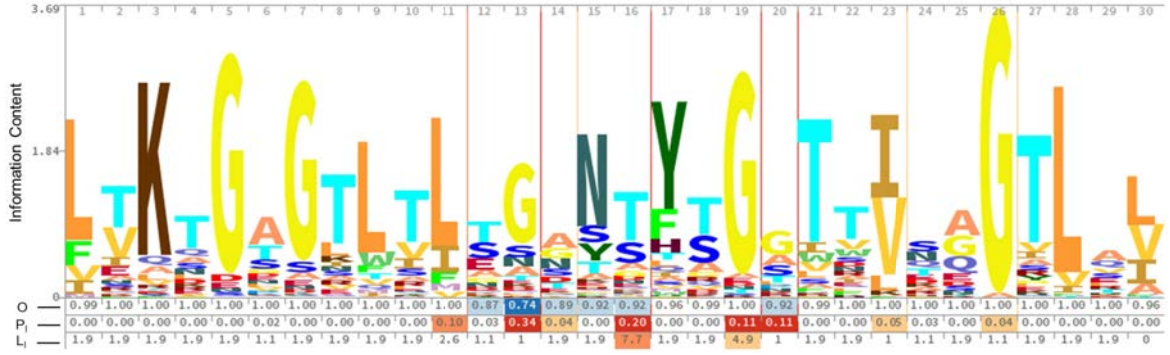


Figure 1.12: Sequence model of a conserved autotransporter repeat motif.

Pfam Hidden Markov Model (HMM) logo (pfam.xfam.org/family/PF12951) of a conserved autotransporter repeat motif with unknown function. This is present in IcsA comprising the region 526-557 (see further Figure 5.1). The letters divide the stack height according their estimated probability of occurrence contributing to the information content. Occupancy (O) = probability of observing the letter instead of an insert. Insert probability (P_I) = probability of insert between this, and the next, letter. Insert length (L_I) = probable insert length if one occurred.

Therefore, it certainly remains possible that mechanisms of passenger secretion are specifically augmented depending on the size, fold-type, function, or modifying sequence feature of the autotransporter in question. After considerable scrutiny of common databases conducted during the development of this Doctoral work (for example the UniProt Knowledgebase), I identified multiple conserved but uncharacterised sequence features and motifs within the passengers of autotransporters. The most interesting of these was simply designated as an 'autotransporter β-repeat' and pictured in Figure 1.12. Although models of

this repeat had been deposited into both TIGRFAM (TIGR02601) and Pfam (PF12951) conserved model databases, and it appeared to be present in thousands of autotransporter virulence factors, its function remained completely uncharacterised. Importantly, IcsA contains one of these repeats ($IcsA_{526-557}$; Figure 1.2) making it an ideal model autotransporter for the investigation of this repeat. Thus, this became a key topic of investigation which is presented in Chapter 5 of this thesis.

1.6. Research Questions and Aims

It is important to understand how bacterial pathogens precisely place and organise their virulence factors. This is true for the *S. flexneri* virulence factor IcsA, which must be both polarly localised and secreted to the cell surface, before it can appropriately contribute to virulence. However, the question of what mechanism distributes IcsA to the pole, has remained unanswered. Hypotheses and models have been proffered, but have not yet been properly verified. While it is suggested that the protease of IcsA (IcsP) is evenly distributed over the cell surface, there are no reports validating this hypothesis directly. Further, the confounding nature of the models of LPS augmentation of IcsA polarity prompts the requirement for appropriate experimental validation. Additionally, the role of the PT regions in this is still uncertain. What is the importance of the nPT relative the cPT, and what is the importance of specific amino acid residues in these regions in augmenting polar localisation? Regarding autotransport, there remains many uncharacterised or under-characterised conserved autotransporter sequence features. For IcsA, the most pressing question is – what is the role of the 'autotransporter β-repeat' sequence?

1.6.1. Thesis Research Aims

- **To investigate the maintenance of IcsA polar localisation at the bacterial surface.**
  - To discover the localisation of IcsP with respect to the localisation of IcsA over the bacterial outer membrane.
  - To determine whether IcsA present on the surface of *S. flexneri* is masked asymmetrically by lipopolysaccharide O-antigen chains.
  - To ascertain whether there is a difference in IcsA polarity between S-LPS and R-LPS expressing *S. flexneri*.

- **To investigate IcsA-dependent regulation of IcsA polarity.**
  - To re-examine the importance of the N-terminal and central polar targeting regions (nPT; $IcsA_{1-104}$ and cPT; $IcsA_{506-620}$ respectively) in IcsA-dependent polarisation.
  - To identify amino acid residues required for appropriate IcsA-dependent polar targeting.

- **To investigate the role of conserved passenger elements in IcsA autotransport.**
  - To determine the function of the uncharacterised 'autotransporter β-repeat' sequence $IcsA_{526-557}$.
  - To explore the wider importance of the repeat sequence in other autotransporters.

Chapter Two

Materials and Methods

Chapter 2: Materials and Methods

2.1. Growth Media

2.1.1. Liquid broth growth media

All *E. coli* and *S. flexneri* strains were routinely grown in Luria Bertani (LB) broth [10 *g* L^{-1} tryptone (Becton, Dickinson and Co.; BD), 5 *g* L^{-1} yeast extract (BD), 5 *g* L^{-1} NaCl]. Bacterial concentration was measured by optical absorbance at 600 nm (OD_{600}) where an OD_{600} reading of one was equivalent to 5 x 10^8 bacteria mL^{-1}.

2.1.2. Solid growth media

For spread, streak, and patch plating, all *E. coli* strains were grown on LB agar [LB, 15 *g* L^{-1} agar (BD)] and all *S. flexneri* strains were grown on Tryptic Soy (TS) agar [15 *g* L^{-1} agar (BD), 30 *g* L^{-1} TSB Soybean-casein digest (BD)].

2.1.3. Selection, differentiation, and additives

For selection, the following antibiotics were used as appropriate; ampicillin at 50-100 μg mL^{-1}, tetracycline at 10 μg mL^{-1}, kanamycin at 50 μg mL^{-1}, chloramphenicol at 25 μg mL^{-1}, and gentamycin at 40 μg mL^{-1}. For differentiation of *S. flexneri* expressing virulence-plasmid proteins, Congo Red dye was added to TS agar at concentration of 0.01 % (w/v). When grown at the permissive temperature of 37 °C, virulence-plasmid positive isolates develop red colonies and virulence-plasmid negative isolates develop white colonies. For blue-white colony screening when cloning through pGEM-T Easy (Promega), LB agar plates were supplemented with substrate 5-bromo-4-chloro-3-indolyl-β-D-galactoside (X-gal) at 20 μg mL^{-1} and inducer isopropyl-β-D-thiogalactopyranoside (IPTG) at 0.5 mM. When using constructs controlled by a P_{BAD} promoter, expression was repressed with 0.2 % (w/v) glucose and induced with 0.2 % (w/v) L-arabinose.

2.2. Bacterial Strains

2.2.1. Lists of strains and plasmids

All *E .coli* and *S. flexneri* strains utilised or constructed in this work are listed in Appendix B and Appendix C. Each chapter also contains a list of bacterial strains and plasmids used that pertain to that chapter (Chapter 3, Table 3.1; Chapter 4, Table 4.1; Chapter 5, Table 5.1S,: Chapter 6, Table 6.1).

2.2.2. Storage of strains and plasmids

Strains were grown on solid media as a lawn, harvested by a sterile loop, and stored at -80 ˚C in 30 % (v/v) glycerol, 1 % (w/v) peptone. Freeze-thawing was avoided when retrieving strains from glycerol storage. All constructed plasmids were stored in *E. coli* DH5α.

2.3. Antibodies and Antisera

Polyclonal affinity purified rabbit anti-IcsA, rabbit anti-N-WASP, rabbit anti-IcsP, rabbit anti-$WzzB_{SF}$, and rabbit anti-$WzzB_{pHS2}$ were all produced in-house and validated as described previously (Van Den Bosch *et al*, 1997, Daniels & Morona, 1999, May & Morona, 2008, Purins *et al*, 2008, Teh *et al*, 2012a). Polyclonal rabbit anti-SurA and rabbit anti-BamA were generous gifts from Carol Gross (University of California) and Thomas Silhavy (Princeton University) respectively. Monoclonal mouse anti-HA (H3663) and monoclonal mouse anti-DnaK (ADI-SPA-880) were purchased from Sigma and Enzo Life Sciences respectively. For western immunoblotting, horse radish peroxidase conjugated goat ant-rabbit or goat anti-mouse secondary antibodies were used (KPL). For immunofluorescence microscopy, donkey anti-mouse Qdot 525, donkey anti-rabbit Qdot 625, donkey anti-rabbit AlexaFluor 488, and donkey anti-rabbit AlexaFluor 594 secondary antibodies were purchased from Life Technologies.

2.4. DNA Techniques

2.4.1. Plasmid isolation

Plasmid DNA was purified from over-night bacterial cultures (LB, 10 mL) according to the QIAprep Spin Miniprep kit (Qiagen) protocol. MilliQ (MQ) H_2O (18.2 MΩcm) was used for DNA elution and sample was stored at -4 °C.

2.4.2. DNA quantitation

NanoDrop 2000c Spectrophotometer (Thermo Scientific) was used for measurements of DNA sample concentration by absorption at 260 nm.

2.4.3. Restriction endonuclease digests

All restriction enzymes were purchased from New England Biolabs (NEB) and used according to the supplied protocols. Following digestion, restriction enzymes were heat inactivated where possible. Otherwise restriction enzymes were removed using a PCR clean-up kit.

2.4.4. Oligonucleotides

All oligonucleotides were ordered from Integrated DNA Technologies (IDT), resuspended in MQ H_2O to a storage stock concentration of 100 μM, and stored at -20 °C. Oligonucleotides were diluted to a stock of 10 μM before use in polymerase chain reactions (PCR) at 0.5 – 2 μM.

2.4.5. PCR

PCR reactions were conducted according to the supplied protocols using Taq DNA polymerase from NEB in a 20 μL volume. Deoxynucleic triphosphates (dNTPs) were purchased from Sigma and used at a final reaction concentration of 200 μM. An Eppendorf Mastercycler Gradient thermocycler was used for all reactions with standard cycles of 30 x denaturation (98 °C, 30 seconds), annealing (55 °C, 30 seconds), and extension (68 °C, 1 minute kb^{-1}).

In general, *Taq* DNA polymerase was only used for screening purposes. For cloning and maintenance of sequence fidelity, Phusion *Pfu* High-Fidelity DNA Polymerase (NEB) was used as it has 50-fold lower error rate compared to *Taq*.

2.4.6. Electrophoretic separation of DNA and visualisation

For DNA size analysis, loading dye [1 mg mL^{-1} bromophenol blue, 20 % (v/v) glycerol, 0.1 mg mL^{-1} RNase] was added to DNA samples at a 10 times dilution. Samples were separated with 0.8, 1, or 1.2% (w/v) agarose TBE [70 mM Tris, 20 mM boric acid, 1 mM EDTA] gels supplemented with the required volume of 20,000 x RedSafe nucleic acid staining solution (iNtRON Biotechnology). Size makers used were SPP1 phage DNA EcoRI fragments made in house (sizes [kb]: 8.51, 7.35, 6.11, 4.84, 3.59, 2.81, 1.95, 1.86, 1.51 1.39, 1.16, 0.98, 0.72, 0.48 0.36 and 0.09). Gels were run at 100 V for 30 minutes and visualised using a GelDoc XR sytem (Biorad).

2.4.7. DNA purification

2.4.7.1. PCR product purification

A QIAquick PCR purification kit (Qiagen) was used according to the supplied protocols for purifying PCR products following PCR. DNA was eluted in 20 µL of MQ H_2O and stored at -4 °C.

2.4.7.2. Gel DNA extraction

A QIAquick Gel Extraction kit (Qiagen) was used according to the supplied protocols for purifying DNA fragments following electrophoretic separation. DNA was eluted in 20 µL of MQ H_2O and stored at -4 °C.

2.4.8. DNA sequencing

Oligonucleotide primer was added at a final concentration of 0.8 µM to 1000 – 1500 ng of purified DNA and adjusted to a volume of 12 µL using MQ H_2O. The samples were sequenced by the Australian Genome Research Facility (AGRF) sequencing service.

2.4.9. Oligonucleotide annealing for insertions

1 nmol of each of the oligonucleotides to be annealed were added to annealing buffer [100 mM potassium acetate, 30 mM HEPES, 2 mM magnesium acetate, pH 7.5] for a volume of 50 µL. Samples were then cycled through temperatures in an Eppendorf Mastercycler Gradient thermocycler (98 °C, 5 minutes; 85 °C, 5 minutes; 70 °C, 10 minutes) then left to cool to room temperature. Samples were then further cooled with 5 minute cycles at 15 °C, 10 °C, and 4 °C. Annealed fragments were then phosphorylated and ligated to dephosphorylated plasmids.

2.4.10. DNA Phosphorylation

Linear DNA fragments were 5' phosphorylated using T4 polynucleotide kinase (NEB) according to the supplied protocols. Samples were then heat treated to 70 °C for 20 minutes to deactivate T4 polynucleotide kinase.

2.4.11. DNA Dephosphorylation

Linear DNA fragments were dephosphorylated using Antarctic phosphatase (NEB) according to the supplied protocols. Samples were then heat treated to 70 °C for 20 minutes to deactivate Antarctic phosphatase.

2.4.12. DNA cohesive- and blunt-end ligations

Ligation of linear DNA fragments was conducted using T4 DNA Ligase (NEB) according to the supplied protocols. Samples were incubated at 25 °C for 2 hours or 16 °C over night.

2.4.13. Cloning into pGEMT-Easy

DNA fragments that were amplified with Phusion *Pfu* polymerase were first treated with *Taq* DNA polymerase to introduce 3' A-tails according to the supplied protocols (Promega). A-tailed fragments were then ligated into pGEMT-Easy using the supplied protocols (Promega).

2.4.14. Deletions via Inverse PCR

For targeted in-frame deletion of protein coding regions within plasmids, oligonucleotide primers were designed to anneal immediately adjacent the region to be deleted. The entire plasmid was PCR amplified (using Phusion *Pfu)* with extension proceeding outwards from the region to be deleted, thus excluding it from the amplicon. The amplicon was simply phosphorylated and blunt end ligated.

2.4.15. Site directed mutagenesis

Single amino acid substitutions were conducted using the QuikChange Lightning Site-directed Mutagenesis kit (Agilent Technologies) according to the supplied protocols.

2.4.16. Domain-targeted random mutagenesis

To generate random mutants specifically over a designated region, the GeneMorph II EZclone Domain Mutagenesis kit (Agilent Technologies) was used according to the supplied protocols. Random mutant plasmid libraries were created by pooling all transformant colonies from the mutagenesis protocol into a glycerol stock and stored at -80 °C.

2.5. Bacterial Transformations

2.5.1. Preparation of chemically competent *E. coli*

Mid-exponential phase (OD_{600} ~ 0.5) *E. coli* grown in 10 mL LB broths were harvested by centrifugation (2,200 x *g*, 10 min, 4 °C), washed in 5 mL of ice-cold 100 mM $MgCl_2$, resuspended in 1 mL of ice-cold 100 mM $CaCl_2$, and incubated on ice for 1 hour. Bacteria were then centrifuged (16,000 x *g*, 1 min, 4 °C), resuspended in 200 µL of 100 mM $CaCl_2$ containing 20 % (v/v) glycerol, and split into 100 µL aliquots. Aliquots were used immediately or stored at -80 °C.

2.5.2. Heat-shock transformation of chemically competent *E. coli*

Chemically competent *E. coli* aliquots were gently thawed on ice if necessary. Either 2 µL of purified plasmid or the entire volume of a ligation reaction was added to the

bacteria and incubated on ice for at least 20 min. Bacteria were then heat-shocked at 42 °C for 30 sec and then incubated on ice for 5 min. 900 µL of SOC recovery medium [20 g L^{-1} tryptone (BD), 5 g L^{-1} yeast extract (BD), 8.6 mM NaCl, 2.5 mM KCl, 10 mM $MgSO_4$, 0.2 5 (v/v) glucose] was added and mixture incubated at 37 °C for 30 – 60 min before plating.

2.5.3. Preparation of electrocompetent *S. flexneri*

Mid-exponential phase (OD_{600} ~ 0.5) *S. flexneri* grown in 10 mL LB broths were harvested by centrifugation (2,200 x *g*, 10 min, 4 °C), washed once with 10 mL of ice-cold MQ H_2O and then again with 5 mL of ice-cold MQ H_2O, resuspended in 200 µL of 20 % (v/v) glycerol, and split into 100 µL aliquots. Electrocompetent *S. flexneri* were always used fresh for electro-pulse transformation.

2.5.4. Electro-pulse transformation of *S. flexneri*

2 µL of purified DNA was added to electrocompetent *S. flexneri* aliquots and mixture incubated on ice for 5 minutes. Bacteria were then transferred to 0.2 cm gap electroporation cuvettes (BioRad) and electro-pulsed using a BioRad Gene Pulser using pre-set *E. coli* conditions. 900 µL of SOC recovery medium was immediately added to shocked cells before incubation at 37 °C for 1 h and then plating.

2.6. Protein Techniques

2.6.1. Total bacterial protein samples

A total of 2.5 x 10^8 bacteria from a mid-exponential phase (OD_{600} = 0.5) culture were harvested (16,000 x *g*, 1 min, 4 °C) and resuspended in 50 µL of sample buffer [2 % (w/v) SDS, 10 % (v/v) glycerol, 5 % (v/v) β-mercaptoethanol, 0.02 % (w/v) bromophenol blue, 62.5 mM Tris pH 7]. Samples were heated to 100 °C for 15 min and either stored at -20 °C or used immediately for SDS-PAGE.

2.6.2. Precipitation of culture supernatant protein

200 mL volume bacterial cultures of *S. flexneri* were grown to an OD_{600} of 0.5 Cultures were centrifuged (10,000 x *g*, 15 min, 4 °C) and 150 mL of supernatant was harvested. Ice-cold trichloroacetic acid [TCA; 100% (w/v) solution] was added to a final

concentration of 6.5% (v/v) and incubated on ice for 1 - 2 h. Precipitates were pelleted by centrifugation (10,000 x *g*, 30 min, 4 °C), washed with 1 mL of ice-cold acetone, re-pelleted by centrifugation (16,000 x *g*, 30 min, 4 °C) before the supernatant was removed and the pellet air dried. Dried pellets were resuspended in 150 µL of sample buffer and pH corrected with 1 M TrisHCl (pH 7.0) if required. Samples were heated to 100 °C for 15 min and either stored at -20 °C or used immediately for SDS-PAGE separation.

2.6.3. Bacterial sub-cellular protein fractionation and differential detergent treatment

Isolation of protein from specific cellular compartments was conducted over ice at all times, unless otherwise stated. Homogenisations were conducted using dounce-homogenisers. 250 mL *S. flexneri* LB broths were grown to an OD_{600} of 0.5, and 5 x 10^{10} bacteria were harvested (3,000 x *g*, 20 min, 4 °C). Culture supernatants were used in TCA precipitations where required. The pellet was resuspended in 15 mL of 10 mM HEPES (pH 7.5) and lysed by either sonication or French press. Debris was removed by centrifugation (10,000 x *g*, 10 min, 4 °C) and the supernatant ultracentrifuged (149,000 x *g* average, 1 h, 4 °C). The cytoplasmic/periplasmic protein containing supernatant was stored where required. The whole membrane pellet was washed by homogenisation in 15 mL of 10 mM HEPES (pH 7.5) and re- ultracentrifuged (149,000 x *g* average, 1 h, 4 °C). The whole membrane pellet was then homogenised into 500 µL of 10 mM HEPES (pH 7.5) supplemented with either 1 % (w/v) sodium N-lauroyl sarcosinate (Sarkosyl; Sigma) or 2 % (v/v) Triton X-100 (Sigma), and incubated for 30 min at 37 °C with continuous inversion. Treated membranes were ultracentrifuged again (106,000 x *g* average, 1 h, 4 °C), the soluble fraction containing inner membrane proteins was stored at -20 °C, while the outer membrane protein pellet was washed by homogenisation in 1 mL of 10 mM HEPES (pH 7.5) and re- ultracentrifuged (106,000 x *g* average, 1 h, 4 °C). The outer membrane pellet was then homogenised into 500 µL of 10 mM HEPES (pH 7.5) and stored at -20 °C.

When the focus was solely on obtaining outer membrane protein, the whole membrane pellet stage was instead treated with 15 mL of 10 mM HEPES (pH 7.5) containing sodium N-lauroylsarcosinate (Sarkosyl) or Triton X-100 (at concentrations as above) and incubated for 30 minutes at 37 °C with continuous inversion. The outer membrane protein pellet was collected by ultracentrifugation (149000 x *g* average, 1 hour, 4 °C), washed by homogenisation in 15 mL of 10 mM HEPES (pH 7.5), re-ultracentrifuged (149000 x *g* average, 1 hour, 4 °C), and homogenised into 250 – 500 µL of 10 mM HEPES (pH 7.5) and stored at -20 °C.

Before analysis of fractions using SDS-PAGE, samples of fractions were diluted in 10 x sample buffer and heated to 100 °C for 15 min. The success and purity of fractions was always assessed by both Coomassie blue staining and western immunoblotting against proteins of known sub-cellular localisation; DnaK (cytoplasm), Wzz (inner membrane), SurA (periplasm), BamA (outer membrane).

2.6.4. Small scale *in vitro* peripheral outer membrane protein disassociation assay

100 µL aliquots of previously isolated outer membrane protein homogenised in 10 mM HEPES (pH 7.5) were collected by ultracentrifugation (106,000 x *g* average, 1 h, 4 °C). Pellets were treated for 30 min at 25 °C by homogenisation into 100 µL volumes of a range solutions: 10 mM HEPES (pH 7.5) control buffer, 8 M urea in 10 mM HEPES (pH 7.5), 100 mM Na_2CO_3 (pH 11.3), 1 M NaCl in 10 mM HEPES (pH 7.5), or 2 M NaI in 10 mM HEPES (pH 7.5). Each treatment was re-ultracentrifuged (106,000 x *g* average, 1 h, 4 °C), the soluble fraction (peripherally associated proteins) was stored, and the insoluble fraction (integral proteins) was homogenised into 100 µL of 10 mM HEPES (pH 7.5). Before analysis of fractions using SDS-PAGE, samples were diluted in 10 x sample buffer and heated to 100 °C for 15 min.

2.6.5. *In vitro* outer membrane proteolysis assay

This assay was conducted over ice at all times. Six 20 µL aliquots of previously isolated outer membrane protein homogenised in 10 mM HEPES (pH 7.5) were treated with a final concentration of 10 µg mL^{-1} Proteinase K (PK; Sigma). At time points 0, 5, 10, 15, 30, and 60 min. the reaction of one of the aliquots was stopped by the addition of a final concentration of 12 % (v/v) TCA and heating to 70 °C for 5 min. All samples were cooled on ice (5 min), precipitated protein collected by centrifugation (16,000 x *g*, 20 min, 4 °C), washed in 250 µL of acetone, pelleted (16,000 x *g*, 20 min, 4 °C), and air dried. Pellets were finally resuspended in 100 µL of sample buffer and heated to 100 °C for 15 min. Samples were pH corrected with 1 M TrisHCl (pH 7) if required.

2.6.6. *In vivo* limited protease shaving

A total of 10^9 *S. flexneri* bacteria were collected from mid- to late-exponential phase (OD_{600} = 0.5 – 1.0) LB cultures by centrifugation (16,000 x *g*, 1 min, 4 °C). Bacterial

pellets were washed twice with 1 mL of phosphate buffered saline (PBS) before a final resuspension in 1 mL of PBS and cooling on ice.

This assay was conducted on ice at all times. A 100 µL aliquot was precipitated with 12 % (v/v) final concentration of TCA (0 minutes; untreated). PK (Sigma) was added to the remaining 900 µL of bacteria for a final concentration of 1 µg mL^{-1} and 100 µL aliquots were taken at 1, 5, 30, and 60 min time points, treated with 4 mM final concentration of phenylmethanesulfonyl fluoride (PMSF; 200 mM stock in ethanol) to stop proteolysis, and immediately precipitated with 12 % (v/v) final concentration of TCA. All precipitated samples were centrifuged (16,000 x *g*, 20 min, 4 °C), washed with 250 µL of acetone, re-pelleted (16,000 x *g*, 20 min, 4 °C), air dried, resuspended with 25 µL of loading buffer, and heated to 100 °C for 15 min. Samples were pH corrected with 1 M TrisHCl (pH 7) if required.

2.6.7. *In vivo S. flexneri* pulse-chase proteolysis assay

The glucose repressible, arabinose inducible P_{BAD} promoter expression switch of the vector pBAD30 (Guzman *et al*, 1995) was utilized for controlled expression essentially as previously described (Leyton *et al*, 2014). *S. flexneri* harbouring pBADIcsA and derivatives were sub-cultured to mid-exponential phase (OD_{600} = 0.5) in 100 mL of LB supplemented with 0.2 % (v/v) glucose. Bacteria were collected by centrifugation (4,000 x *g*, 4 min, 4 °C), washed with 50 mL of LB, re-pelleted (4,000 x *g*, 4 min, 4 °C), resuspended in LB containing 0.2 % (v/v) arabinose, and incubated (5 min, 25 °C constant temperature block) for IcsA production (pulse phase). Bacteria were immediately collected (4,000 x *g*, 4 min, 4 °C), resuspended in 30 mL of LB containing 0.2 % (v/v) glucose, and placed on a 25 °C constant temperature block for 60 min (chase phase). During the chase phase, four 1 mL aliquots were taken at each of the times; 0 (resuspension), 1, 5, 10, 20, 40, and 60 min. The first aliquot was precipitated by the addition of 12 % (v/v) final concentration of TCA and placed on ice. The second aliquot was treated with a final concentration of 10 µg mL^{-1} PK (Sigma) on ice for 10 min. Proteolysis was then stopped by addition of 4 mM final concentration of PMSF before TCA precipitation (12 % (v/v) final concentration). Bacteria from the third aliquot were collected by centrifugation (16,000 x *g*, 1 min, 4 °C), treated with 20 µL of chloroform to selectively permeabilize the outer membrane (Ames *et al*, 1984, Wagner *et al*, 2009), and then PK treated, stopped, and precipitated as described for the second aliquot. All precipitated samples were pelleted (16,000 x *g*, 20 min, 4 °C), washed with 250 µL of acetone, re-pelleted (16,000 x *g*, 20 min, 4 °C), air dried, and resuspended with 50 µL of loading buffer per 1 OD_{600}

unit (measured using the fourth aliquot). Samples were heated to 100 °C for 10 min, cooled, and immediately used in SDS-PAGE and western immunoblotting analysis.

2.6.8. *S. flexneri in situ* cross-linking and membrane protein fractionation

A total of 5 x 10^{10} *S. flexneri* bacteria were collected from mid- to late-exponential phase (OD_{600} = 0.5 – 1.0) LB cultures by centrifugation (3,000 x *g*, 20 min, 4 °C). The bacterial pellet was washed with 50 mL of DSP crosslinking buffer [120 mM NaCl, 20 mM Na_2HPO_4/NaH_2PO_4 pH 7.2 buffer], re-pelleted (3,000 x *g*, 10 min, 4 °C), and resuspended in 20 ml of DSP crosslinking buffer. A final concentration of 0.2 mM dithiobis(succinimidyl propionate) (DSP; Pierce) crosslinking agent was added to the bacteria and incubated at 37 °C for 30 min. Excess DSP was quenched by the addition of a final concentration of 20 mM TrisHCl (pH 7.5). Bacteria were pelleted (3,000 x *g*, 10 min, 4 °C) and resuspended in 15 mL of 10 mM HEPES (pH 7.5). Whole membrane, inner membrane, and outer membrane protein samples were then isolated from the cross-linked bacteria using the method described in part 2.6.3.

For analysis of fractions using SDS-PAGE, samples of fractions were diluted 1:1 in either 2 x sample buffer or 2 x β-mercaptoethanol negative sample buffer and heated to 100 °C for 15 min. Samples containing β-mercaptoethanol will have crosslinks reversed and β-mercaptoethanol negative samples will have crosslinks retained.

2.6.9. Electrophoretic separation of protein

SDS-PAGE was conducted using BioRad self-cast Mini-Protean system III and a Tris-Glycine buffer system. Running buffer (pH 8.3) consisted of 25 mM TrisHCl, 200 mM glycine, and 0.1 % (w/v) SDS. Gels were generally electrophoresed between 100 – 200 V for 1 – 2 h.

2.6.10. Coomassie blue staining

SDS-PAGE separated proteins were stained by incubating the gel in Coomassie blue stain solution [0.01 % (w/v) Coomassie blue G250, 10 % (v/v) acetic acid, 50 % (v/v) methanol] at room temperature with shaking overnight. Gels were destained with repeated washes of Destain solution [10 % (v/v) acetic acid, 50 % (v/v) methanol].

2.6.11. Western immunoblot and detection

SDS-PAGE separated proteins were transferred to nitrocellulose membranes (NitroBind, pure nitrocellulose, 0.45 µm, GE Water & Process Technologies) using BioRad Trans-Blot Electrophoretic Transfer Cells. Transfer was conducted for 1 h at 200 – 400 mA in transfer buffer [25 mM TrisHCl, 200 mM glycine, 5 % (v/v) methanol]. Blots were blocked with 5 % (w/v) skim milk in TTBS buffer [16 mM TrisHCl, 120 mM NaCl, 0.05 % (v/v) Tween-20 (Sigma)] for 1 h before overnight incubation with primary antibody diluted in 5 % (w/v) skim milk in TTBS. Blots were then washed three times with TTBS and incubated with HRP-conjugated secondary antibody diluted in 5 % (w/v) skim milk in TTBS for 2 h. Blots were then washed three times with TTBS and once with TBS [TTBS without Tween-20]. Blots were incubated with Chemiluminescent substrate (Sigma) and then developed by either X-ray film exposure using a Curix 60 automatic X-ray film processor (Agfa) or digitally using a BioRad GelDoc XR+.

2.7. Tissue Culture Techniques

2.7.1. Tissue culture and maintenance

The cell line used in this thesis was HeLa cells (Human cervical epithelial adenocarcinoma cells ATCC CCL-2). HeLa cells were grown in Modified Eagle's medium (MEM, Life Technologies) supplemented with 10 % (v/v) foetal calf serum (FCS), 4 mM L-glutamine, 100 U mL^{-1} penicillin, and 100 U mL^{-1} streptomycin. Cells were maintained at 37 °C in a humidified incubator with a constant 5 % CO_2.

2.7.2. Plaque assays

Plaque formation assays were conducted as described previously (Oaks *et al*, 1985, Teh & Morona, 2013). HeLa cells were grown overnight to a confluent monolayer in a 6-well tray. Prior to infection with *S. flexneri*, monolayers were washed twice with Dulbecco's PBS [D-PBS; PBS containing 0.1 % (w/v) $CaCl_2$ and 0.1 % (w/v) $MgCl_2$] and once in Dulbecco's modified Eagle medium (DMEM, Life Technologies). A total of 10^8 mid-exponential phase (OD_{600} = 0.5) bacteria were harvested by centrifugation (16,000 × *g*, 2 min, 20°C), resuspended in 1 mL of DMEM, and serial diluted to 10^5 bacteria mL^{-1} in DMEM before 200 µL was added to each well. Monolayers were incubated for 2 h at 37 °C with 5 % CO_2 with gentle rocking every 15 min. Well media was then aspirated and 3 mL of DMEM

supplemented with 5 % (v/v) FCS, 20 μg mL^{-1} gentamycin, and 0.5 % (w/v) agarose (Seakem ME) was added to each well (overlay one) before further incubation at 37 °C with 5 % CO_2 to set. The second overlay containing DMEM, 5 % (v/v) FCS, 20 μg mL^{-1} gentamycin, 0.5 % (w/v) agarose, and 0.1 % (w/v) Neutral Red (MP Biomedicals) was added at 48 hours post-infection before further incubation at 37 °C with 5 % CO_2. Wells were photographed at 70 h post infection from a standardized height while illuminated with a light box. Plaque diameters were quantitated using MetaMorph image analysis software (Molecular Devices, Version 7.7.3.0).

2.8. Microscopy Techniques

2.8.1. Immunofluorescence staining of infected tissue culture samples

Infection of semi-confluent HeLa cell monolayers with *S. flexneri* was conducted as described previously (Teh *et al*, 2012a). HeLa cells were grown on sterile round glass coverslips (at the bottom of 24-well trays). Mid-exponential phase (OD_{600} = 0.5) bacteria were harvested by centrifugation (16,000 × *g*, 2 min, 20 °C) and diluted to 3 × 10^8 bacteria mL^{-1} in D-PBS. HeLa cells were washed with MEM containing 10% (v/v) FCS, and then 80 μL of bacteria was added before the tray was centrifuged (500 × *g*, 5 min, 20 °C) to assist adhesion and invasion. After incubation at 37 °C with 5 % CO_2 for 1 h to allow invasion, wells were washed three times with D-PBS, and incubated a further 90 min with 500 μL of MEM supplemented with 10 % (v/v) FCS and 40 μg ml^{-1} gentamycin to allow time for actin-based motility initiation. Cells were then washed three times with D-PBS, treated for 15 min with 3.7 % (v/v) formaldehyde solution (Sigma) in PBS, and washed twice with PBS. Cells were incubated with 50 mM NH_4Cl in D-PBS for 10 min, washed with PBS, permeablized with 0.1 % (v/v) Triton X-100 in PBS for 5 min, and washed once with PBS.

In general, infected HeLa cells were most commonly stained for N-WASP, F-actin, and DNA. Cells were blocked with 10 % (v/v) FCS in PBS for 20 min, before incubation with rabbit anti-N-WASP antibodies diluted at 1:100 in PBS containing 10 % (v/v) FCS for 30 minutes at 37 °C. Cells were washed three times with PBS, and then incubated for 1 h at 37 °C (with protection from light) with donkey anti-rabbit AlexaFluor 594 antibody and AlexaFluor 488 phalloidin (Invitrogen) diluted to 1:100 and 1:200 respectively in PBS containing 10 % (v/v) FCS. These were then washed three times with PBS. DNA was stained

using 10 μg ml^{-1} DAPI for 1 min, washed three times with PBS again. All samples were mounted on slides for microscopy (see 2.8.3.).

2.8.2. Immunofluorescence staining of bacteria

All solutions to be used were filtered through a 0.22 μm nitrocellulose filter. A total of 10^8 mid-exponential phase bacteria (OD_{600} = 0.5) were harvested by centrifugation (16,000 x *g*, 2 min, 20 °C), fixed for 20 min in 3.7 % (v/v) formaldehyde solution in PBS at room temperature, washed twice in PBS, and resuspended in 100 μL of PBS. Sterile round coverslips (at the bottom of a 24-well tray) were pre-treated with 10 % (v/v) poly-L-lysine solution (Sigma). Fixed bacteria were spotted (5 μL) onto the coverslips before centrifugation of the tray to assist adherence (775 x *g*, 5 min, 20 °C).

Bacteria were then incubated with dilutions of the desired antibodies – this was commonly anti-IcsA. Bacteria were incubated for 2 h at 20 °C with anti-IcsA antibodies diluted 1:100 in PBS containing 10 % (v/v) FCS. Bacteria were then washed three times with PBS, incubated (with light protection) for 30 min at 37 °C with anti-rabbit AlexaFluor 488 diluted 1:100 in PBS containing 10 % (v/v) FCS. Bacteria were washed three more times with PBS before mounting on microscope slides for microscopy (see part 2.8.3.).

2.8.3. Mounting and Microscope

To mount sample coverslips, mounting media was prepared fresh each time containing 20 % (w/v) Mowiol 4–88 (Calbiochem) and 4 mg mL^{-1} *p*-phenylenediamine. A volume of 5 μL of mounting media was spotted onto microscope slides before the use of hooked forceps to extract sample coverslips from the trays and inverting them onto the mounting media spot. Coverslip edges were then sealed using clear nail polish sparingly.

All microscopy images were captured using an Olympus IX-70 microscope with a phase contrast 100 × oil immersion objective and with the varied use of a 1.5 × enlarger. For fluorescence imaging, an X-Cite 120Q lamp was used set at high intensity. An Omega XF67-1 excitation filter set and a narrow band excitation filter (Omega X67) was commonly used in the Pinkel configuration. For quantum dot (Qdot) microscopy, a specialist set up was used consisting of FF510-Di01-25x36 dichroics (Semrock), and an FF01-435/40-25 (Semrock) excitation filter used with FF01-525/15-25 and FF01-625/15-25 (Semrock) emission filters. Fluorescence images were captured using consistent and precise exposure times such that intensity quantitation could be conducted at a later time point if required. Fluorescence images

of blank fields were also taken using the same exposure times to enable background fluorescence correction. Images were fine focused with the aid of an MFC-2000 digital z-axis controller (Applied Scientific Instruments). The microscope was controlled using MetaMorph image analysis software (Molecular Devices, Version 7.7.3.0).

2.8.4. Image analysis, fluorescence quantification, and directional scanning

All bacterial IcsA and N-WASP fluorescence images were acquired with predetermined exposures; 100 and 500 msec respectively. MetaMorph image analysis software was used to subtract the background fluorescence from individual micrographs before further analysis.

For total bacterial fluorescence, bacteria were first selected (at least 50 per experiment) using phase contrast micrographs and MetaMorph region tools. Region selections were ported to corresponding fluorescence micrographs and average bacterial fluorescence intensities automatically tallied. From this, mean bacterial fluorescence intensity was calculated for a given bacterial population.

To quantitate fluorescence intensities with respect to the longitudinal measurement of the bacterium, bacteria were first selected (at least 50 per experiment) using phase contrast micrographs and MetaMorph line-scan tools. The software was instructed to scan from pole-to-pole and average fluorescence intensities across the perpendicular axis at a pixel width resolution. From this, mean bacterial fluorescence intensity profiles were calculated for a given bacterial population.

2.9. Bioinformatic Techniques

2.9.1. Databases, metadata, data-mining, and analysis

To identify conserved sequence features of proteins the Pfam (pfam.xfam.org) search engine and InterProScan 5 (ebi.ac.uk/interpro) was used interchangeably (Hunter *et al*, 2012, Finn *et al*, 2014a, Jones *et al*, 2014). For investigating AT common sequence features the UniProt knowledgebase (uniprot.org) was used due to its ease of use and its extensive architecture annotation of protein entries by InterPro, Pfam, and other databases (Jain *et al*, 2009, Magrane & Consortium, 2011). Custom protein entry lists were easily retrieved from UniProt using basic AND/NOT/OR logic for various criteria; species, domain identifiers,

annotation statistics. To remove redundancy, each list was clustered into UniRef100 sequence clusters (Suzek *et al*, 2007) which groups identical sequences and represents each group with a representative entry. Generated lists of entry identifiers were then used to extract required metadata for each entry including amino acid sequence, sequence length, organism, conserved domains, conserved domain positions, etc. Differences in domain combinations were identified by further grouping lists based on domain combination frequencies and inclusions or exclusions of certain domains to see if this alters the combination frequencies. Statistical changes in frequencies were analysed using Graphpad Prism 6 software. Entry lists were also input to ExPASy-Compute pI/Mw tool (web.expasy.org/compute_pi) to calculate theoretical monoisotopic isoelectric points (pI) and then grouped analyses were again conducted using Graphpad Prism 6.

2.9.2. *In silico* protein tertiary structure prediction

To model protein tertiary structure, the I-TASSER web server (zhanglab.ccmb.med.umich.edu/I-TASSER) was used which uses threading and *ab initio* modelling of input amino acid sequences (Roy *et al*, 2010, Xu *et al*, 2011). Confidence in the model was determined by the C-score which has a possible range from -5 to 2 (where higher values signifies higher confidence in the model) and TM-score where a score of more than 0.5 indicates correct topology and a score of less than 0.17 indicates random similarity. Models were visualised using the PyMOL Molecular Graphics System (Version 1.3).

2.9.3. *In silico* protein-protein alignment via structure

To identify conservation in structure between proteins, the TM-Align web server (zhanglab.ccmb.med.umich.edu/TM-align) was used which is able to align two input structures (Zhang & Skolnick, 2005). Protein folds were shared when the TM-score was greater than 0.5.

Chapter Three

ARTICLE 1:

LPS unmasking of *Shigella flexneri* reveals preferential localisation of tagged outer membrane protease IcsP to septa and new poles.

Elizabeth Ngoc Hoa Tran*, **Matthew Thomas Doyle*** and Renato Morona

Chapter 3: *Research Article 1:* LPS unmasking of *Shigella flexneri* reveals preferential localisation of tagged outer membrane protease IcsP to septa and new poles.

3.1. Statement of Authorship

Title of Paper	LPS unmasking of *Shigella flexneri* reveals preferential localisation of tagged outer membrane protease IcsP to septa and new poles.
Status	Published
Citation	Tran EN*, Doyle MT* & Morona R (2013) LPS unmasking of *Shigella flexneri* reveals preferential localisation of tagged outer membrane protease IcsP to septa and new poles. *PloS one* 8: e70508. doi: 10.1371/journal.pone.0070508. * Equal Contribution

Author Contributions: By signing the Statement of Authorship, each author certifies that: (i) the candidate's stated contribution to the publication is accurate, (ii) permission is granted for the candidate in include the publication in the thesis; and (iii) the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

Author	Elizabeth Ngoc Hoa Tran		
Contribution	Construction of strains and plasmids, conducted all experiments pertaining to Figures 3.1, 3.2, 3.3, and 3.4 (Western blot, OMP purification, LPS analysis) and constructed these figures, data analysis, conception of model, and writing of manuscript.		
Signature		Date	28/07/15

Author	Matthew Thomas Doyle		
Contribution	Designed, optimised, and conducted all quantum dot microscopy experiments and quantitation pertaining to Figures 3.5, 3.6, 3.7, 3.8, 3.9, 3.10S, 3.11S, 3.12S, and 3.13S and constructed these figures, idea of tunicamycin treatment, data analysis, conception of model, and writing of manuscript.		
Signature		Date	28/07/15

Author	Renato Morona		
Contribution	Supervised development of work, provision of laboratory and materials, design of quantum dot experiment, manuscript evaluation and editing, acting corresponding author.		
Signature		Date	28/07/15

3.2. Purpose of the Article

As previously described in the Introduction, the polarity of IcsA is maintained by multiple factors. One of these factors is the IcsA-specific protease IcsP which is an outer membrane localised Omptin-type barrel protein that cleaves the IcsA passenger domain from the bacterial surface. This action appears to be required for the refined polar localisation of IcsA. However, whether IcsP is evenly localised over the bacterial surface, or is asymmetrically localised, had not been established. Therefore, this article investigated the surface localisation of IcsP with respect to that of IcsA providing a greater understanding of how IcsA is maintained specifically at the bacterial pole.

3.2.1. Thesis Research Aims Addressed:

- **To investigate the maintenance of IcsA polar localisation at the bacterial surface.**

  - To discover the localisation of IcsP with respect to the localisation of IcsA over the bacterial outer membrane.

3.3. Article Abstract

The *Shigella flexneri* outer membrane (OM) protease IcsP (SopA) is a member of the enterobacterial Omptin family of proteases which cleaves the polarly localised OM protein IcsA that is essential for *Shigella* virulence. Unlike IcsA however, the specific localisation of IcsP on the cell surface is unknown. To determine the distribution of IcsP, a haemagglutinin (HA) epitope was inserted into the non-essential IcsP OM loop 5 using Splicing by Overlap Extension (SOE) PCR, and $IcsP^{HA}$ was characterised. Quantum Dot (QD) immunofluorescence (IF) surface labelling of $IcsP^{HA}$ was then undertaken. Quantitative fluorescence analysis of *S. flexneri* 2a 2457T treated with and without tunicaymcin to deplete lipopolysaccharide (LPS) O antigen (Oag) showed that $IcsP^{HA}$ was asymmetrically distributed on the surface of septating and non-septating cells, and that this distribution was masked by LPS Oag in untreated cells. Double QD IF labelling of $IcsP^{HA}$ and IcsA showed that $IcsP^{HA}$ preferentially localised to the new pole of non-septating cells and to the septum of septating cells. The localisation of $IcsP^{HA}$ in a rough LPS *S. flexneri* 2457T strain (with no Oag) was also investigated and a similar distribution of $IcsP^{HA}$ was observed. Complementation of the rough LPS strain with *rmlD* resulted in restored LPS Oag chain expression and loss of $IcsP^{HA}$ detection, providing further support for LPS Oag masking of surface proteins. Our data presents for the first time the distribution for the Omptin OM protease IcsP, relative to IcsA, and the effect of LPS Oag masking on its detection.

3.4. Article Introduction

Shigella flexneri is an intracellular pathogen which causes bacillary dysentery, a disease characterised by the presence of severe mucoid bloody diarrhoea and by invasion of the gut epithelium (LaBrec *et al*, 1964, Suzuki & Sasakawa, 2001). IcsA (VirG) is a 120 kDa outer membrane (OM) protein localised at the cell pole (Robbins *et al*, 2001). It mediates intracellular cytoplasmic movement of *S. flexneri* in epithelial cells, and cell-to-cell spread, by the assembly of an F-actin comet-tail at one pole of the bacterium (Bernardini *et al*, 1989, Lett *et al*, 1989, Goldberg *et al*, 1993). This type of movement is described as actin-based motility (ABM). IcsA is secreted primarily at the 'old pole' of *Shigellae* (Jain *et al*, 2006) which is opposite the 'new pole' (the pole derived from the site of septation of the parent cell (Dworkin, 2009). The 36.9 kDa IcsP (SopA) OM protease of *S. flexneri* slowly cleaves IcsA at the Arg_{758} – Arg_{759} bond position (Fukuda *et al*, 1995) resulting in the release of a 95 kDa amino-terminal IcsA fragment that can be detected in culture supernatants (Goldberg *et al*, 1993, Goldberg & Theriot, 1995). Analysis of *icsP* / *sopA* mutants has shown that IcsA is detected across the entire surface of these bacteria with polar reinforcement (Egile *et al*, 1997, Shere *et al*, 1997). Over-expression of IcsP results in the complete removal of IcsA from the cell surface (Steinhauer *et al*, 1999).

IcsP belongs to the Omptin family of proteases which consists of 6 members; OmpT and OmpP of *Escherichia coli*, Pla of *Yersinia pestis*, PgtE of *Salmonella enterica*, Pla endopeptidase A of *Erwinia pyrifoliae*, and IcsP of *Shigella flexneri*. Immunogold labelling of overexpressed OmpP has shown that OmpP is symmetrically distributed over the cell surface (Kaufmann *et al*, 1994). However to date, no studies have attempted to describe the surface localisation of Omptins expressed at native levels. While it has been suggested that IcsP may also be located uniformly across the cell surface (Steinhauer *et al*, 1999), its specific distribution is currently unknown. In contrast to many inner membrane proteins, such as FtsZ (Lutkenhaus & Addinall, 1997) and MreB (Tamaki *et al*, 1980) involved in cell division and cell shape, few OM proteins have had their subcellular distribution determined. An exception to this is the *E. coli* OM protein LamB which has been characterised to exist as two populations: one that diffuses in a helical pattern, and one that is relatively immobile (Gibbs *et al*, 2004, Chatterjee & Rothenberg, 2012). The *E. coli* Iss and Bor proteins have been detected on the cell surface with no distinct pattern (Lynne *et al*, 2007). A number of non-specific *E. coli* OM proteins were suggested to be organised in stable helical swaths (Ghosh & Young, 2005), and data by Shiomi *et al.* (2006) suggested that the general protein translocation Sec machinery

itself may also be arranged in a helical array. Whether IcsP possesses a distribution similar to these OM proteins, or has an asymmetric distribution like IcsA, is the subject of this study.

In addition to the above, mutations affecting lipopolysaccharide (LPS) have also been shown to affect the observed distribution of OM proteins (Van der Ley *et al*, 1986a, Van der Ley *et al*, 1986b, Voorhis *et al*, 1991, Morona & Van Den Bosch, 2003b). LPS is composed of three distinct regions: lipid A, core sugars and O antigen (Oag) polysaccharide chains. Strains with LPS containing all 3 regions intact are known as smooth LPS strains. *Shigella* mutants lacking Oag are known as rough LPS strains. Such strains have been shown to have high levels of circumferentially distributed IcsA on the cell surface (at both cell poles and on lateral regions) (Sandlin *et al*, 1995, Van Den Bosch *et al*, 1997), compared to the polar localisation of IcsA seen in smooth LPS strains. Treatment of Y serotype derivatives of smooth LPS *S. flexneri* with bacteriophage Sf6 tailspike protein (TSP) endorhamnosidase results in the hydrolysis of Oag chains and an increased detection of circumferential IcsA on the cell surface by indirect immunofluorescence (IF) staining (Morona & Van Den Bosch, 2003b). This suggests that the presence of LPS Oag masks the observed distribution of IcsA on the cell surface and supports the idea that LPS Oag structure may block antibody accessibility to the detection of surface proteins (Van der Ley *et al*, 1986a, Van der Ley *et al*, 1986b). The effect of LPS Oag structure on the detection and distribution of IcsP has not been investigated.

In this study, we investigated the distribution of IcsP by cell surface quantum dot (QD) IF labelling of functional, HA-tagged IcsP ($IcsP^{HA}$) in *S. flexneri* 2457T and establish that LPS Oag masks detection of $IcsP^{HA}$ on the cell surface by using tunicamycin to inhibit Oag synthesis. Additional IF labelling with anti-IcsA antibodies to mark the location of the old pole suggested that IcsP is preferentially localised to the new pole of non-septating cells and to the septa of septating cells. We also investigated the distribution of IcsP in a rough LPS 2457T strain to provide further support for the LPS Oag masking hypothesis. Overall, our data presents for the first time the cell surface distribution of the Omptin OM protease IcsP and the effect of LPS Oag masking on its detection. This distribution has implications for IcsA polarity determination, and a model is described to explain IcsP's contribution to IcsA polarity in *S. flexneri*.

3.5. Article Methods

3.5.1. Ethics Statement.

The anti-IcsP and anti-IcsA antibodies were produced under the National Health and Medical Research Council (NHMRC) Australian Code of Practice for the Care and Use of Animals for Scientific Purposes and were approved by the University of Adelaide Animal Ethics Committee.

3.5.2. Bacterial strains, plasmids and media.

Bacterial strains and plasmids used in this study are listed in Table 3.1. Bacteria were routinely grown at 37 ºC in Luria-Bertani (LB) broth with aeration or on Congo red agar (Van Den Bosch *et al*, 1997). Antibiotics used were as follows: 50 µg ampicillin (Amp) ml^{-1}; 25 µg chloramphenicol (Cml) ml^{-1}; 50 µg kanamycin (Kan) ml^{-1}.

3.5.3. DNA methods.

E. coli K-12 DH5α was used for all cloning experiments. DNA manipulation, PCR, transformation and electroporation was performed as previously described (Morona *et al*, 1995, Purins *et al*, 2008).

3.5.4. Antibodies.

Anti-HA monoclonal antibody (#H3663) was purchased from Sigma. Rabbit anti-IcsP and anti-IcsA antibodies were prepared as described previously (Van Den Bosch *et al*, 1997, Teh *et al*, 2012b). The antibodies were produced under the National Health and Medical Research Council (NHMRC) Australian Code of Practice for the Care and Use of Animals for Scientific Purposes and were approved by the University of Adelaide Animal Ethics Committee.

3.5.5. Insertion of HA epitope into IcsP.

The sequence encoding the HA epitope (YPYDVPDYA) was inserted into the putatively non-essential IcsP OM loop 5 (based on sequence alignments with *E. coli* OmpT) using SOE PCR (Horton *et al*, 1990, Horton, 1993). In the first part of this two-step PCR technique, upstream and downstream amplicons were amplified from *S. flexneri* 2457T

genomic DNA using HA encoding primers (ET18/ET19) and *icsP* specific primers (ET3/ET10) (Table 3.2). The two amplicons from this primary PCR were then mixed and used as a DNA template for the second round of PCR with primers ET3/ET10. In this second reaction, the HA encoding regions of the primary PCR amplicons overlap and prime one another to give the final *icsP* PCR product with an inserted HA epitope. The *icsP*^{HA} fragment was then cloned into pGEMT-Easy and primers ET22/ET25 (Table 3.2) were used to amplify an *icsP*^{HA} product with *Kpn*I and *Hind*III restriction enzyme sites from this construct. The resultant *Kpn*I-*Hind*III fragment was digested and sub-cloned into likewise digested pBAD30 to give pBAD30::*icsP*^{HA}, also referred to as pIcsP^{HA} in text (Table 3.1). Primers ET22/ET25 were also used to amplify the *icsP* gene from 2457T genomic DNA, and cloned into pBAD30 to give pBAD30::*icsP*, referred to as pIcsP in text (Table 3.1). DNA sequencing was used to confirm that no mutation had been introduced by PCR into the sequence, and the presence of the in-frame HA epitope tag sequence.

3.5.6. Construction of *S. flexneri icsP* mutant.

The *S. flexneri* 2457T *icsP* mutant strain was constructed using allelic exchange mutagenesis (Morona *et al*, 1995) to inactivate the *icsP* gene by insertion of a kanamycin resistance gene (*kan*^{R}). Initially, the *icsP* gene was PCR amplified with primers ET3/ET4 containing *Bam*HI and *Sac*I restriction sites from 2457T genomic DNA. The resultant PCR fragment was digested with *Bam*HI and *Sac*I and sub-cloned into likewise digested pSL1180 (Table 3.1). Further digestion with *Cla*I allowed insertion of the *Acc*I-*Acc*I digested *kan*^{R} gene from pKTUWE (Table 3.1) to give pSL1180-*icsP*::*kan*^{R} (Table 3.1). Following re-digestion with *Bam*HI and *Sac*I, the *icsP*::*kan*^{R} fragment was cloned into pCACTUS (Table 3.1) and transformed into *S. flexneri* 2457T via electroporation. Allelic exchange mutagenesis was performed as previously described (Morona *et al*, 1995). The *icsP*::*kan*^{R} mutation in the virulence plasmid was confirmed by PCR with primers ET3/ET4 (Table 3.2) to give the 2457T *icsP* mutant ETRM22 (Table 3.1).

3.5.7. Construction of *S. flexneri icsP*/*rmlD* double mutant.

The *S. flexneri* 2457T *icsP*/*rmlD* mutant strain was constructed using a modification of the λ Red recombinase system to initially delete the *rmlD* gene (Datsenko & Wanner, 2000). Primers ET28/ET29 containing *Nhe*I restriction enzyme sites (Table 3.2) were used to PCR amplify the *kan*^{R} gene from pKD4 (Table 3.1). The amplified product was ligated into pGEMT-Easy and pGEMT-Easy::*kan*^{R} was digested with *Nhe*I. The *Nhe*I-*kan*^{R}-*Nhe*I

fragment was then subcloned into likewise digested pRMA718 (Van Den Bosch *et al*, 1997) to give pRMA718-*rmlD*::*kan*^{R} (Table 3.1). This plasmid was then digested with *Bam*HI and the *rmlD*::*Kan*^{R} fragment was cloned into the *Bam*HI site of pCACTUS. The pCACTUS-*rmlD*::*Kan*^{R} construct was then electroporated into *S. flexneri* 2457T and allelic exchange mutagenesis was induced to give the 2457T *rmlD*::*kan*^{R} mutant ETRM230 (Table 3.1). ETRM230 was transformed with pCP20 at 30 ºC to flip out the FRT flanked *kan*^{R} gene and give the 2457T *rmlD*$^{-}$ mutant ETRM233 (Table 3.1). The *rmlD* mutation was confirmed by LPS analysis. ETRM233 was further electroporated with pCACTUS-*icsP*::*kan*^{R} (Table 3.1) and another round of allelic exchange mutagenesis was performed to give the final 2457T *rmlD*$^{-}$/*icsP*$^{-}$ double mutant ETRM240 (Table 3.1).

3.5.8. Analysis of IcsP/IcsP^{HA} protein production.

For detection of native IcsP, strains were grown at 37 °C in LB broth with aeration for 16 h, subcultured 1/20 into fresh broth and grown for another 3 h to an OD_{600} of ~1. Strains harbouring pIcsP or pIcsP^{HA} were grown in LB broth containing 0.2 % (w/v) glucose for 16 h with aeration, subcultured 1/20 into fresh broth and grown for 1.5 h to an OD_{600} reading of ~0.4. Cultures were then pelleted by centrifugation (2,219 x *g*, 10 min, Sigma 3K15 centrifuge), washed 3 times in LB, and unless otherwise stated, induced with 0.03% (w/v) arabinose for 1 h to an OD_{600} of ~1. Cells (5 x 10^{8}) were then harvested by centrifugation and resuspended in 2X sample buffer (Lugtenberg *et al*, 1975). Protein samples were solubilised at 100 ºC for 5 min, separated on SDS 15 % polyacrylamide gels, and stained with Coomassie R-250, or subjected to Western immunoblotting on nitrocellulose membrane (Medos) with either polyclonal rabbit anti-IcsP antiserum (at 1/250 dilution) or monoclonal mouse anti-HA (at 1/500 dilution). Detection was performed with goat anti-rabbit (or anti-mouse) horseradish-peroxidase-conjugated antibodies (KPL) and chemiluminescence reagent (Sigma). Benchmark prestained molecular weight markers (Invitrogen) were used as molecular size markers.

3.5.9. Sucrose gradient density fractionation.

Fractionation of the cell whole membrane (WM) into cytoplasmic membrane (CM) and OM fractions was performed by sucrose gradient centrifugation according to the method of Osborn and Munson (Osborn & Munson, 1974). In brief, 200 ml cultures were grown and induced with arabinose as described above, harvested by centrifugation (9,800 x *g*, 15 min, 4 ºC, JA14 rotor, Beckman centrifuge J2-21M), washed in 50 mM Tris-HCl (pH 8.0)

and resuspended in 5 ml 10 mM HEPES in 1 mM $MgCl_2$. The bacterial suspension was then passed through a pre-cooled French Pressure cell (SLM Aminco) once and re-centrifuged to remove cell debri. WM pellets were collected by ultracentrifugation (115,000 x *g*, 1 h, 4 ºC, 80Ti rotor, Beckman Coulter Optima L-100 XP ultracentrifuge), solubilised in 0.8 ml 25% (w/w) sucrose in 5 mM EDTA and applied to a 10 ml sucrose gradient of 30 – 50 % (w/w) sucrose in 5mM EDTA. Centrifugation to equilibrium was performed with a Beckman SW40Ti swing out rotor (217,000 x *g*, 20 h, 4 ºC, Beckman Coulter Optima L-100 XP ultracentrifuge) and 0.5 ml fractions collected through the pierced bottom of the tube. 10 µl samples of each fraction were resuspended in 2X sample buffer (Lugtenberg *et al*, 1975) and IcsP protein detected as described above.

3.5.10. Detection of cell associated and soluble IcsA.

Whole cell and supernatant bacterial protein extracts were prepared as described previously (May & Morona, 2008). IcsA protein was detected from 10 µl whole cell protein extracts and 20 µl supernatant protein extracts. Western immunoblotting was performed as described above, but with a polyclonal rabbit IcsA antibody and a goat anti-rabbit HRP conjugate.

3.5.11. LPS analysis.

LPS samples and gels were prepared as described previously (Murray *et al*, 2003, Papadopoulos & Morona, 2010).

3.5.12. LPS depletion-regeneration assay.

Depletion and regeneration of LPS was performed as previously described (Teh *et al*, 2012b, Tran *et al*, 2014) with the exception that 0.03 % (w/v) arabinose induction was included in the final hour of tunicamycin/polymyxin B nonapeptide (PMBN) treatment.

3.5.13. Formaldehyde fixation of bacteria for immunofluorescence (IF) microscopy.

Bacteria were grown and induced as described above. 1 x 10^8 cells of induced bacteria were then harvested by centrifugation, washed once in PBS and resuspended in 100 µl 3.7 % (w/v) formaldehyde (Sigma) in PBS for 20 min at room temperature (RT). Fixed bacteria were then pelleted, washed three times in PBS and resuspended in a final volume of 100 µl PBS.

3.5.14. Quantum dot (QD) IF staining and epi-fluorescence microscopy.

Sterile glass coverslips were placed into wells of a 24-well tray and coated with 10 % (v/v) poly-L-lysine solution (Sigma) in PBS for 1 h at RT. Coating solution was aspirated and 5 μL of formaldehyde fixed bacteria were spotted onto coverslips. The tray was then centrifuged to assist adherence of bacteria (Heraeus Labofuge 400R Centrifuge, 2,000 x *g*, 5 min, 20 °C). Bacteria were blocked for 1 h at RT with 10 % (v/v) foetal calf serum (FCS) diluted in PBS. For labelling, bacteria were incubated for 2 h at RT with mouse anti-HA antibody (Sigma) and rabbit anti-IcsA antiserum diluted 1:50 and 1:100 respectively in PBS containing 10 % (v/v) FCS. Bacteria were then washed 3 times with PBS, and then incubated for 1 h at RT with either QD 525 donkey anti-mouse antibody (Invitrogen) or QD 625 donkey anti-rabbit antibody (Invitrogen) diluted 1:50 and 1:100, respectively, in PBS containing 10 % (v/v) FCS. After a final 3 more washes with PBS, coverslips were mounted on glass microscope slides with Mowiol 4-88 (Calbiochem). All microscopy images were captured using an Olympus IX-70 Microscope, with a phase contrast 100X oil immersion objective and a 1.5X enlarger, which was controlled by MetaMorph (Version 7.7.1.0, Molecular Devices). All $IcsP^{HA}$ (525 nm) and IcsA (625 nm) channel images were acquired with 1 sec and 0.1 sec exposures respectively using an X-Cite 120Q lamp set at high intensity as the excitation source. The excitation filter used was FF01-435/40-25 (Semrock) and the emission filters were FF01-525/15-25 and FF01-625/15-25 (Semrock). Semrock FF510-Di01-25x36 dichroics were used.

3.5.15. Fluorescence quantification of QD labelled surface $IcsP^{HA}$.

Data for intensity profile plots were extracted from images using MetaMorph's Line-scan function which averages intensities across the perpendicular axis of a point-to-point scan. Single scans were conducted from pole-to-pole with width (perpendicular axis) equal to the bacterium (approx. 20 pixels). Cumulative scans of the 525 nm wavelength ($IcsP^{HA}$) for non-septating cells were conducted for 50 bacteria from each independent sample. Each bacterium was scanned from the new pole to old pole where IcsA was used as a marker of the old pole. From the same samples, all septating cells from captured images were scanned starting from the septum to the old pole of one daughter cell chosen at random. An average of 26.8 septating bacteria were scanned for each sample. Intensity data was then exported to MS Excel using MetaMorph's Dynamic Data Exchange and subsequently analysed using GraphPad Prism. Statistical significances were tested by Student's two-tailed t-test.

Table 3.1: Bacterial strains and plasmids.

Strain or plasmid	Relevant characteristics	LPS*	Source/reference
E. coli K-12			
DH5α	endA hsdR supE44 thi-1 recA1 gyrA relA Δ(lacZYA-argF) U169 [ϕ80dlacΔ(lacZ) M15)	R	Gibco-BRL
S. flexneri 2a			
2457T	wild type strain	S	(Van Den Bosch *et al*, 1997)
ETRM22	2457T *icsP*- mutant; Kan^R	S	This study
ETRM230	2457T *rmlD::kan*^{R} mutant; Kan^R	R	This study
ETRM233	2457T *rmlD*- mutant; Kan^R	R	This study
ETRM240	2457T *icsP*-/*rmlD*- mutant; Kan^R	R	This study
ETRM143	ETRM22 (pIcsP)	S	This study
ETRM117	ETRM22 ($pIcsP^{HA}$)	S	This study
ETRM118	ETRM22 (pBAD30)	S	This study
ETRM243	ETRM240 ($pIcsP^{HA}$)	R	This study
ETRM245	ETRM240 (pBAD30)	R	This study
RMA4376	ETRM243 (pRMA727)	S	This study
RMA4377	ETRM243 (pACYC184)	S	This study
Plasmids			
pCACTUS	Suicide vector; Cml^R; 30°C		(Morona *et al*, 1995)
pCACTUS-*icsP*::*kan*^{R}	pCACTUS with *icsP::kan*^{R} gene		This study
pSL1180	Cloning vector; Amp^R		(Brosius, 1989)
pSL1180-*icsP*::*kan*^{R}	pSL1180 with *icsP::kan*^{R} gene		This study
pACYC184	Cloning vector; Cml^R, Tet^R		(Rose, 1988)
pKTUWE	pACYC184 with *kan*^{R} gene; Kan^R		(Murray *et al*, 2003)
pKD4	Vector containing FRT-flanked *kan*^{R} gene		(Datsenko & Wanner, 2000)
pKD46	Red lambda plasmid; Amp^R; 30°C		(Datsenko & Wanner, 2000)
pCP20	FLP recombinase; Amp^R, Cml^R; 30°C		(Datsenko & Wanner, 2000)
pGEMT-Easy	Cloning vector; Amp^R		Promega
pGEMT-Easy::*icsP*^{HA}	pGEMT-Easy with *icsP*^{HA} gene; Amp^R		This study
pGEMT-Easy::*kan*^{R}	pGEMT-Easy with *kan*^{R} gene; Amp^R		This study
pBAD30	Arabinose-inducible pBAD promoter vector, Amp^R		(Guzman *et al*, 1995)
pIcsP	pBAD30 with *icsP* gene; Amp^R		This study
$pIcsP^{HA}$	pBAD30 with *icsP*^{HA} gene; Amp^R		This study
pRMA718	pUC1318 containing *S. flexneri rfb* region; Amp^R		(Van Den Bosch *et al*, 1997)
pRMA727	pACYC184 with *rmlD* gene; Cml^R		(Van Den Bosch *et al*, 1997)

* S, smooth LPS; R, rough LPS

Table 3.2: Oligonucleotides used in this study.

Primer name	Oligonucleotide sequence (5' - 3')*	Target	Genebank	nt position†
HA-encoding primers				
ET18	tacccgtacgacgtcccggactacgccagtaccaatatatctggcac	icsP gene	AF386526	221158
ET19	ggcgtagtccgggacgtcgtacgggtattgctcataaagagatgtatc	icsP gene	AF386527	221157
***icsP*-specific primers**				
ET3	gcggatccgtattgcttctgccatttcc	484 bp upstream *icsP*	AF386526	219783
ET4	gcgagctcgtccctgatagcactgttc	371 bp downstream *icsP*	AF386526	221618
ET10	gcggatccaaaaatatacttttatacctgcg	icsP gene	AF386526	221223
ET22	gcggtaccataaagtaagaagatcatggac	16 bp upstream *icsP*	AF386526	220251
ET25	gggaagctttcaaaaaatatacttttatacctg	*icsP* gene	AF386526	221225
pKD4 specific primers				
ET28	ccgggctagctgtgtaggctggagctgcttcg	FRT-*kan*^{R} priming site 1	AY048743	31
ET29	gcccgctagccatatgaatatcctcctta	FRT-*kan*^{R} priming site 2	AY048743	1488

* Underlined sequences indicate the nucleotides that encode the HA epitope

† nt, nucleotide

3.6. Article Results

3.6.1. Detection of IcsP expression by Western immunoblotting and immunofluorescence.

In initial experiments to detect cell surface IcsP, the optimal time point for IcsP expression in *S. flexneri* was determined. Western immunoblotting with a rabbit anti-IcsP was performed on whole cell lysates collected from wild-type *S. flexneri* 2457T and the 2457T *icsP* mutant (*icsP*) grown for 0.5, 1, 1.5, 2, 2.5 and 3 h after subculture. A band consistent with the size of the IcsP protein (~36 kDa) was detected at time points after 1.5 h for 2457T (Figure 3.1A, lanes 5, 7, 9 & 11) with high expression observed at 3 h (Figure 3.1A, lane 11). We reasoned that IcsP levels at this time point were high enough for subsequent immunofluorescent detection of IcsP. No expression of IcsP was observed for *icsP* as expected (Figure 3.1A, lanes 2, 4, 6, 8, 10 & 12). Subsequent attempts to detect IcsP on the surface of 2457T with rabbit anti-IcsP however were unsuccessful, even in a rough LPS strain (data not shown). We speculate that as the IcsP protein used to raise antisera was purified under denaturing conditions, this may have affected the resulting antibody's ability to detect native IcsP. However, difficulties with IF detection of cell surface IcsP with a polyclonal antibody have also been reported by others (Steinhauer *et al*, 1999). Hence, in an alternative approach to investigate the distribution of IcsP on the bacterial cell surface, a HA epitope was inserted into the IcsP protein.

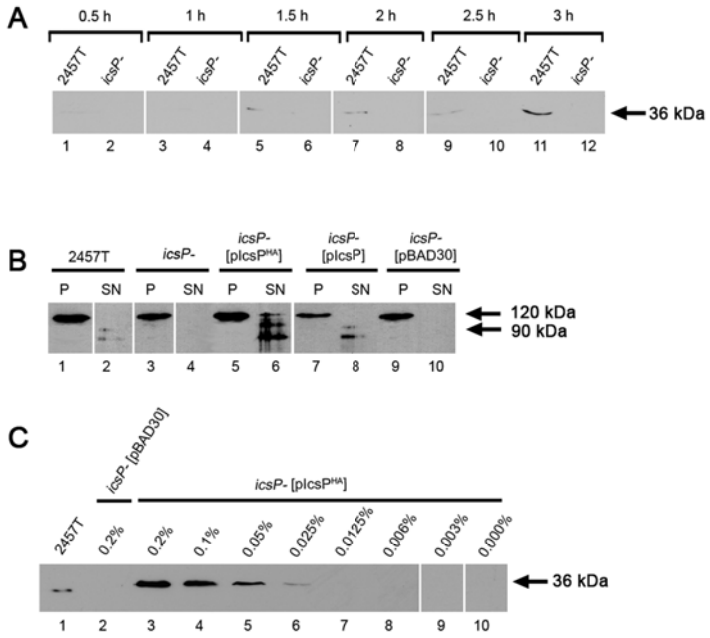


Figure 3.1: Detection of IcsP/IcsP^{HA} expression and activity on IcsA by Western immunoblotting.

(A) *S. flexneri* strains 2457T and 2457T *icsP*- mutant (*icsP*-) were grown in LB and whole cell lysate samples were taken at 0.5, 1, 1.5, 2, 2.5 and 3 h after subculture, followed by electrophoresis on a SDS 15 % polyacrylamide gel and Western immunoblotting with rabbit anti-IcsP antiserum; **(B)** *S. flexneri* strains 2457T, *icsP*-and *icsP*- harbouring pIcsP, pIcsP^{HA} or pBAD30 (as indicated) were grown in LB for 1.5 h to an OD600 reading of ~0.4, washed 3 times, and induced with arabinose for 1 h. Pellet and supernatant protein samples were then prepared and electrophoresed on a SDS 15 % polyacrylamide gel, followed by Western immunoblotting with rabbit anti-IcsA antibodies. The size of the full length IcsA protein (120 kDa) and the cleaved form of IcsA (95 kDa) are indicated; **(C)** *S. flexneri* strains 2457T and *icsP*- harbouring pIcsP^{HA} or pBAD30 were grown in LB as described in **(B)**, followed by induction with 0 %, 0.003 %, 0.006 %, 0.0125 %, 0.025 %, 0.05 %, 0.1 % or 0.2 % (w/v) arabinose for 1 h. Whole cell lysates were prepared and electrophoresed on a SDS 15 % polyacrylamide gel, followed by Western immunoblotting with rabbit anti-IcsP antiserum. The size of the full length IcsP protein (36 kDa) is indicated in **(A)** and **(C)**. Each lane contains 5 x 10^7 bacterial cells of each strain.

3.6.2. Insertion of a HA epitope into IcsP.

IcsP is 60 % identical in primary amino acid sequence to the *E. coli* protease OmpT and computer structure modelling predicts that both proteins possess similar β-barrel structures (Figure 3.2A). Based on sequence alignments with *E. coli* OmpT (Figure 3.2B), a HA epitope tag (YPYDVPDYA) was hence inserted into the OM loop 5 of IcsP (Figure 3.2B) using SOE PCR (as described in the *Methods*). An area within this loop region with some sequence diversity between IcsP and OmpT was selected as we reasoned that this sequence variability might allow the protein to accommodate the epitope insertion with little disturbance to the overall structure. The OM loop 5 region was also selected to increase the chance of surface detection by antibodies. OM loops 2 and 4 were not selected for HA tag insertion to avoid the proposed catalytic residues present in OmpT (Vandeputte-Rutten *et al*,

2001) which also exist in IcsP (Figure 3.2C). The IcsP^{HA} (as well as IcsP) coding regions were placed in front of the pBAD promoter in pBAD30 (Guzman *et al*, 1995) to allow expression control with arabinose. Expression of IcsP/IcsP^{HA} was confirmed by Western immunoblotting with anti-IcsP or anti-HA antibodies (Figure 3.3A & B, lanes WM and Figure 3.1A & C).

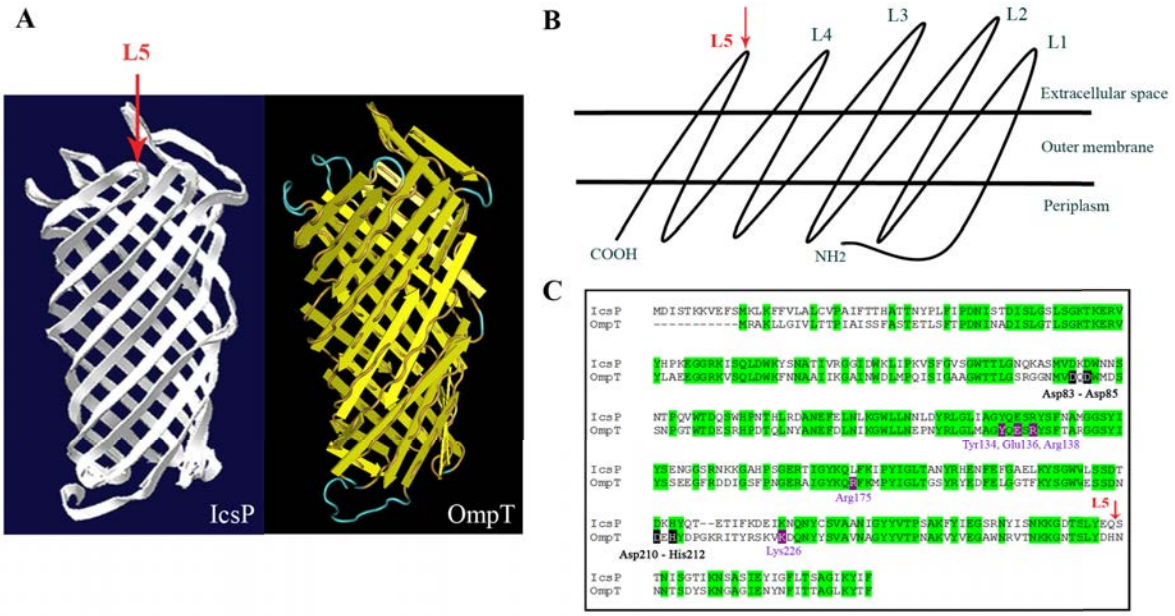


Figure 3.2: Putative structure of IcsP and location of HA epitope insertion in IcsP.

(A) IcsP was modelled using the SWISS-MODEL Protein Modelling Server (swissmodel.expasy.org//SWISS-MODEL) (left) and compared to the structure of OmpT (PDB 1I78) (right); **(B)** Schematic diagram of IcsP showing the location of the OM loops 1 - 5; **(C)** Amino acid sequence alignment of IcsP (AF001633) and OmpT (P09169) showing 60 % identity (green shaded regions). The black-boxed amino acids in the sequence of OmpT refer to active site residues found in OM loops 2 and 4 38], and the purple-boxed amino acids refer putative LPS binding sites 38]. The location of the HA epitope (YPY DVP DYA) insertion into the putatively non-active OM loop 5 (L5) is indicated by the red arrow in **(A)** – **(C)**.

3.6.3. IcsP^{HA} activity on IcsA.

To determine whether insertion of a HA epitope into the OM of IcsP affected IcsP's protease activity on IcsA, pellet and supernatant protein preparations of 2457T, *icsP* and arabinose-induced *icsP* strains expressing pIcsP^{HA}, pIcsP and pBAD30 were subjected to Western immunoblotting with an anti-IcsA antibody. The full length 120 kDa IcsA protein was detected in cell pellet samples of all strains as expected (Figure 3.1B, lanes 1, 3, 5, 7 and 9), while the 95 kDa cleaved form of IcsA was only detected in the supernatant sample of 2457T, *icsP* [pIcsP^{HA}] and *icsP* [pIcsP] (Figure 3.1B, lanes 2, 6 & 8). These results suggest that the insertion of the HA tag into the OM loop 5 of IcsP does not affect its ability to cleave IcsA, and hence the IcsP^{HA} protein is functional.

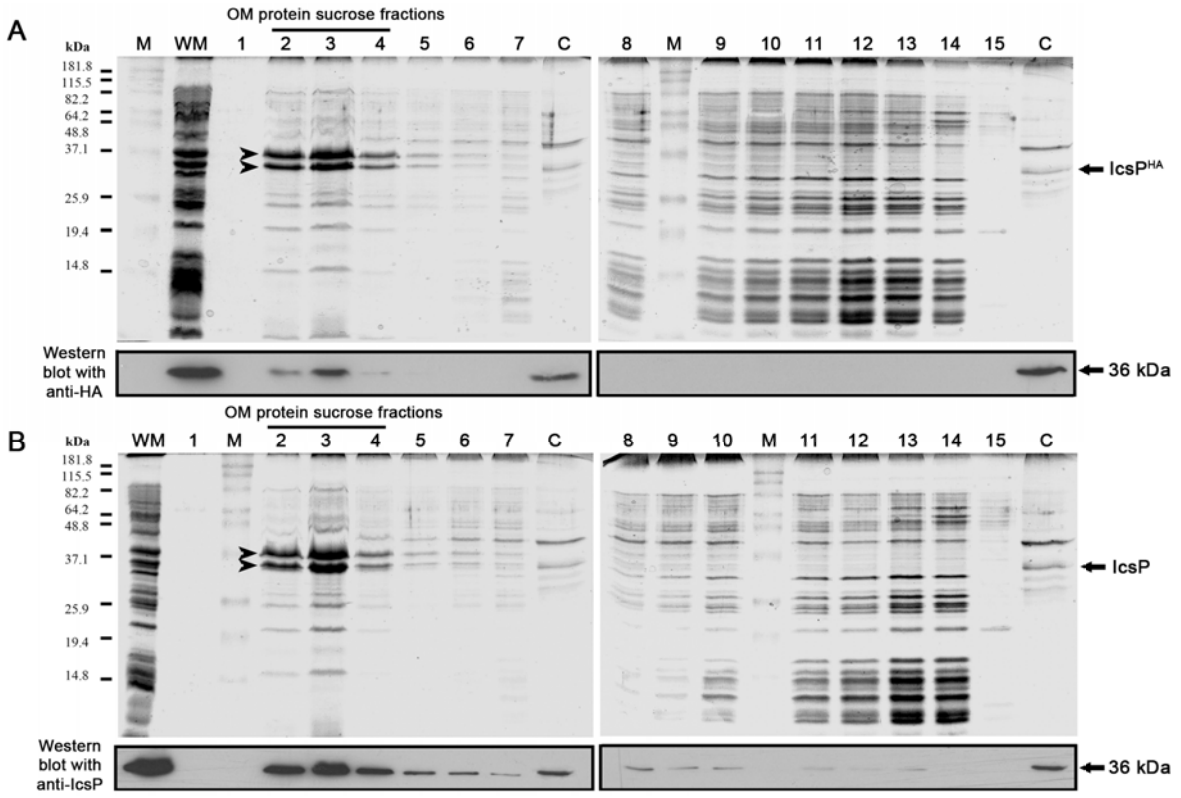


Figure 3.3: Analysis of IcsP/IcsP^{HA} subcellular localisation by sucrose density gradient centrifugation of WM.

S. flexneri icsP$^-$ strains harbouring either pIcsP^{HA} and pIcsP were grown in LB for 1.5 h to an OD_{600} reading of ~0.4, washed 3 times, and induced with arabinose for 1 h (OD_{600} ~1). Whole membranes (WM) were prepared by French press lysis, and subjected to sucrose density gradient centrifugation as described in the *Methods*. Fractions (0.5 ml; numbered 1-15) were collected and samples of each were electrophoresed on SDS 15 % polyacrylamide gels for staining with Coomassie Blue and Western immunoblotting with anti-HA or anti-IcsP antibodies; **(A)** The results obtained for *icsP*$^-$ [pIcsP^{HA}]; **(B)** The results obtained for *icsP*$^-$ [pIcsP]. The migration positions of the Benchmark Prestained Marker (M) Standards (Invitrogen) are indicated on the left in kDa. The major OM proteins OmpF+OmpC and OmpA are indicated by the two arrowheads in lane 2. Lane C corresponds to the control lanes containing WC samples of *icsP*$^-$ [pIcsP^{HA}] induced with 0.2 % (w/v) arabinose for protein overexpression (positive with both anti-HA and anti-IcsP by Western immunoblotting) as indicated by the arrows on the right. The sucrose fractions in lanes 2 – 4 contain most of the OM proteins, and the sucrose fractions in lanes 8 – 14 contain most of the inner (cytoplasmic) membrane proteins.

3.6.4. Localisation of IcsP/IcsP^{HA} protein to the OM.

Since IcsP^{HA} was modified compared to IcsP and expressed from the pBAD promoter, the presence of IcsP^{HA} protein exclusively in the OM was confirmed by using sucrose density gradient centrifugation. The WM of *icsP*$^-$ [pIcsP^{HA}] and *icsP*$^-$ [pIcsP] were fractionated into CM and OM on sucrose gradients, and fractions subjected to SDS 15 % polyacrylamide gel electrophoresis, prior to visualisation by Coomassie Blue staining and Western immunoblotting with anti-HA and anti-IcsP. Analysis of the sucrose gradient samples showed that fractions which were enriched with OM proteins (OmpF, OmpC, and OmpA)

(Fehniger *et al*, 1986) contained the majority of the 36 kDa IcsP^{HA} and IcsP proteins (Figure 3.3A & B, lanes 2 to 5). These results indicate that IcsP^{HA} is localised to the OM similar to IcsP, and that the HA insertion did not result in any dramatic disruption of IcsP protein localisation.

3.6.5. Detection of arabinose induced IcsP^{HA} expression by Western immunoblotting.

Wild-type *S. flexneri* 2457T showed optimal expression levels of native IcsP at 3 h (Figure 3.1A). To investigate the conditions required for comparable IcsP^{HA} expression, whole cell lysate samples were prepared from *icsP* [pIcsP^{HA}] and *icsP* [pBAD30] induced with 0 %, 0.003 %, 0.006 %, 0.0125 %, 0.025 %, 0.05 %, 0.1 % or 0.2 % (w/v) arabinose for 1 h. Western immunoblotting with anti-IcsP showed that a band of ~36 kDa was detected for 2457T and *icsP* [pIcsP^{HA}] induced with 0.025 %, 0.05 %, 0.1 % and 0.2 % (w/v) arabinose (Figure 3.1C, lanes 1, 3 - 6), with expression levels comparable to native IcsP for *icsP* [pIcsP^{HA}] observed between 0.025 % - 0.05 % (w/v) arabinose induction (Figure 3.1C, lanes 5 & 6). Induction at 0.03 % (w/v) arabinose was hence chosen for all subsequent experiments.

3.6.6. Cell surface detection of IcsP^{HA} distribution in *S. flexneri* 2457T and the effect of tunicamycin treatment.

Having established an induction protocol that closely approximates the level of expression of IcsP^{HA} to native IcsP, we next attempted to detect IcsP^{HA} at these levels on the surface of 2457T *icsP* using indirect QD IF microscopy. Strains *icsP* [pIcsP^{HA}] and *icsP* [pBAD30] were cultured with 0.03 % (w/v) arabinose, fixed, and then probed for HA using a primary anti-HA antibody and a secondary QD 525 conjugated antibody. However, no IcsP^{HA} was detected on the cell surface, and intensity scans of *icsP* [pIcsP^{HA}] were equivalent to *icsP* [pBAD30] (Figure 3.10[S]).

Since IcsP^{HA} could not be detected on the cell surface of *icsP* [pIcsP^{HA}], we reasoned that the presence of LPS Oag may mask the detection of surface IcsP as this has previously been shown for IcsA (Morona & Van Den Bosch, 2003b). An LPS Oag depletion-regeneration assay (Teh *et al*, 2012b) was hence carried out on *icsP* [pIcsP^{HA}] and *icsP* [pBAD30] induced with 0.03 % (w/v) arabinose, followed by IF labelling. This assay involves the use of an inhibitor of the WecA enzyme necessary for Oag subunit biosynthesis using tunicamycin, and polymyxin B nonapeptide (PMBN) was used to improve OM penetration. Upon removal of these two chemicals from growing bacteria, LPS Oag is regenerated (Teh *et*

al, 2012b). Analysis of the resulting LPS by SDS-PAGE and silver staining showed that *icsP*⁻ [pIcsP^{HA}] and *icsP*⁻ [pBAD30] samples treated with tunicamycin/PMBN (TP) had depleted LPS Oag (Figure 3.4A, lanes 5 & 6), with Oag production restored (R) upon removal of TP (Figure 3.4A, lanes 7 & 8). Untreated (U) and PMBN treated (P) samples showed no inhibition of Oag biosynthesis as expected (Figure 3.4A, lanes 1 - 4).

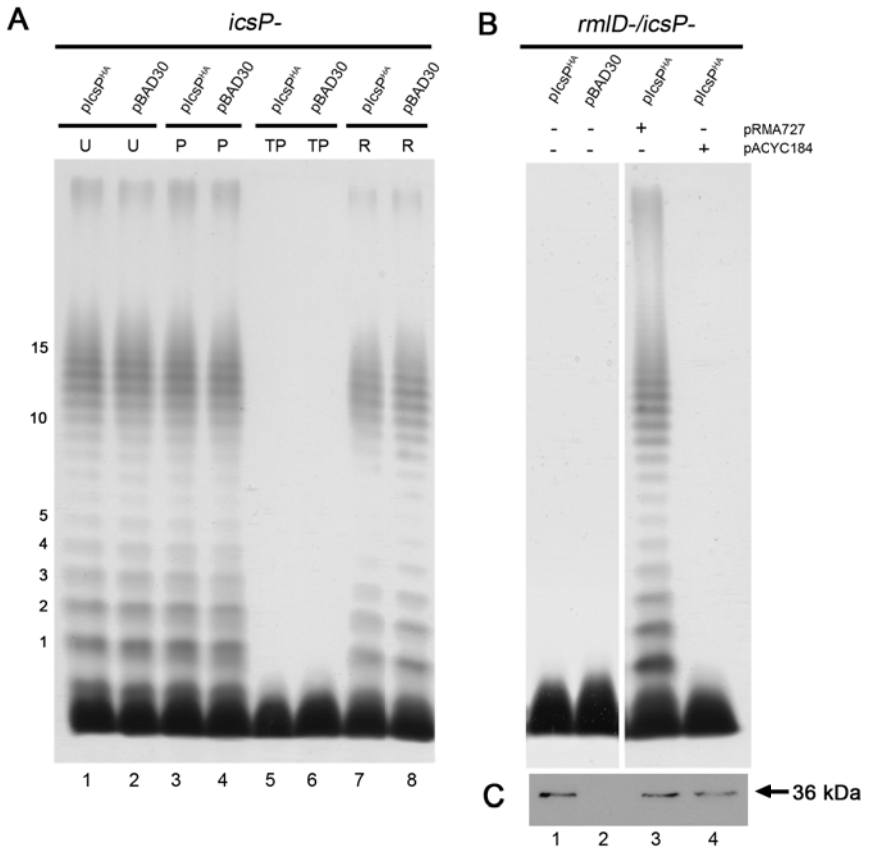


Figure 3.4: Effect of tunicamycin on the LPS of strains expressing IcsP^{HA}.

(A) Smooth LPS 2457T *icsP*⁻ strains harbouring pIcsP^{HA} or pBAD30 were grown to an OD_{600} reading of ~0.8 in LB, washed 3 times, and treated without TP (U), with PMBN only (P) or with TP treatment for 2 h. Strains were then induced with 0.03 % (w/v) arabinose for 1 h, washed 3 times, and grown for an additional 3 h for restoration (R) of LPS Oag; **(B)** Rough LPS 2457T *icsP*⁻/*rmlD*⁻ strains harbouring pIcsP^{HA} and either pRMA727 or pACYC184 (as indicated) were grown to an OD_{600} reading of ~0.4 in LB, washed 3 times, and induced with arabinose for 1 h. LPS from strains described in **(A)** and **(B)** were isolated and detected by silver staining as described in the *Methods*. The first 15 Oag RUs are indicated on the side of each gel. Each lane contains ~2 x 10^8 bacterial cells of each strain; **(C)** Western blots on whole cell lysates obtained from strains in **(B)** were probed with rabbit anti-IcsP antiserum. The size of the full length IcsP^{HA} protein (~36 kDa) is indicated. Each lane contains 5 x 10^7 bacterial cells of each strain.

Following successful depletion of LPS Oag, QD IF microscopy was performed on the above mentioned samples. IcsP^{HA} could not be detected on untreated or PMBN treated *icsP*⁻ [pIcsP^{HA}] cells, as expected (Figure 3.11[S]A). However, surface IcsP^{HA} was detected on *icsP*⁻ [pIcsP^{HA}] cells treated with TP, suggesting that LPS Oag is able to mask

antibody accessibility to IcsP^{HA} (Figure 3.11[S]A). Interestingly, IcsP^{HA} appeared to be distributed asymmetrically over the cell surface of the majority of cells examined and fluorescence intensity line-scans revealed that IcsP^{HA} localised preferentially to one pole of non-septating cells and the septa of septating cells (Figure 3.11[S]A). To determine if the distribution of IcsP^{HA} on the cell surface segregated to either the new pole or the old pole, additional staining of the bacteria was conducted with anti-IcsA antibodies since IcsA is known to localise to the old cell pole (Goldberg *et al*, 1993). Again, IcsP^{HA} could not be detected on untreated or PMBN treated *icsP*$^{-}$ [pIcsP^{HA}] cells (Figure 3.5A & B) but could be detected after TP treatment (Figure 3.5C). Peak IcsP^{HA} detection was consistently observed at the new pole (opposing IcsA at the old pole) and at the septum of septating bacteria as shown by line-scans (Figure 3.5C). As expected, no surface IcsP^{HA} was detected for untreated, PMBN treated, or TP treated samples of *icsP*$^{-}$ [pBAD30] when stained for IcsP^{HA} (Figure 3.11[S]B) or double stained for IcsP^{HA} and IcsA (Figure 3.12[S]). Notably, IcsA on *icsP*$^{-}$ [pBAD30] was detected at higher amounts laterally and at the septa of bacteria (Figure 3.12[S]) as previously seen for ⊿*icsP* strains (Steinhauer *et al*, 1999).

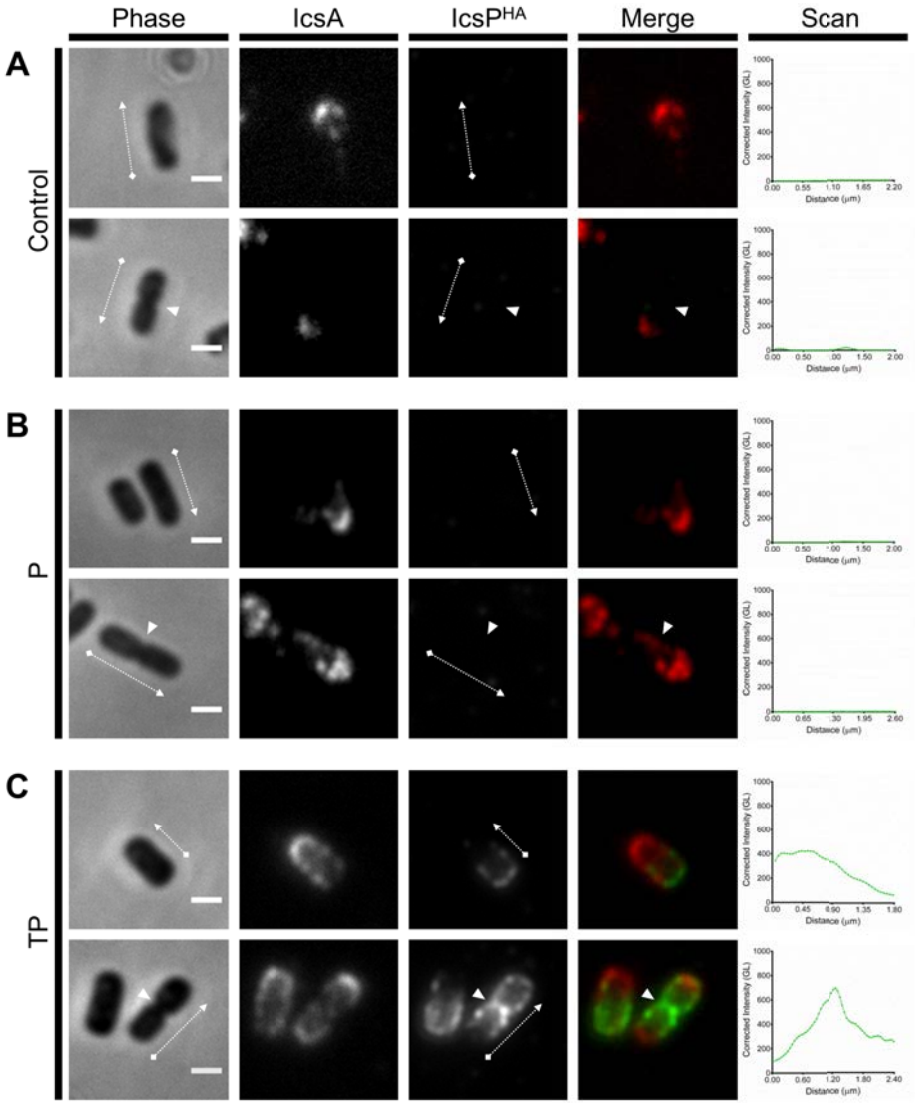


Figure 3.5: Effect of LPS Oag-depletion in detection of surface IcsP^{HA}.

Smooth LPS 2457T *icsP*- strains harbouring pIcsP^{HA} were subcultured in LB broth to an OD_{600} reading of ~0.8, washed 3 times in LB, and then further cultured for 2 h; **(A)** in the absence of TP, **(B)** in the presence of PMBN only, or **(C)** in the presence of TP. Arabinose was included in the final hour of treatment at a concentration of 0.03 % (w/v). Samples were then fixed and subjected to QD IF using antibodies to HA epitope and IcsA. Non-septating and septating cells (upper and lower rows respectively) are shown for each treatment group. Representative bacteria are shown. Scan = Single line-scans measuring the fluorescence intensity of IcsP^{HA} detected along the surface of the bacterium, Bars = 1 µm, Arrows = direction of line-scan, Arrow heads = septa, Control = grown in absence of both tunicamycin and PMBN, P = PMBN, TP = tunicamycin/PMBN, GL = Grey level, Phase = phase contrast image, IcsP^{HA} = image of fluorescence at 525 nm, IcsA = image of fluorescence at 625 nm, Merge = overlay of IcsP^{HA} and IcsA images.

3.6.7. Cell surface detection of IcsP^{HA} distribution in rough LPS *S. flexneri* 2457T.

To further investigate the effect of Oag on IcsP distribution, a 2457T *icsP*/*rmlD*- double mutant was constructed to independently assess the distribution of IcsP^{HA} in a rough LPS background. Mutation of the *rmlD* gene in *S. flexneri* results in a strain which is unable to

synthesise the precursor deoxythymidine diphosphate (dTDP)-rhamnose required for Oag repeat units and, hence, results in a rough LPS phenotype (Reeves *et al*, 1996, Teh *et al*, 2012b). Analysis of the resulting LPS conferred by *icsP*$^{-}$/*rmlD*$^{-}$ strains expressing pIcsP^{HA} and pBAD30 showed that rough LPS was observed for both (Figure 3.4B, lanes 1 & 2), with a band consistent with the size of the IcsP^{HA} protein (36 kDa) detected only in the *icsP*$^{-}$/*rmlD*$^{-}$ [pIcsP^{HA}] sample by Western immunoblotting, as expected (Figure 3.4C, lane 1).

Fixed samples of *icsP*$^{-}$/*rmlD*$^{-}$ [pIcsP^{HA}] and *icsP*$^{-}$/*rmlD*$^{-}$ [pBAD30] cells were then probed for both IcsP^{HA} and IcsA using the same QD IF staining protocol as previously conducted for smooth strains. Similarly to LPS Oag-depleted *icsP*$^{-}$ [pIcsP^{HA}], IcsP^{HA} was detected on the bacterial surface most predominately at the new pole (Figure 3.6A) and the septum (Figure 3.6B). Again, the majority of septating cells had higher peak IcsP^{HA} intensity at the septum than new poles of non-septating cells (line-scans Figure 3.6A & B). Single staining of these cells for IcsP^{HA} was also conducted and yielded the same localisation results (Figure 3.13[S]). As expected, IcsP^{HA} was not detected on the *icsP*$^{-}$/*rmlD*$^{-}$ [pBAD30] strain in IF microscopy experiments when either single (Figure 3.13[S]) or double stained (Figure 3.6B). Again, IcsA on *icsP*$^{-}$/*rmlD*$^{-}$ [pBAD30] was detected at higher amounts laterally and at the septum (Figure 3.6A & B).

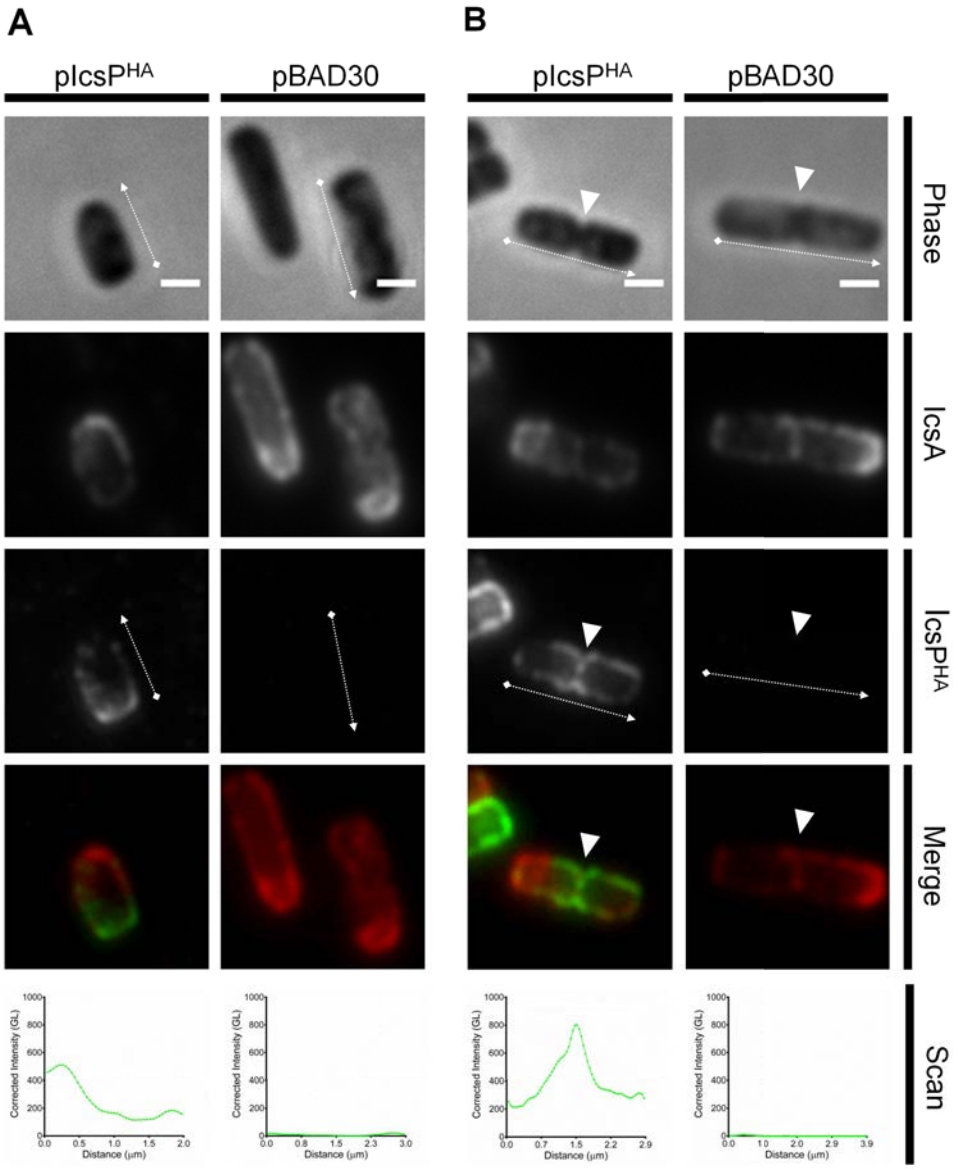


Figure 3.6: Localisation of IcsP^{HA} on the surface of rough LPS 2457T *icsP-/rmlD-*.

Rough LPS 2457T *icsP-/rmlD-* strains harbouring pIcsP^{HA} (left columns) or pBAD30 (right columns) were subcultured in LB broth for 1.5 h to an OD_{600} reading of ~0.4. Cultures were then washed 3 times in LB, induced with 0.03 % (w/v) arabinose for 1 h, fixed, and subjected to QD IF using antibodies for HA epitope and IcsA. Non-septating and septating life stages are shown in **A** and **B** respectively. Representative bacteria are shown. Scan = Single line-scans measuring the fluorescence intensity of IcsP^{HA} detected along the surface of the bacterium, Bars = 1 μm, Arrows = direction of line-scan, Arrow heads = septa, GL = Grey level, Phase = phase contrast image, IcsP^{HA} = image of fluorescence at 525 nm, IcsA = image of fluorescence at 625 nm, Merge = overlay of IcsP^{HA} and IcsA images.

To again demonstrate the effect of LPS Oag masking of IcsP^{HA}, smooth LPS structure was restored in the *icsP-/rmlD-* strain expressing pIcsP^{HA} by transforming pRMA727 carrying a functional *rmlD* gene. Analysis of the resulting LPS conferred by *icsP-/rmlD-* [pIcsP^{HA}][pRMA727] showed restored smooth LPS phenotype (Figure 3.4B, lane 3), and when probed for IcsP^{HA} by QD IF microscopy, was barely detectable (Figure 3.7A). The control

icsP/*rmlD*⁻ [pIcsP^{HA}] strain carrying pACYC184 conferred a rough LPS phenotype (Figure 3.4B, lane 4) and IcsP^{HA} was detected by IF microscopy with the same distribution observed previously (Figure 3.7B).

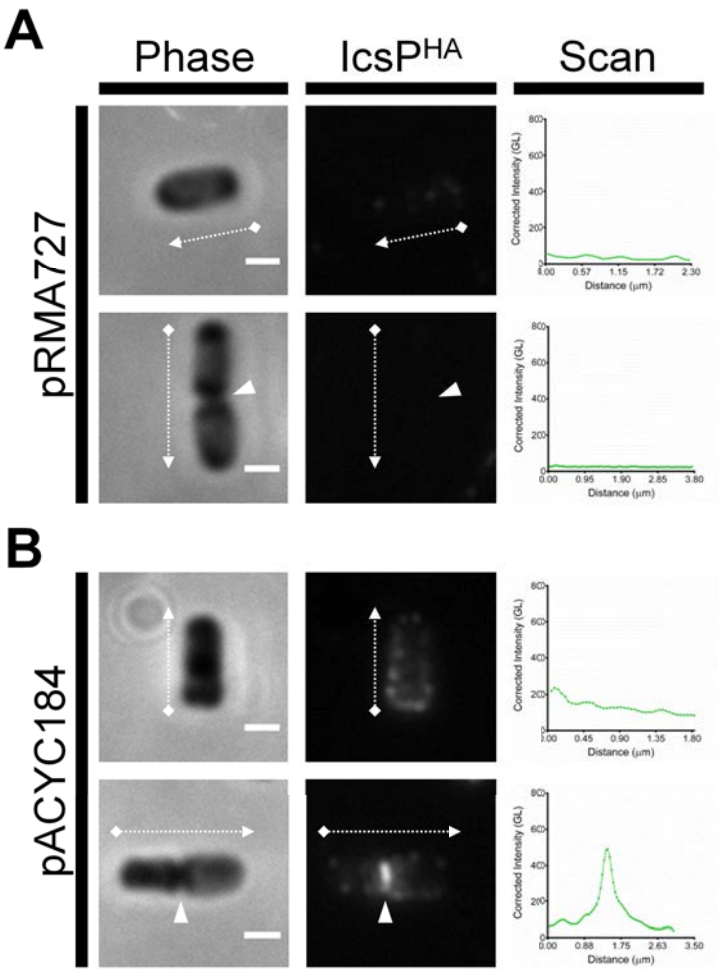


Figure 3.7: Complementation of *rmlD* restores LPS masking effect on IcsP^{HA} detection.

Rough LPS 2457T *icsP*⁻/*rmlD*⁻ strains harbouring pIcsP^{HA}, and either **(A)** pRMA727 or **(B)** pACYC184, were subcultured in LB broth for 1.5 h to an OD_{600} reading of ~0.4. Cultures were then washed 3 times in LB, induced with 0.03 % (w/v) arabinose for 1 h, fixed, and subjected to QD IF using antibodies for the HA epitope. Representative bacteria are shown. Scan = Single line-scans measuring the fluorescence intensity of IcsP^{HA} detected along the surface of the bacterium, Bars = 1 µm, Arrows = direction of line-scan, Arrow heads = septa, GL = Grey level, Phase = phase contrast image, IcsP^{HA} = image of fluorescence at 525 nm.

3.6.8. Multi-cell line-scan analysis of IcsP^{HA} surface distribution.

To determine if the observed sub-cellular preference of IcsP^{HA} to the new poles and septa is a statistically significant phenomenon, cumulative QD IF line-scan analyses on both IcsP^{HA} / IcsA double stained LPS Oag-depleted *icsP*⁻ strains and rough *icsP*⁻/*rmlD*⁻ strains was conducted for a larger population of cells. Scanning was initiated from the new pole to

the old pole (as marked by IcsA) for non-septating cells, and from the septum to the old pole for septating cells (Figure 3.8C). From 9 independent samples of LPS Oag-depleted *icsP*⁻ [pIcsP^{HA}] and *icsP*⁻ [pBAD30], a total of 450 non-septating cells each were line-scanned. Additionally a total of 355 and 256 septating cells were scanned for *icsP*⁻ [pIcsP^{HA}] and *icsP*⁻ [pBAD30] respectively. The resultant mean fluorescence intensity profiles show that IcsP^{HA} is preferentially localised at the new pole and tends to gradually decrease towards the old pole on non-septating cells expressing pIcsP^{HA} (Figure 3.8A). For septating cells, IcsP^{HA} mean intensity is localised highest at the septum and declines more steeply towards the old pole (Figure 3.8A). As expected, the intensity profiles of *icsP*⁻ [pBAD30] cells were at a negligible level (Figure 3.8A). Statistical analysis of mean IcsP^{HA} intensity at discrete bacterial positions of LPS-depleted *icsP*⁻ [pIcsP^{HA}] confirmed that the localisation of IcsP^{HA} is: (i) 2-fold higher at the new pole of non-septating cells than the old pole ($P = 0.004$), (ii) 2.9-fold higher at the septum of septating cells than the old pole ($P < 0.0001$), and (iii) 1.5-fold higher at septa of septating cells compared to new poles of non-septating cells ($P = 0.038$) (Figure 3.8D).

For the rough LPS strains, 9 independent samples of *icsP*⁻/*rmlD*⁻ [pIcsP^{HA}] and *icsP*⁻/*rmlD*⁻ [pBAD30] were investigated with a total of 450 non-septating cells each line-scanned. Additionally, a total of 172 and 173 septating cells were scanned for *icsP*⁻/*rmlD*⁻ [pIcsP^{HA}] and *icsP*⁻/*rmlD*⁻ [pBAD30] respectively. Again, the mean intensity profiles of non-septating cells expressing pIcsP^{HA} shows IcsP^{HA} is preferentially localised at the new pole and tends to gradually decrease towards the old pole (Figure 3.8B). Likewise, IcsP^{HA} mean intensity of septating cells is localised highest at the septum and declines very steeply towards the old pole (Figure 3.8B). As expected, the mean intensity profiles of *icsP*⁻/*rmlD*⁻ [pBAD30] were at a negligible level (Figure 3.8B). Statistical analysis of mean IcsP^{HA} intensity at discrete bacterial positions of *icsP*⁻/*rmlD*⁻ [pIcsP^{HA}] again confirmed that the localisation of IcsP^{HA} is: (i) 2.2-fold higher at the new pole of non-septating cells than the old pole ($P < 0.0001$), (ii) 3.6-fold higher at the septum of septating cells than the old pole ($P < 0.0001$), and (iii) 1.7-fold higher at septa of septating cells compared to new poles of non-septating cells ($P = 0.0002$) (Figure 3.8D). Although, IcsP^{HA} fluorescence intensity was an average of 2.5 times higher on *icsP*⁻/*rmlD*⁻ [pIcsP^{HA}] cells than LPS-depleted *icsP*⁻ [pIcsP^{HA}] cells, the ratios of intensity between bacterial positions were comparable to the respective fold-changes observed for LPS Oag-depleted *icsP*⁻ [pIcsP^{HA}], suggesting that IcsP^{HA} is localised similarly in both types of Oag deficient cells. The lower level of IcsP^{HA} detection on LPS-depleted *icsP*⁻ [pIcsP^{HA}] compared to *icsP*⁻/*rmlD*⁻ [pIcsP^{HA}] indicates that TP treatment is not 100 % efficient in inhibiting Oag synthesis.

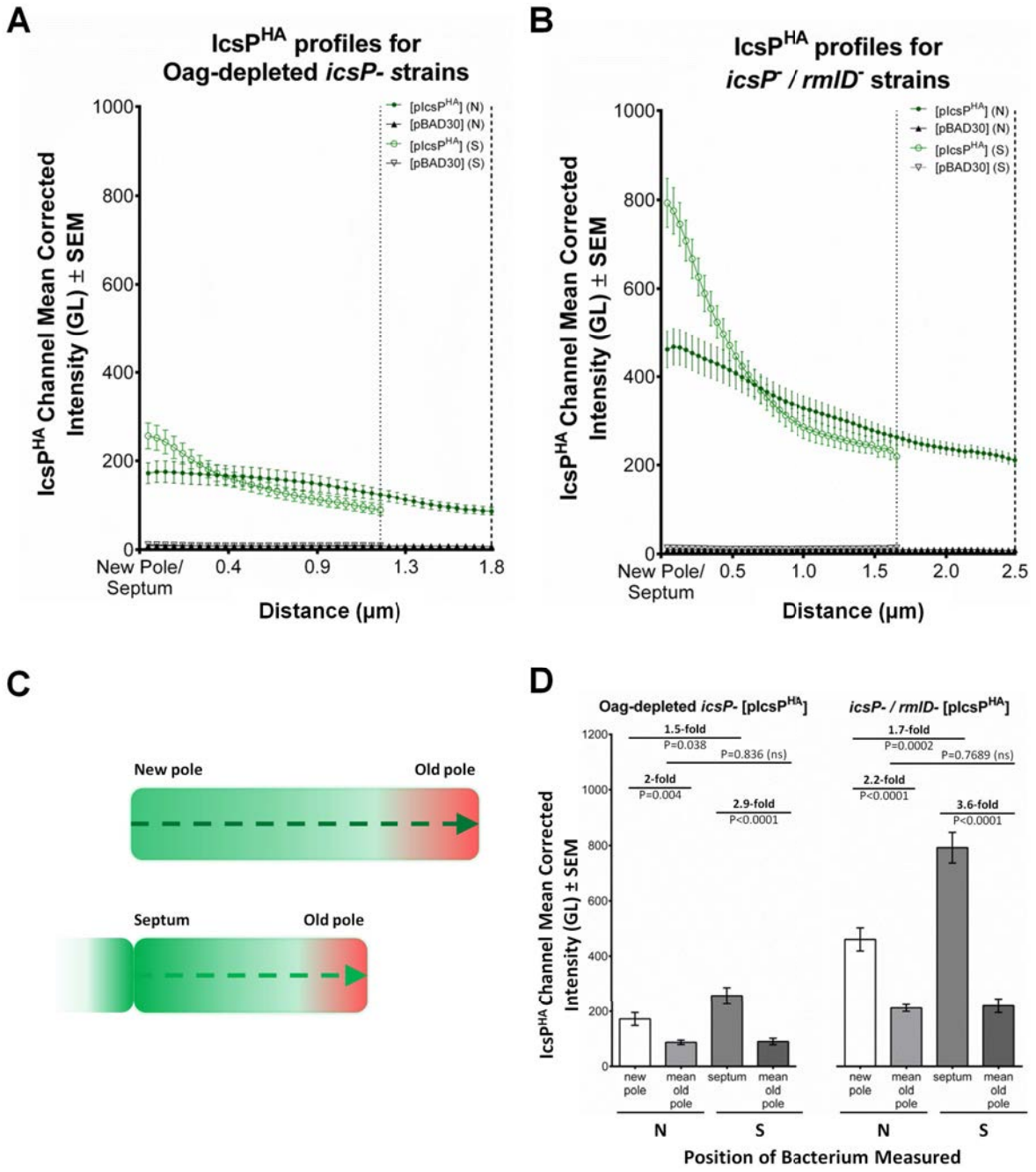


Figure 3.8: Quantification and statistical analysis of IcsP^{HA} surface distribution.

A total of 9 independent cultures of Smooth LPS 2457T *icsP*- strains harbouring pIcsP^{HA} or pBAD30 were subcultured in LB for 1.5 h to an OD_{600} reading of ~0.4, washed 3 times in LB, and then further cultured for 2 in the presence of TP for Oag-depletion. Arabinose was included in the final hour of treatment at a concentration of 0.03 % (w/v). A total of 9 independent cultures of Rough LPS 2457T *icsP*-/*rmlD*- strains harbouring pIcsP^{HA} or pBAD30 were also subcultured LB broth to an OD_{600} reading of ~0.4, washed 3 times in LB, and induced with 0.03 % (w/v) arabinose for 1 h. A sample from each culture was then fixed and subjected to QD IF using antibodies for HA epitope and IcsA. Fluorescence intensity scans of IcsP^{HA} (525 nm wavelength) were conducted on multiple bacteria from each independent sample and accumulated scans were used to create mean intensity profiles. A total 450 non-septating bacteria were scanned for each strain. A further 355, 256, 172, and 173 septating bacteria were scanned for Oag-depleted *icsP*- [pIcsP^{HA}], Oag-depleted *icsP*- [pBAD30], *icsP*-/*rmlD*- [pIcsP^{HA}], and *icsP*-/*rmlD*- [pBAD30] respectively. Resultant IcsP^{HA} mean surface profiles for **(A)** Oag-depleted *icsP*- and **(B)** *icsP*-/*rmlD*- bacteria are shown. Dotted and dashed vertical lines indicate mean positions of old poles for septating and non-septating bacteria respectively. The schematic **(C)** shows

methodology of IcsP^{HA} intensity scan directions from either new poles to old poles (as marked by IcsA) of non-septating bacteria, or septa to old poles of septating bacteria. Student's two-tailed t-tests **(D)** were also conducted on the fold differences of mean IcsP^{HA} intensities between discrete positions (new pole, old pole, and septum) of Oag-depleted *icsP*$^{-}$ and *icsP*$^{-}$/*rmlD*$^{-}$ bacteria. SEM = standard error of the mean, GL = Grey level, N = non-septating cells, S = septating cells, P = p-value, ns = not significant.

3.7. Article Discussion

The cell surface distribution of neither IcsP, or any other member of the Omptin family, had not been previously determined. In this study we investigated the distribution of IcsP on the cell surface of *S. flexneri* 2a 2457T using a HA-tagged IcsP protein (Figure 3.2) under pBAD control. Characterisation of $IcsP^{HA}$ showed that it was functionally able to cleave IcsA, and was secreted into the OM comparably to IcsP (Figure 3.1B & Figure 3.3). However, when $IcsP^{HA}$ was expressed at native IcsP equivalent levels (Figure 3.1A & C), it was undetectable in the OM via QD IF microscopy in smooth LPS *S. flexneri* (Figure 3.10[S]) but detectable on both LPS Oag-depleted and rough LPS *Shigella* bacteria (Figure 3.4, Figure 3.5 & Figure 3.6). Furthermore, this masking effect was restored in rough strains upon complementation of *rmlD* (Figure 3.4B & Figure 3.7). We suggest that the long LPS Oag chains of smooth strains sterically hinder the accessibility of antibodies to the OM surface. This type of protein masking by LPS Oag has also been shown for IcsA in 2457T (Morona & Van Den Bosch, 2003b), and several other OM proteins (Van der Ley *et al*, 1986a, Van der Ley *et al*, 1986b). It is interesting to note that multi-cell line-scan analysis of $IcsP^{HA}$ detection on the cell surface showed that $IcsP^{HA}$ fluorescence intensity was higher for rough LPS cells than for LPS Oag-depleted cells (Figure 3.8). Similar to its Omptin homolog, OmpT, IcsP possesses most of the putative LPS-binding sites found in OmpT (Figure 3.2C) (Vandeputte-Rutten *et al*, 2001) and may also interact with LPS. Since TP is only partly effective at blocking LPS Oag synthesis, it is possible that the few LPS Oag molecules synthesised are still closely bound to IcsP. This may cause a small amount of masking and may explain the lower fluorescence intensity observed on LPS Oag-depleted cells (Figure 3.8A) compared to the rough LPS cells (Figure 3.8B).

IcsA distributions in ⊿*icsP* strains previously provided indirect evidence to suggest that IcsP is distributed equally over the cell surface of *S. flexneri* (d'Hauteville *et al*, 1996, Shere *et al*, 1997, Steinhauer *et al*, 1999). However, once the masking effect of LPS Oag was circumvented, we were able to extensively study the distribution profile of $IcsP^{HA}$ in the OM. This work has shown that $IcsP^{HA}$ (and likely IcsP) has an asymmetrical distribution which may be cell cycle dependent. On dividing cells IcsP accumulates at high concentration at the septum and then declines steeply towards the old pole of the emerging daughter cell. However, once divided, detection of IcsP at the new pole (the pole derived from the septum) of daughter cells decreased significantly and then the decline towards the old pole is more moderate. To explain this, we propose that IcsP is directed to septa of dividing cells and that daughter cells 'inherit' a higher concentration of IcsP at new poles (model Figure 3.9). As the

daughter cell elongates, IcsP is laterally diluted, resulting in the gentle gradient on non-septating cells. In this model, lateral dilution of IcsP may occur by membrane insertion events during cell elongation similarly to that described for *E. coli* LamB (Ursell *et al*, 2012). This model predicts increased cleavage of IcsA at the septum and lateral regions of the cell in order to help set-up a relatively higher concentration of IcsA at the old pole and to maintain the old pole preference of IcsA in daughter cells. Additionally, the model may explain, and also fits, the common observation that IcsA distributions and intensities vary between cells in a given population of the same strain.

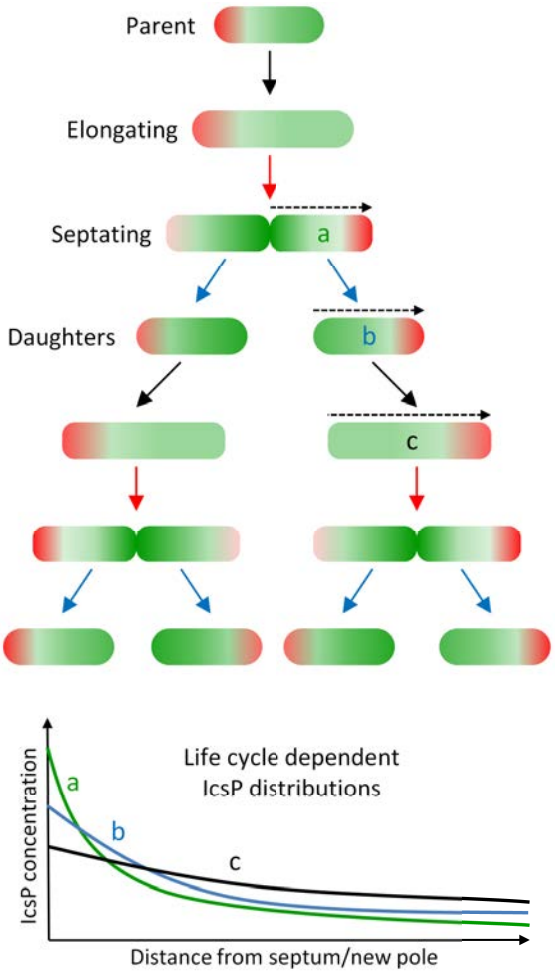


Figure 3.9: Model for differential inheritance of spatially separated OM proteins IcsA and IcsP.

Successive generations of *S. flexneri* are shown including: septating cells, newly formed daughters, and elongating cells. Solid black, red, and blue arrows denote the events of growth/elongation, septation, and division respectively. IcsA localisation is represented by red shading and IcsP by green shading where the darker the shading indicates higher protein concentration in the OM. The bottom graph represents IcsP distributions and gradients that result from hypothetical line-scans of: (a) a septating cell, (b) a newly formed daughter cell, and (c) an elongated cell from this model. IcsP on septating cells (a) declines steeply from the septum towards the old pole, whereas it's gradient on elongated cells (b) is more moderate between poles. A newly formed cell (c) would have an intermediate gradient. Dotted arrow = direction of hypothetical line-scan.

Many proteins are known to accumulate at the inner membrane of the mid-cell in order to mediate septum formation and cell division – for instance, FtsZ polymerises at the cytoplasmic face to form the Z-ring and attract further divisome components (Egan & Vollmer, 2013). However, few OM proteins have been shown to concentrate to the septum. A notable exception to this is the OM lipoprotein LpoB which localises as distinct foci at septa and complexes with periplasmic penicillin binding protein 1B tethering it to other divisome components (Typas *et al*, 2010, Egan & Vollmer, 2013). Interestingly, septal localisation of LpoB is lost when peptidoglycan synthesis is inhibited (Typas *et al*, 2010). Whether the localisation of integral OM protein IcsP to the septal OM requires similar interactions with divisome components, or is dependent on peptidoglycan synthesis, remains to be investigated.

In summary, this work has shown for the first time the surface localisation of IcsP, a member on the Omptin family of OM proteases. We have observed that: (i) the distribution of IcsP is masked by LPS Oag in *S. flexneri* 2457T, and (ii) IcsP is concentrated at new poles and at the septum of dividing cells. This distribution of IcsP explains the observed IcsA localisation defect in ⊿*icsP* strains (Egile *et al*, 1997, Shere *et al*, 1997). We have also proposed a model to explain the inheritance of OM proteins IcsP and IcsA through generations of cell division. Finally, unmasking of surface antigens via LPS Oag-depletion may be useful in the study of other minimally exposed OM proteins.

3.8. Article References

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.

Brosius J (1989) Superpolylinkers in cloning and expression vectors. *DNA* 8: 759-777.

Chatterjee S & Rothenberg E (2012) Interaction of bacteriophage l with its *E. coli* receptor, LamB. *Viruses* 4: 3162-3178.

d'Hauteville H, Dufourcq Lagelouse R, Nato F & Sansonetti PJ (1996) Lack of cleavage of IcsA in *Shigella flexneri* causes aberrant movement and allows demonstration of a cross-reactive eukaryotic protein. *Infect Immun* 64: 511-517.

Datsenko KA & Wanner BL (2000) One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc Natl Acad Sci U S A* 97: 6640-6645.

Dworkin J (2009) Cellular polarity in prokaryotic organisms. *Cold Spring Harb Perspect Biol* 1: a003368.

Egan AJ & Vollmer W (2013) The physiology of bacterial cell division. *Ann N Y Acad Sci* 1277: 8-28.

Egile C, d'Hauteville H, Parsot C & Sansonetti PJ (1997) SopA, the outer membrane protease responsible for polar localization of IcsA in *Shigella flexneri*. *Mol Microbiol* 23: 1063-1073.

Fehniger TE, Radolf JD, Walfield AM, Cunningham TM, Miller JN & Lovett MA (1986) Native surface association of a recombinant 38-kilodalton *Treponema pallidum* antigen isolated from the *Escherichia coli* outer membrane. *Infection and immunity* 52: 586-593.

Fukuda I, Suzuki T, Munakata H, Hayashi N, Katayama E, Yoshikawa M & Sasakawa C (1995) Cleavage of *Shigella* surface protein VirG occurs at a specific site, but the secretion is not essential for intracellular spreading. *J Bacteriol* 177: 1719-1726.

Ghosh AS & Young KD (2005) Helical disposition of proteins and lipopolysaccharide in the outer membrane of *Escherichia coli*. *J Bacteriol* 187: 1913-1922.

Gibbs KA, Isaac DD, Xu J, Hendrix RW, Silhavy TJ & Theriot JA (2004) Complex spatial distribution and dynamics of an abundant *Escherichia coli* outer membrane protein, LamB. *Mol Microbiol* 53: 1771-1783.

Goldberg MB & Theriot JA (1995) *Shigella flexneri* surface protein IcsA is sufficient to direct actin-based motility. *Proc Natl Acad Sci U S A* 92: 6572-6576.

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.

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.

Horton RM (1993) In Vitro Recombination and Mutagenesis of DNA (Chapter 25). *Methods in Molecular Biology, Vol 15: PCR Protocols: Current Methods and Applications*, Vol. 15 (White BA, ed.) p.^pp. 251-261. Humana Press Inc., Totowa, NJ.

Horton RM, Cai ZL, Ho SN & Pease LR (1990) Gene splicing by overlap extension: tailor-made genes using the polymerase chain reaction. *Biotechniques* 8: 528-535.

Jain S, van Ulsen P, Benz I, Schmidt MA, Fernandez R, Tommassen J & Goldberg MB (2006) Polar localization of the autotransporter family of large bacterial virulence proteins. *J Bacteriol* 188: 4841-4850.

Kaufmann A, Stierhof YD & Henning U (1994) New outer membrane-associated protease of *Escherichia coli* K-12. *J Bacteriol* 176: 359-367.

LaBrec E, Schneider H, Magnani T & SB F (1964) Epithelial cell penetration as an essential step in the pathogenesis of bacillary dysentery. *J Bacteriol* 88: 1503-1518.

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.

Lugtenberg B, Meijers J, Peters R, van der Hoek P & van Alphen L (1975) Electrophoretic resolution of the "major outer membrane protein" of *Escherichia coli* K12 into four bands. *FEBS Lett* 58: 254-258.

Lutkenhaus J & Addinall SG (1997) Bacterial cell division and the Z ring. *Annu Rev Biochem* 66: 93-116.

Lynne AM, Skyberg JA, Logue CM & Nolan LK (2007) Detection of Iss and Bor on the surface of *Escherichia coli*. *J Appl Microbiol* 102: 660-666.

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.

Morona R & Van Den Bosch L (2003) Lipopolysaccharide O antigen chains mask IcsA (VirG) in *Shigella flexneri*. *FEMS Microbiol Lett* 221: 173-180.

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.

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.

Osborn MJ & Munson R (1974) Separation of the inner (cytoplasmic) and outer membranes of Gram-negative bacteria. *Method Enzymol* 31: 642-653.

Papadopoulos M & 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.

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 Wzz_{pHS2}. *Microbiol-SGM* 154: 1104-1116.

Reeves PR, Hobbs M, Valvano MA, Skurnik M, Whitfield C, Coplin D, Kido N, Klena J, Maskell D, Raetz CR & Rick PD (1996) Bacterial polysaccharide synthesis and gene nomenclature. *Trends Microbiol*4: 495-503.

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.

Rose RE (1988) The nucleotide sequence of pACYC184. *Nucleic Acids* Res 16: 355.

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.

Shere KD, Sallustio S, Manessis A, D'Aversa TG & Goldberg MB (1997) Disruption of IcsP, the major *Shigella* protease that cleaves IcsA, accelerates actin-based motility. *Mol Microbiol* 25: 451-462.

Steinhauer J, Agha R, Pham T, Varga AW & Goldberg MB (1999) The unipolar *Shigella* surface protein IcsA is targeted directly to the bacterial old pole: IcsP cleavage of IcsA occurs over the entire bacterial surface. *Mol Microbiol* 32: 367-377.

Suzuki T & Sasakawa C (2001) Molecular basis of the intracellular spreading of *Shigella*. *Infect Immun* 69: 5959-5966.

Tamaki S, Matsuzawa H & Matsuhashi M (1980) Cluster of *mrdA* and *mrdB* genes responsible for the rod shape and mecillinam sensitivity of *Escherichia coli*. *J Bacteriol* 141: 52-57.

Teh MY, Tran ENH & Morona R (2012) Absence of O antigen suppresses *Shigella flexneri* IcsA autochaperone region mutations. *Microbiol-SGM* 158: 2835-2850.

Typas A, Banzhaf M, van den Berg van Saparoea B, Verheul J, Biboy J, Nichols RJ, Zietek M, Beilharz K, Kannenberg K, von Rechenberg M, Breukink E, den Blaauwen T, Gross CA & Vollmer W (2010) Regulation of peptidoglycan synthesis by outer-membrane proteins. *Cell* 143: 1097-1109.

Ursell TS, Trepagnier EH, Huang KC & Theriot JA (2012) Analysis of surface protein expression reveals the growth pattern of the gram-negative outer membrane. *PLoS Comput Biol* 8: e1002680.

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 der Ley P, De Graaff P & Tommassen J (1986a) Shielding of *Escherichia* coli outer membrane proteins as receptors for bacteriophages and colicins by O-antigenic chains of lipopolysaccharide. *J Bacteriol* 168: 449-451.

Van der Ley P, Kuipers O, Tommassen J & Lugtenberg B (1986b) O-antigenic chains of lipopolysaccharide prevent binding of antibody molecules to an outer membrane pore protein in Enterobacteriaceae. *Microb Pathog* 1: 43-49.

Vandeputte-Rutten L, Kramer RA, Kroon J, Dekker N, Egmond MR & Gros P (2001) Crystal structure of the outer membrane protease OmpT from *Escherichia coli* suggests a novel catalytic site. *Embo J* 20: 5033-5039.

Voorhis DL, Dillon S, Formal SB & Isberg RR (1991) An O antigen can interfere with the function of the *Yersinia pseudotuberculosis* invasin protein. *Mol Microbiol* 5: 317-325.

3.9. Article Supporting Information

3.9.1. Supporting Figures

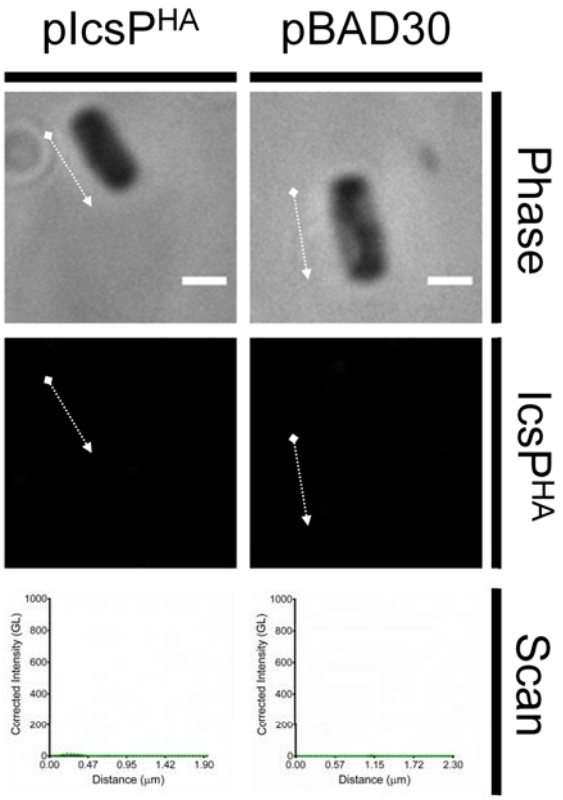


Figure 3.10[S]: Inability to detect IcsP^{HA} on the cell surface of 2457T *icsP-*.

Smooth LPS 2457T *icsP-* strains expressing pIcsP^{HA} (left column) or pBAD30 (right column) were subcultured in LB for 1.5 h to an OD_{600} reading of ~0.4. Cultures were then washed 3 times in LB, induced with 0.03 % (w/v) arabinose for 1 h, fixed, and subjected to QD IF using antibodies for HA epitope and IcsA. Representative bacteria are shown. Scan = Single line-scans measuring the intensity of IcsP^{HA} detected along the surface of the bacterium, Bars = 1 µm, Arrows = direction of line-scan, GL = Grey level, Phase = phase contrast image, IcsP^{HA} = image of fluorescence at 525 nm.

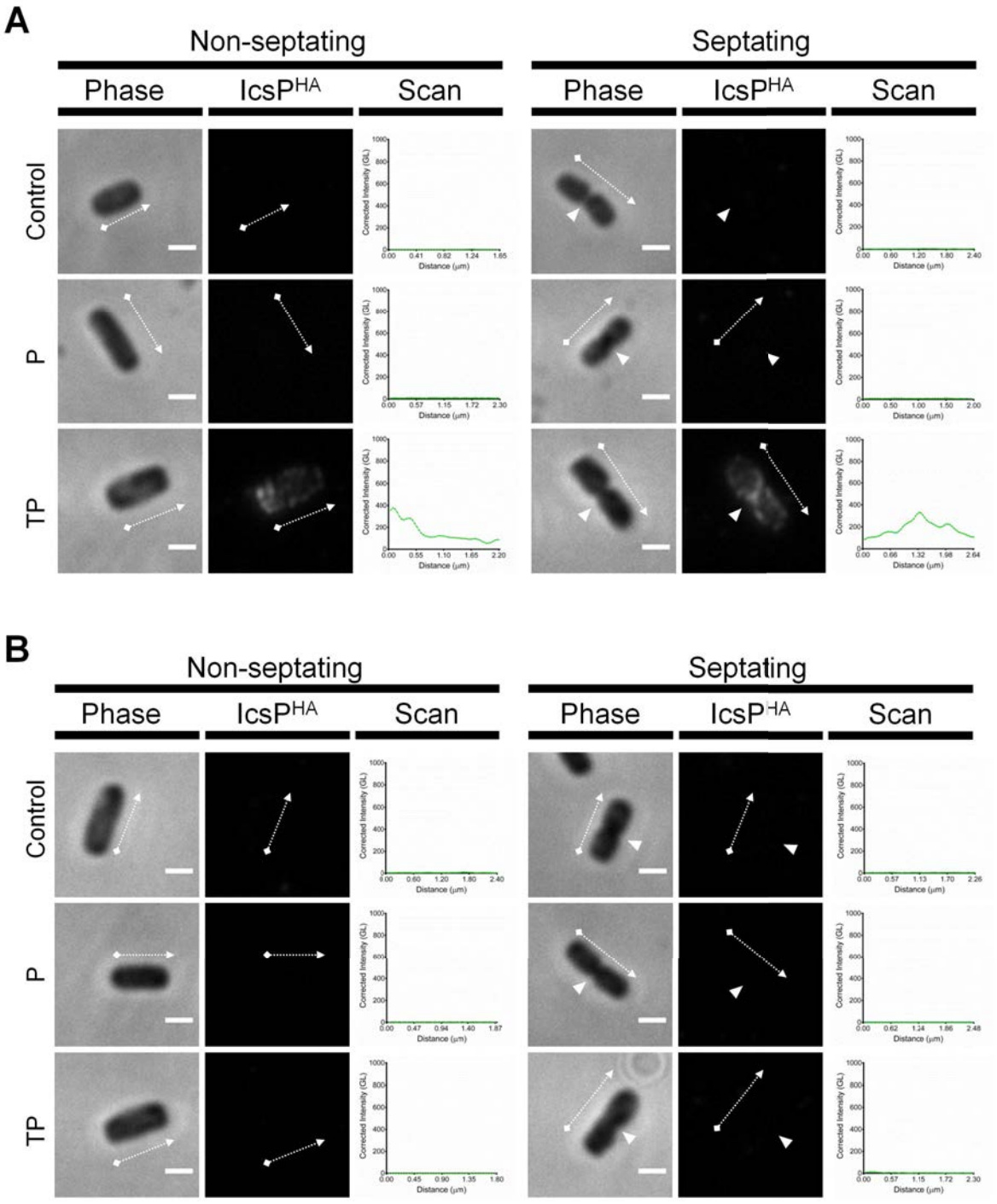


Figure 3.11[S]: Single IcsP^{HA} staining of LPS-depleted 2457T *icsP*-.

Smooth LPS 2457T *icsP*- strains harbouring **(A)** pIcsP^{HA} or **(B)** pBAD30 were subcultured in LB broth for 1.5 h to an OD_{600} reading of ~0.4, washed 3 times in LB, and then further cultured for 2 h in either: the absence of TP, in the presence of PMBN only, or in the presence of TP. Arabinose was included in the final hour of treatment at a concentration of 0.03 % (w/v). Samples were then fixed and subjected to QD IF using antibodies to HA epitope. Representative non-septating and septating cells are shown for each treatment group. Scan=Single line-scans measuring the fluorescence intensity of IcsP^{HA} detected along the surface of the bacterium, Bars=1 μm, Arrows=direction of line-scan, Arrow heads=septa, Control=grown in absence of both tunicamycin and PMBN, P=PMBN, TP=tunicamycin/PMBN, GL=Grey level, Phase=phase contrast image, IcsP^{HA}=image of fluorescence at 525 nm.

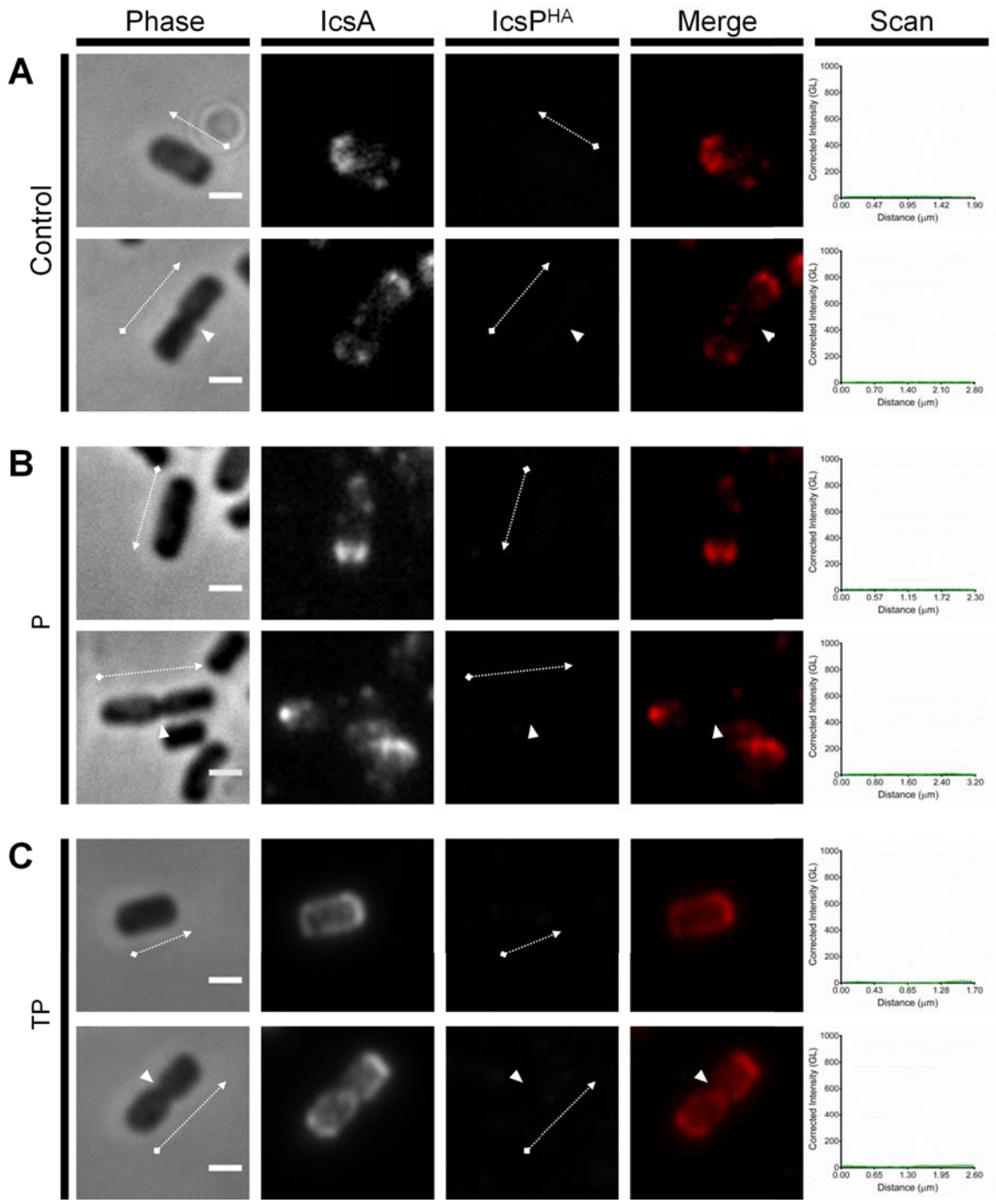


Figure 3.12[S]: Double stained IF of LPS depleted 2457T *icsP*- [pBAD30].

Smooth LPS 2457T *icsP*- strains harbouring pBAD30 were subcultured in LB broth for 1.5 h to an OD_{600} reading of ~0.4, washed 3 times in LB, and then further cultured for 2 h; **(A)** in the absence of TP, **(B)** in the presence of PMBN only, or **(C)** in the presence of TP. Arabinose was included in the final hour of treatment at a concentration of 0.03 % (w/v). Samples were then fixed and subjected to QD IF using antibodies to HA epitope and IcsA. Non-septating and septating cells (upper and lower rows respectively) are shown for each treatment group. Representative bacteria are shown. Scan = Single line-scans measuring the fluorescence intensity of $IcsP^{HA}$ detected along the surface of the bacterium, Bars = 1 μm, Arrows = direction of line-scan, Arrow heads = septa, Control = grown in absence of both tunicamycin and PMBN, P = PMBN, TP = tunicamycin/PMBN, GL = Grey level, Phase = phase contrast image, $IcsP^{HA}$ = image of fluorescence at 525 nm, IcsA = image of fluorescence at 625 nm, Merge = overlay of $IcsP^{HA}$ and IcsA images.

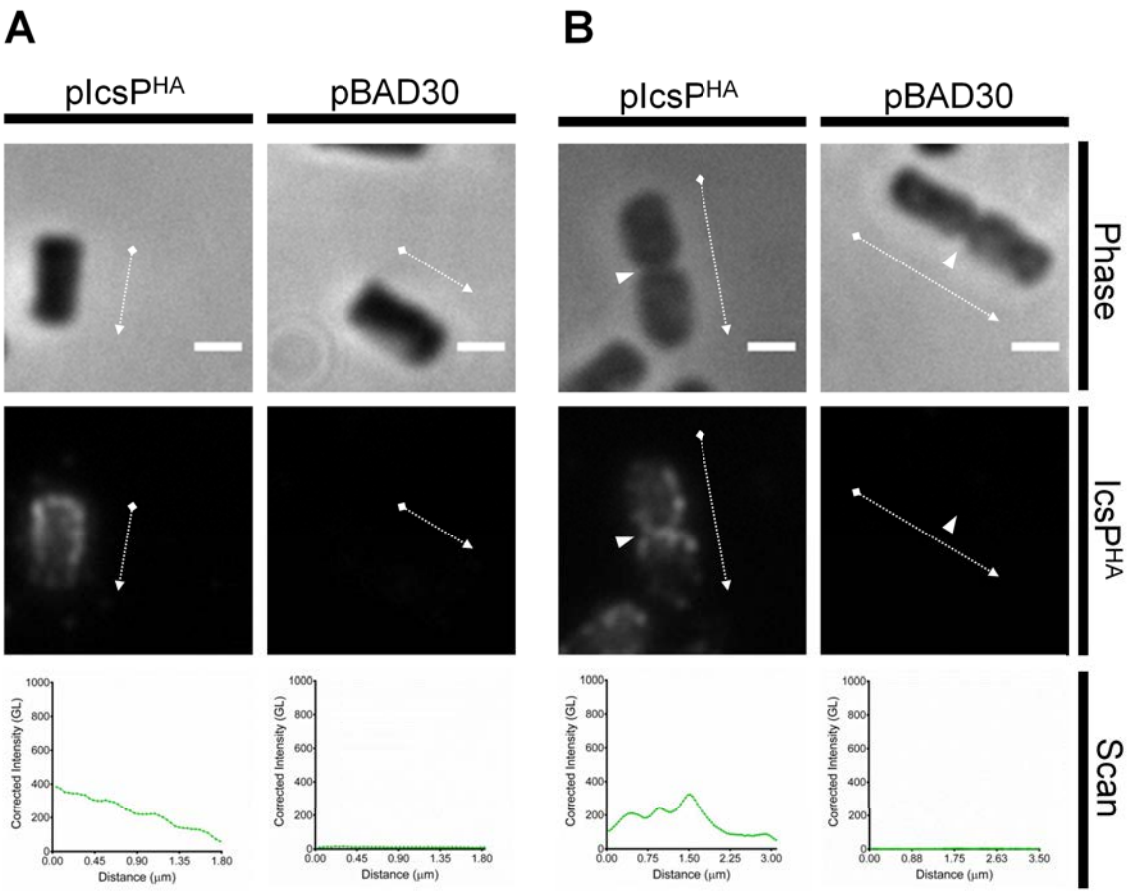


Figure 3.13[S]: Single IcsP^{HA} staining of 2457T *icsP-/rmlD-*.

Single IcsP^{HA} staining of 2457T *icsP*$^{-}$*/rmlD-*. Rough LPS 2457T *icsP*$^{-}$*/rmlD-* strains harbouring pIcsP^{HA} (left columns) or pBAD30 (right columns) were subcultured LB broth for 1.5 h to an OD_{600} reading of ~0.4. Cultures were then washed 3 times in LB, induced with 0.03 % (w/v) arabinose for 1 h, fixed, and subjected to QD IF using antibodies for HA epitope. Non-septating and septating life stages are shown in **A** and **B** respectively. Representative bacteria are shown. Scan = Single line-scans measuring the fluorescence intensity of IcsP^{HA} detected along the surface of the bacterium, Bars = 1 μm, Arrows = direction of line-scan, Arrow heads = septa, GL = Grey level, Phase = phase contrast image, IcsP^{HA} = image of fluorescence at 525 nm.

Chapter Four

ARTICLE 2:

Lipopolysaccharide surface structure does not influence IcsA polarity.

Matthew Thomas Doyle, Marcin Grabowicz, Kerrie Leanne May, and Renato Morona

Chapter 4: *Research Article 2:* Lipopolysaccharide surface structure does not influence IcsA polarity.

4.1. Statement of Authorship

Title of Paper	Lipopolysaccharide surface structure does not influence IcsA polarity
Status	Published
Citation	Doyle MT, Grabowicz M, May KL & Morona R (2015) Lipopolysaccharide surface structure does not influence IcsA polarity. *FEMS Microbiol Lett* 362: fnv042. doi: 10.1093/femsle/fnv042.

Author Contributions: By signing the Statement of Authorship, each author certifies that: (i) the candidate's stated contribution to the publication is accurate, (ii) permission is granted for the candidate in include the publication in the thesis; and (iii) the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

Author	Matthew Thomas Doyle		
Contribution	Conducted all experiments pertaining to Figure 4.3, conceived manuscript theme, constructed all figures (including Figure 4.2) and tables, and wrote the manuscript.		
Signature		Date	28/07/15

Author	Marcin Grabowicz		
Contribution	Construction of *S. flexneri* 2457T *minD* knock-out, conducted all experiments pertaining to Figure 4.2 (minicell/whole-cell isolation, membrane purification, LPS analysis), manuscript editing.		
Signature		Date	28/07/15

Author	Kerrie Leanne May		
Contribution	Preliminary minicell experiments, construction of plasmids for *minD* knock-out, manuscript editing.		
Signature		Date	28/07/15

Author	Renato Morona		
Contribution	Conceived the use of minicells, supervised development of work, provision of laboratory and materials, manuscript evaluation and editing, acting corresponding author.		
Signature		Date	28/07/15

4.2. Purpose of the Article

As previously described in the Introduction, the polarity of IcsA is influenced by multiple factors. The lipopolysaccharide (LPS) of *S. flexneri* is thought to be one of these factors. IcsA polar localisation is seemingly reduced in rough (R)-LPS strains compared to smooth (S)-LPS strains. This has led to the proposal that LPS alterations change membrane fluidity which subsequently alters IcsA stable localisation at the pole. Since, S-LPS *S. flexneri* also have two O-antigen chain modal lengths, there also remained the possibility that shorter chains were localised at the IcsA pole and longer chains were localised laterally in order to 'mask' off-target IcsA placement. Due to the confounding nature of these ideas, the actual influence of LPS on IcsA polarisation remained poorly understood. Hence, this article re-examined the role of LPS surface structure on IcsA polarity.

4.2.1. Thesis Research Aims Addressed:

- **To investigate the maintenance of IcsA polar localisation at the bacterial surface.**

  - To determine whether IcsA present on the surface of *S. flexneri* is masked asymmetrically by lipopolysaccharide O-antigen chains.

  - To ascertain whether there is a difference in IcsA polarity between S-LPS and R-LPS expressing *S. flexneri*.

4.3. Article Abstract

Shigella species are the causative agents of human bacillary dysentery. These bacteria spread within the lining of the gut via a process termed actin-based motility whereby an actin 'tail' is formed at the bacterial pole. The bacterial outer membrane protein IcsA initiates this process, and crucially, is precisely positioned on the bacterial polar surface. Lipopolysaccharide (LPS) O-antigen surface structure has been implicated as an augmenting factor of polarity maintenance due to the apparent dysregulation of IcsA polarity in O-antigen deficient strains. Due to *Shigellae* having long and short O-antigen chains on their surfaces, it has been proposed that O-antigen chain lengths are asymmetrically distributed to optimize IcsA exposure at the pole and mask exposure laterally. Additionally, it has been proposed that LPS O-antigen restricts IcsA diffusion from the pole by maintaining minimal membrane fluidity. This study utilizes minicells and quantitative microscopy providing data refuting the models of asymmetric masking and membrane diffusion, and supporting a model of symmetric masking of IcsA. We contend that IcsA surface distribution is equivalent between wild-type and O-antigen deficient strains, and that differences in cellular IcsA levels have confounded previous conclusions.

4.4. Article Introduction

Shigella species such as *Shigella flexneri* are human specific Gram negative bacterial pathogens that are adapted to the invasion of colonic mucosa leading to dysentery (Niyogi, 2005, Lima *et al*, 2015). The outer membrane autotransporter protein IcsA is essential for intra- and inter-cellular spreading of *S. flexneri* in epithelia via the process of actin-based motility (Bernardini *et al*, 1989, Lett *et al*, 1989, Goldberg & Theriot, 1995, Kocks *et al*, 1995, Egile *et al*, 1999). IcsA is localized to the surface of the old bacterial pole (that which is not derived from the septum of the parent cell) where it binds host cell actin recruiting / polymerizing complexes required for this motility (Egile *et al*, 1999, Steinhauer *et al*, 1999, Snapper *et al*, 2001, Suzuki *et al*, 2002, May & Morona, 2008, Valencia-Gallardo *et al*, 2015). Hence, maintenance of an asymmetrical spatial surface distribution is critical for appropriate functioning of IcsA in all *Shigellae* species. By mechanisms that are yet to be fully elucidated, new IcsA is secreted to the pole after pre-secretion cytoplasmic accumulation (Charles *et al*, 2001, Rokney *et al*, 2009). IcsA surface polarity is also refined by the actions of its specific outer membrane protease IcsP which is localised to the new cell pole and the septa of dividing bacteria (Egile *et al*, 1997, Tran *et al*, 2013). This opposing distribution results in asymmetric IcsA cleavage and refines IcsA surface polarity (Tran *et al*, 2013).

Lipopolysaccharide (LPS) structure has also been implicated as a modulating factor in IcsA biogenesis, polarity, and function. Certainly, *S. flexneri* spreading is abrogated upon changes in LPS structure (Sandlin *et al*, 1995, Sandlin *et al*, 1996, Hong & Payne, 1997, Van Den Bosch *et al*, 1997). However, there is disagreement in the literature concerning the specific mechanisms by which LPS effects IcsA. For instance, immunofluorescence microscopy and immunogold electron microscopy studies have reported that IcsA can be found at increased levels along the lateral surface of rough (R-LPS) *S. flexneri* (strains that lack the O-antigen repeat chain component of LPS), as opposed to the refined polar detection of smooth (S-LPS) wild-type *S. flexneri* (Sandlin *et al*, 1995, Van Den Bosch *et al*, 1997, Robbins *et al*, 2001). In explanation, it was proposed that R-LPS strains have higher membrane fluidity (Figure 4.1A) resulting in easier diffusion of IcsA away from the pole and down the sides of the bacterium (Robbins *et al*, 2001). However, this is confounded by the realization that LPS O-antigen chains mask detection of IcsA by limiting antibody access (Morona *et al*, 2003, Morona & Van Den Bosch, 2003a, Morona & Van Den Bosch, 2003b). Therefore, the refined polar detection of IcsA observed on *S. flexneri* may not be the complete picture of its actual surface localization. Further complicating is that *S. flexneri* decorates its surface with two modal lengths of O-antigen repeats; short type (^{S}LPS; 11-17 repeats) (Morona *et al*, 1995) and

very long type (^{VL}LPS; 90+ repeats) (Hong & Payne, 1997) which are regulated by the $WzzB_{SF}$ and $WzzB_{pHS2}$ inner membrane co-polymerases respectively (Morona *et al*, 1995, Stevenson *et al*, 1995, Hong & Payne, 1997). It has been hypothesized on multiple occasions that *S. flexneri* has two types of O-antigen modal lengths to counteract the steric hindrance effect of LPS, whilst retaining protection from host defences and colicins (Morona *et al*, 2003, Pugsley & Buddelmeijer, 2004, Scribano *et al*, 2014, Tran *et al*, 2014). In this model, ^{VL}LPS is required for serum resistance, whereas ^{S}LPS minimizes IcsA masking at the pole such that it can access external actin recruiting complexes (Figure 4.1B).

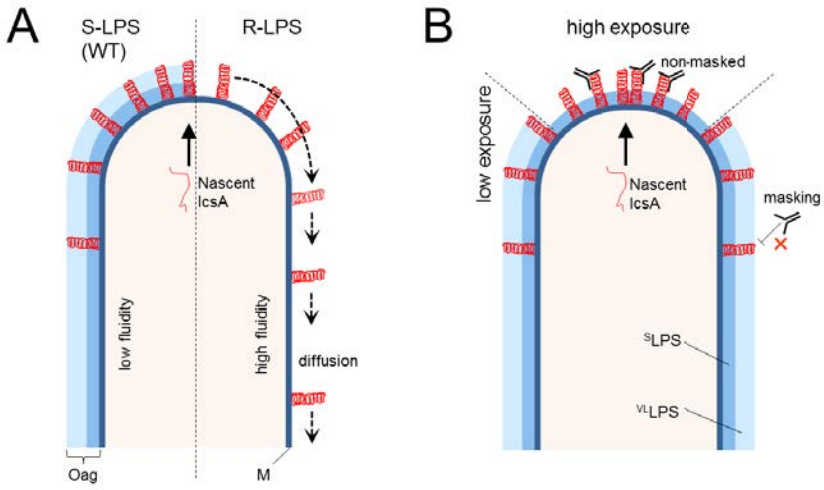


Figure 4.1: Models of IcsA surface polarity augmentation by LPS.

(A) Lateral diffusion model. It has been proposed that R-LPS *Shigellae* (without O-antigen; Oag) has a higher membrane (M) fluidity causing IcsA (red) deposited at the pole to defuse away from the pole and down the lateral edge. **(B)** Model of asymmetrical masking due to O-antigen modal lengths distribution. To optimize the pathogenic role of IcsA in recruiting host actin polymerizing complexes at the pole, it is thought that LPS O-antigen chain lengths may be useful in optimizing IcsA exposure at the pole by the use of ^{S}LPS (medium blue) at the pole and ^{VL}LPS (light blue) on the lateral edges to restrict IcsA exposure. The red cross depicts the notion that LPS O-antigen chains can inhibit access of antibodies to IcsA via steric hindrance.

Due to the confounding nature of these models (masking, lateral diffusion / membrane fluidity, asymmetric O-antigen chain lengths), the exact effects of LPS on IcsA surface localisation remains enigmatic. This work unravels the IcsA-LPS relationship in *S. flexneri* by first examining whether LPS O-antigen modal chain lengths are asymmetrically distributed in the outer membrane. IcsA localizations in the rough and wild-type membrane are then quantified and directly compared allowing a re-evaluation of the asymmetrical masking and lateral diffusion models. The results obtained challenge current thoughts concerning the LPS-IcsA relationship and provide further insights into IcsA polar positioning.

4.5. Article Materials and Methods

4.5.1. Bacterial strains, plasmids and culture.

Lists of strains and plasmids utilized in this study are included in Table 4.1. *S. flexneri* colonies were grown on Congo Red agar for confirmation of virulence plasmid presence before routine growth in Luria-Bertani (LB) media at 37 °C with shaking. Unless otherwise stated, bacteria were sub-cultured to a log-phase OD_{600} reading of 0.5 before experimental use. When required, broths were supplemented with the following additives at respective concentrations; tetracycline (10 μg mL^{-1}), kanamycin (50 μg mL^{-1}), and ampicillin (50 μg mL^{-1}).

Table 4.1: Strains and plasmids used in this study.

Strain or Plasmid	Description	LPS	Source
Strain			
2457T	Wild-type *S. flexneri* 2a	S	
RMA2041	2457T ⊿*icsA*::Tc^R	S	(Van den Bosch & Morona, 2003)
RMA2043	2457T ⊿*icsA*::Tc^R ⊿*rmlD*::Km^R	R	(Van den Bosch & Morona, 2003)
ETRM230	2457T ⊿*rmlD*::Km^R	R	(Tran *et al*, 2013)
MG292	2457T *minD*::Km^R	S	This study
Plasmids			
pBR322	medium copy number, colE1 *ori*, Ap^R, Tc^R		(Bolivar *et al*, 1977)
pIcsA	pBR322 derivative containing cloned *icsA* gene, P_{IcsA} promoter, Ap^R		(Morona & Van Den Bosch, 2003a)
pKD4	FLP Km^R template		(Datsenko & Wanner, 2000)
pKD46	λ red recombinase, Ap^R		(Datsenko & Wanner, 2000)
pKMRM96	pGEMT::*minCDE*, Ap^R		This study
pKMRM161	pGEMT::*minCD*::Km^R*E*, Ap^R		This study

Tc^R = tetracycline resistance, Km^R = kanamycin resistance, Ap^R = ampicillin resistance, S = S-LPS, R = R-LPS

4.5.2. Construction of *minD* mutant.

The *minCDE* locus of *S. flexneri* 2457T was PCR amplified using oligonucleotides minF (gacttgcctcaatataatcc) and minR (tctgtgcgtgggaacagc) that anneal to nt positions 1210181-1210200 and 1208137-1208154 respectively on the 2457T chromosome (Wei *et al*, 2003). The amplicon was cloned into pGEMT-Easy (Promega) creating pKMRM96 (Table 4.1). To disrupt the *minD* gene, the kanamycin resistance (Km^R) cassette from pKD4 (Datsenko & Wanner, 2000) was amplified using P1PacI (ccttaattaagtgtaggctggagctgcttc) and

P2PacI (ccttaattaacatatgaatatcctccttag) incorporating flanking PacI sites which were used to insert the Km^R cassette into the native PacI site within the *minD* gene in pKMRM96 resulting in pKMRM161 (Table 4.1). The *min* locus containing disrupted *minD::Km*^{R} was then amplified using minF/R and the amplicon used in recombineering mutagenesis of 2457T *minD* genomic copy via the λ red recombinase system (Datsenko & Wanner, 2000).

4.5.3. Antibodies.

Polyclonal rabbit anti-IcsA (passenger), rabbit anti-$WzzB_{SF}$, and rabbit anti-$WzzB_{pHS2}$ were produced and validated as described previously (Van Den Bosch *et al*, 1997, Daniels & Morona, 1999, Purins *et al*, 2008). Mouse anti-DnaK monoclonal antibody was from Enzo Life Sciences.

4.5.4. Total bacterial protein samples.

1:50 sub-cultures were grown to log-phase. 5 x 10^8 of log-phase bacteria were collected by centrifugation (16,000 x *g*, 1 min, 4 °C), resuspended in 100 µL of SDS-PAGE loading buffer (Lugtenberg *et al*, 1975), and heated to 100 °C for 10 min before SDS-PAGE and immunoblot analysis.

4.5.5. Bacterial IcsA labelling.

Immunofluorescence (IF) staining was conducted essentially as described previously (Tran *et al*, 2013). All solutions were filtered through a 0.2 µm nitrocellulose filter. 10^8 log-phase bacteria were harvested from a 1:50 sub-culture by centrifugation (16,000 x *g*, 2 min, 20 °C), resuspended in 3.7 % (v/v) formaldehyde solution (Sigma) in phosphate buffered saline (PBS), and incubated at 20 °C for 20 min. Fixed bacteria were washed twice in PBS before resuspension in 100 µL of PBS. 5 µL of the bacteria were spotted onto sterile round coverslips (at the bottom of a 24-well tray) that were pre-treated with 10 % (v/v) poly-L-lysine solution (Sigma) in PBS. Bacteria were centrifuged (775 x *g*, 5 min, 20 °C) and then incubated for 2 h with anti-IcsA diluted 1:100 in PBS containing 10 % (v/v) fetal calf serum (FCS). Bacteria were washed three times with PBS and then incubated for 30 min at 37 °C with donkey anti-rabbit Alexa Fluor 488 antibody (Invitrogen) diluted 1:100 in PBS containing 10 % (v/v) fetal calf serum (FCS). Bacteria were washed three more times with PBS before mounting with 20 % Mowiol 4-88 (Calbiochem), 4 mg mL^{-1} *p*-phenylenediamine.

4.5.6. Minicell and whole-cell purification.

Separation of minicells and whole-cells was conducted as described previously (Achtman *et al*, 1979). The minicell strain was sub-cultured (1:20) until log-phase, or sub-cultured for 16 h to produce stationary phase cultures. A volume of 250 mL bacteria from both log-phase and stationary-phase cultures were pelleted by centrifugation (8,600 x *g*, 20 min, 4 °C;) and washed in 10 mL of buffered saline gelatin (BSG; 0.85 % (w/v) NaCl, 0.03 % (w/v) KH_2PO_4, 0.06 % (w/v) Na_2HPO_4, 100 μg/mL gelatin). Bacteria were pelleted again (20,400 x *g*, 8 min, 4 °C) and resuspended in 2 mL of BSG. Bacteria were layered onto sucrose gradients and centrifuged (3,300 x *g*, 30 min, 4 °C). The minicell fraction in the middle of the tube was extracted using a syringe. The whole-cell fraction at the bottom of the tube was also collected and diluted in 50 mM Tris pH 7.5. The minicells were pelleted (20,400 x *g*, 8 min, 4 °C), resuspended in 1 mL of BSG and purified once more on a sucrose gradient as described. The minicells were then re-pelleted (as above) and resuspended in 2 mL of 50 mM Tris pH 7.5. Cell concentrations were normalised on the basis that an OD_{600} = 1.0 represents 5 x 10^8 whole cells and 2 x 10^9 minicells.

4.5.7. Minicell and whole-cell membrane protein and LPS analysis.

As described previously (Achtman *et al*, 1979), purified minicells and whole-cells were lysed by sonication in 20 mM Tris-HCl pH 8.0, 10 mM NaCl buffer containing 0.1 mg mL^{-1} DNase, 0.1 mg mL^{-1} RNase, and 0.1 mM phenylmethanesulfonyl fluoride. Unbroken cells were removed by centrifugation (5,500 x *g*, 25 min, 4 °C) and the lysate was ultracentrifuged (100,000 x *g*, 60 min, 4 °C). The whole membrane pellet was rinsed with buffer, homogenised in 20 mM Tris-HCl pH 8.0, 10 mM NaCl buffer containing 1 % (v/v) SDS, and incubated on ice for 1 h. This was then ultracentrifuged (as above) and the resulting supernatant collected. Protein content was assessed using a BCA Protein Estimation assay (Pierce). Membrane samples from minicells and whole-cells were standardised to equivalent total membrane protein concentration for protein analysis by immunoblot. For LPS analysis, samples were treated with 0.5 mg mL^{-1} Proteinase K in SDS-PAGE loading buffer at 56 °C for 16 h and analysed by SDS-PAGE and silver stain.

4.5.8. Microscopy and quantitation.

All images of IF labelled bacteria were captured using an Olympus IX-7 Microscope and MetaMorph software (Molecular Devices) with a phase contrast 100 x oil

immersion objective and a 1.5 x enlarger. For fluorescence imaging an X-Cite 120Q lamp was used set at high intensity. All live bacterial imaging was conducted on custom made 1 % (w/v) agarose-LB solid media mounts with 37 °C incubation. All bacterial IcsA fluorescence images were acquired with 100 millisecond exposures. Fluorescence images for background correction were taken for each experiment. IcsA fluorescence images for presentation were recoloured using the ICA LUT in ImageJ such that the full intensity spectrum can be easily observed. MetaMorph line-scan measurement tools were used to quantitate fluorescence intensities across the perpendicular axis of a point-to-point scan. Scans were conducted from pole-to-pole starting from intense pole, with scan width (perpendicular axis) equal to the bacterium (approx. 20 pixels). For each strain under investigation, cumulative scans were conducted of many bacteria (50 bacteria from each independent experiment '*n*') that were without a visible septum, resulting in distribution profiles representative of the population.

Any asymmetry in LPS O-antigen chain lengths would dramatically change the apparent IcsA polarity between S-LPS and R-LPS strains and may allow increased exposure of IcsA at the pole. To investigate LPS asymmetry we constructed an *S. flexneri minD*- strain (MG292; Table 4.1). MinD (along with MinC and MinE) regulates appropriate positioning for septum formation in bacterial division (Treuner-Lange & Sogaard-Andersen, 2014). Mutants in this system form minicells that result from mislocalized septation at the poles (de Boer *et al*, 1989). As such, minicells are rich in polar membrane material compared to whole-cells and have been vital for investigations on the polar cytology (Koppelman *et al*, 2001, Lai *et al*, 2004). The *minD*- strain behaved as expected with the formation of free minicells and observed polar budding of minicells (Figure 4.2A). We then purified both whole-cell and minicell populations of this strain based on density and assessed purity microscopically. The whole-cell fraction was 98.9 % pure (one budding minicell observed per 94 bacteria), and whole-cells were not observed in the minicell fraction (Figure 4.2B).

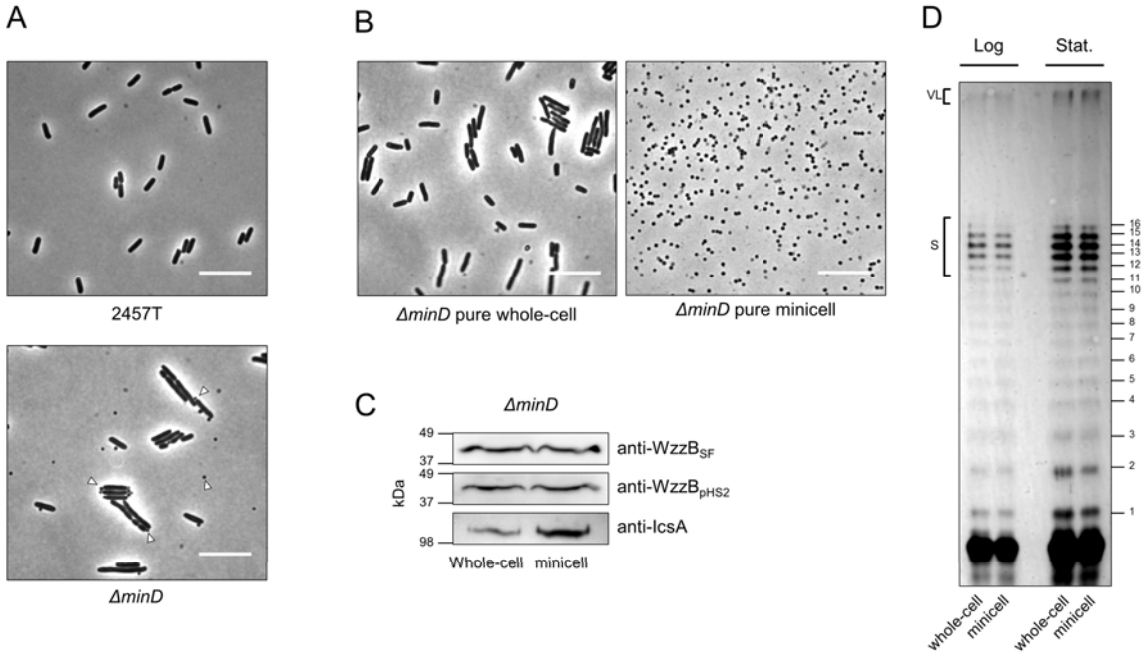


Figure 4.2: Lateral and polar LPS has equivalent O-antigen modal lengths.

(A) Phase micrographs of live *S. flexneri* showing wild type (2457T, top) and *minD*- phenotype (MG292, bottom). Arrowheads in the latter indicate free minicells and minicells budding from whole-cell poles. All scale bars represent 10 μm. **(B)** Phase micrographs of purified whole-cells and minicells from the *minD*- strain. **(C)** Western immunoblot analysis of standardized whole membrane samples extracted from purified whole-cells and minicells. Levels of both LPS-Oag modal length modulators WzzB$_{SF}$ and WzzB$_{pHS2}$ were assessed. Anti-IcsA served as a control that minicells were derived from polar material. **(D)** LPS was isolated from the standardized whole membrane samples of whole-cells and minicells from both log-phase and stationary phase cultures and analyzed by SDS-PAGE and silver staining. S = ^{S}LPS, VL = ^{VL}LPS.

Upon assessment of extracted membrane protein (Figure 4.2C), we observed no discernible difference between whole-cells and minicells in the abundance of O-antigen chain length modulators $WzzB_{SF}$ and $WzzB_{pHS2}$. As expected, minicell membranes were more abundant in IcsA than whole-cells showing that minicells represent polar material of the IcsA pole. Additionally, we also observed no differences in the relative abundances of ^{S}LPS and ^{VL}LPS between minicells and whole-cells. This was true for purified populations from both log-phase and stationary phase cultures (Figure 4.2D). Therefore, these results do not support a model of enhancement of IcsA exposure at the pole due to an asymmetric distribution of LPS O-antigen chain lengths between the pole and lateral surfaces (Figure 4.1B). Consequently, the previously observed changes in apparent IcsA distributions between S-LPS and R-LPS bacteria must be due to one or more of the effects of symmetrical masking, membrane fluidity and lateral diffusion, or other factors. It should also be noted here that, to our knowledge, this is the first observation of LPS O-antigen modal length distribution using minicells.

To thoroughly model IcsA distributions and the effects of LPS, we devised methods to quantitate the average IcsA surface population distribution for a given strain removing biases of qualitative assessment and artificial selection of bacteria (see *Materials and Methods*). We first investigated IcsA differences between the wild-type and R-LPS derivative strains (Figure 4.3A, B, and C). Our R-LPS strain is unable to make O-antigen due to the absence of RmlD which synthesizes dTDP-rhamnose (a precursor sugar for O-antigen synthesis, see Table 4.1). Unexpectedly, we observed a large increase in IcsA levels in the R-LPS strain relative wild-type (Figure 4.3A) which had not previously been reported. However, qualitative IcsA surface distributions replicated previous reports with the R-LPS strain displaying higher lateral and bipolar IcsA detection compared to wild-type (Figure 4.3B). We quantified these distributions (Figure 4.3C) and found that IcsA surface detection was significantly more intense for the R-LPS strain (Figure 4.3C_i; $P = 0.0002$), yet was still highly localized to the old pole (Figure 4.3C_{ii}). Direct comparisons of S-LPS and R-LPS IcsA distributions (Figure 4.3C_{iii}) revealed that the R-LPS strain had significantly higher placement of IcsA at the new pole, whether assessed relative to the old pole or the mid-cell ($P = 0.0053$ and $P < 0.0001$ respectively). There was no significant change in IcsA old pole localization relative the mid-cell between S-LPS and R-LPS strains.

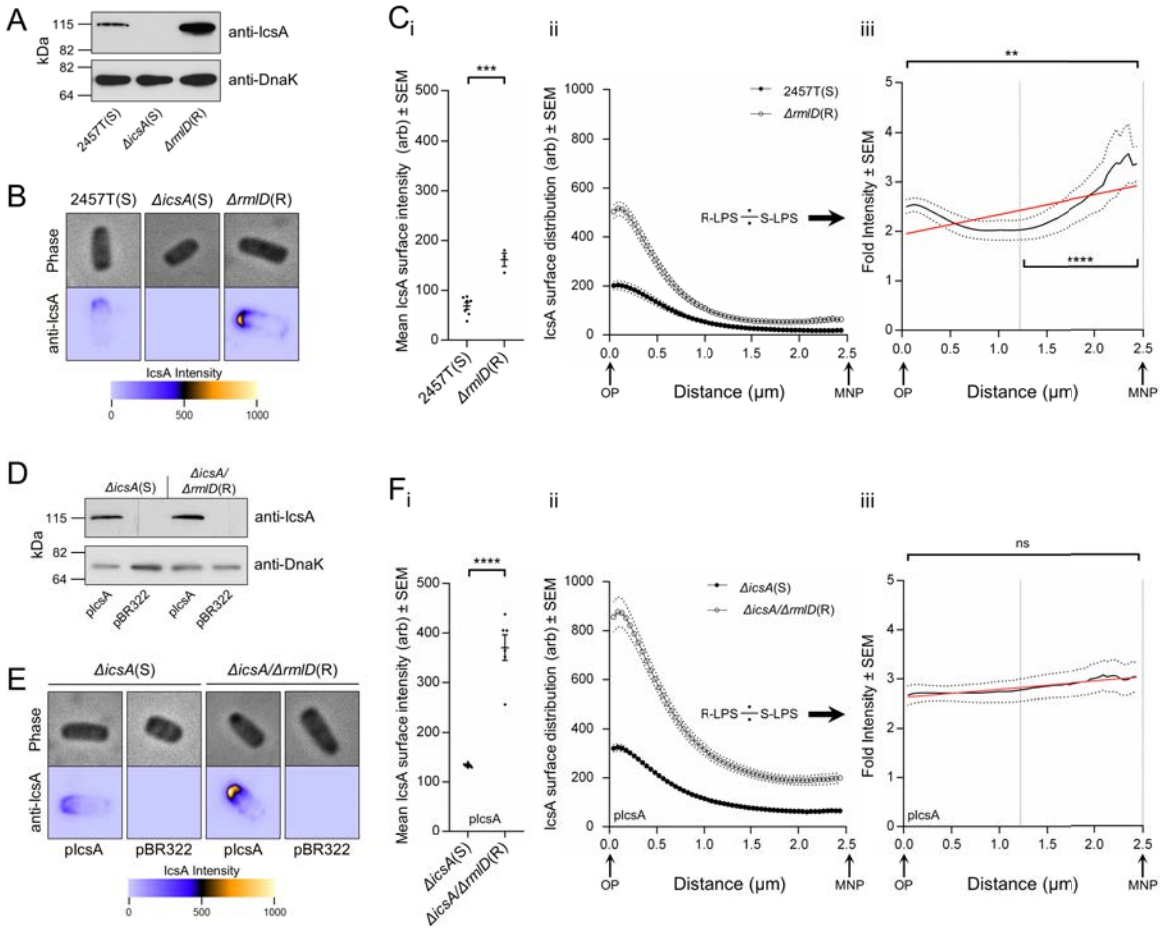


Figure 4.3: Removal of LPS O-antigen does not change IcsA surface distribution.

IcsA expression levels and surface distributions were investigated in both single *icsA* copy and multi-*icsA* copy conditions. Panels **(A – C)** show data generated using wild type *S. flexneri* 2457T, ⊿*icsA* (RMA2041), and ⊿*rmlD* (ETRM230) strains, and panels **(D – F)** from strains ⊿*icsA* and ⊿*icsA* ⊿*rmlD* (RMA2043) complemented with either pIcsA or base vector pBR322 (Bolivar *et al*, 1977) (see Table 1). 'S' and 'R' denote smooth-LPS (with O-antigen) and rough-LPS (without O-antigen) respectively. Anti-IcsA Western immunoblots **(A and D)** show IcsA protein expression levels in total bacterial protein samples (n = 3). Chaperone DnaK served as a loading control. **(B and E)** Phase (top) and anti-IcsA IF micrographs (bottom) of representative bacteria. Fluorescence intensities for panels **C and F** are average pixel grey levels scaled equally relative to each strain. Each image is 4 μm by 4 μm. **(C and F)** IF experiments were repeated (n = 3-7) and IcsA surface detection **(i)** and surface distributions **(ii)** were measured for each IcsA expressing strain on a population basis. The quotients of the R-LPS and S-LPS IcsA distributions are also shown in **(iii)** with mean mid-cell indicated by the vertical line and red line indicating fitted linear functions (R^2 = 0.3974 and 0.8924 for $\mathbf{C_{iii}}$ and $\mathbf{F_{iii}}$ respectively). OP = old pole, MNP = mean new pole, ns = not significant. Differences in mean surface detection for **(i)** were analyzed by two-tailed t-test, and differences in distribution between OP, MNP and mid-cell in **(iii)** analyzed by one-way ANOVA (** = $P < 0.01$, *** = $P < 0.001$, **** = $P < 0.0001$).

These data support previous reports that R-LPS strains have an increased propensity for bipolarity and a reduction in polar refinement, yet it is difficult to assess whether this is due to the increase in overall IcsA expression or due to changes in membrane diffusion of IcsA. Therefore we repeated this investigation using strains expressing IcsA from a plasmid (pIcsA; see Table 4.1). These conditions equalized IcsA levels between S-LPS and

R-LPS strains as shown (Figure 4.3D). Qualitatively, IcsA surface distributions on R-LPS bacteria again appeared more intense than S-LPS, but had similar overall distributions (Figure 4.3E). This was recapitulated when quantitated (Figure 4.3F_i and F_{ii}), but unexpectedly, the quotient of these distributions (Figure 4.3F_{iii}) did not show any significant shifts in IcsA localization for any point between the poles.

Contrary to the current literature, the results presented in Figure 4.3 show that upon IcsA cellular levels being equal, IcsA surface distribution remains indistinguishable regardless of the presence of LPS O-antigen on the membrane. This supports the notion that the masking effect of LPS is exerted symmetrically over the surface of *S. flexneri*, and is further supported by our observations of equivalent O-antigen chain lengths between whole and minicells (indicating symmetrical chain length distributions for wild type) presented in Figure 4.2. Furthermore, since LPS changes do not affect IcsA polarity, it also shows that R-LPS does not consequently increase the fluidity of IcsA molecules in the outer membrane (Figure 4.1A). Lateral diffusion of IcsA from the pole is either unchanged or does not occur.

It is also interesting that IcsA levels are increased when O-antigen synthesis is blocked. Although previously utilized *S. flexneri* strains were deficient in O-antigen due to varied mutations (Sandlin *et al*, 1995, Sandlin *et al*, 1996, Robbins *et al*, 2001), it is possible that previous attributions of LPS effecting IcsA polarity were due to overlooked changes in cellular IcsA concentration. The reason for this change in IcsA level is uncertain but it is plausible that degradases responsible for normal IcsA turnover are functionally altered in R-LPS strains resulting in higher steady-state levels. Indeed, we have previously shown that periplasmic protease DegP has altered activities with respect to IcsA maintenance in R-LPS *S. flexneri* (Teh *et al*, 2012a). Nevertheless, it is intriguing that increased IcsA levels increase the tendency for abnormal placement of IcsA at the new pole. It has been proffered that cytoplasmic accumulation at the pole seeds initial placement of IcsA (Charles *et al*, 2001, Rokney *et al*, 2009) – it is possible that changes in IcsA abundance can influence this accumulation and increase the tendency for off target accumulation. This notion is consistent with the increases of IcsA at the new pole observed in this work (Figure 4.3C).

In summary, this study reveals that; (i) *S. flexneri* IcsA polarity, and any diffusion of IcsA in the outer membrane, is **not** affected by LPS O-antigen presence, (ii), IcsA is affected by symmetrical masking, (iii) O-antigen chain lengths are symmetrically distributed, and (iv) changes in O-antigen synthesis can deregulate IcsA levels effecting polarity.

4.7. Article Acknowledgements

MTD is the recipient of a Doctor of Philosophy scholarship from the University of Adelaide. We thank the Research Centre for Infectious Diseases (RCID) for support during this work. We also thank Elizabeth Ngoc Hoa Tran for critical reading of the manuscript.

4.8. Article References

Achtman M, Manning PA, Edelbluth C & Herrlich P (1979) Export without proteolytic processing of inner and outer-membrane proteins encoded by F-Sex factor tra cistrons in *Escherichia coli* minicells. *P Natl Acad Sci USA* 76: 4837-4841.

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.

Charles M, Perez M, Kobil JH & Goldberg MB (2001) Polar targeting of *Shigella* virulence factor IcsA in *Enterobacteriacae* and *Vibrio*. *Proc Natl Acad Sci U S A* 98: 9871-9876.

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.

atsenko KA & Wanner BL (2000) One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc Natl Acad Sci U S A* 97: 6640-6645.

de Boer PA, Crossley RE & Rothfield LI (1989) A division inhibitor and a topological specificity factor coded for by the minicell locus determine proper placement of the division septum in *E. coli*. *Cell* 56: 641-649.

Egile C, d'Hauteville H, Parsot C & Sansonetti PJ (1997) SopA, the outer membrane protease responsible for polar localization of IcsA in *Shigella flexneri*. *Mol Microbiol* 23: 1063-1073.

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.

Goldberg MB & Theriot JA (1995) *Shigella flexneri* surface protein IcsA is sufficient to direct actin-based motility. *Proc Natl Acad Sci U S A* 92: 6572-6576.

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.

Kocks C, Marchand JB, Gouin E, d'Hauteville H, Sansonetti PJ, Carlier MF & Cossart P (1995) The unrelated surface proteins ActA of *Listeria monocytogenes* and IcsA of *Shigella flexneri* are sufficient to confer actin-based motility on *Listeria innocua* and *Escherichia coli* respectively. *Mol Microbiol* 18: 413-423.

Koppelman CM, Den Blaauwen T, Duursma MC, Heeren RM & Nanninga N (2001) *Escherichia coli* minicell membranes are enriched in cardiolipin. *J Bacteriol* 183: 6144-6147.

Lai EM, Nair U, Phadke ND & Maddock JR (2004) Proteomic screening and identification of differentially distributed membrane proteins in *Escherichia coli*. *Mol Microbiol* 52: 1029-1044.

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.

Lima IF, Havt A & Lima AA (2015) Update on molecular epidemiology of *Shigella* infection. *Curr Opin Gastroen* 31: 30-37.

Lugtenberg B, Meijers J, Peters R, van der Hoek P & van Alphen L (1975) Electrophoretic resolution of the "major outer membrane protein" of *Escherichia coli* K12 into four bands. *FEBS Lett* 58: 254-258.

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.

Morona R & Van Den Bosch L (2003a) Lipopolysaccharide O antigen chains mask IcsA (VirG) in *Shigella flexneri*. *FEMS Microbiol Lett* 221: 173-180.

Morona R & Van Den Bosch L (2003b) 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 & Manning PA (1995) Molecular, genetic, and topological characterization of O-antigen chain length regulation in Shigella flexneri. *J Bacteriol* 177: 1059-1068.

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

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

Pugsley AP & Buddelmeijer N (2004) Traffic spotting: poles apart. *Mol Microbiol* 53: 1559-1562.

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. *Microbiol-SGM* 154: 1104-1116.

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.

Rokney A, Shagan M, Kessel M, Smith Y, Rosenshine I & Oppenheim AB (2009) *E. coli* transports aggregated proteins to the poles by a specific and energy-dependent process. *J Mol Biol* 392: 589-601.

Sandlin RC, Goldberg MB & Maurelli AT (1996) Effect of O side-chain length and composition on the virulence of *Shigella flexneri* 2a. *Mol Microbiol* 22: 63-73.

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.

Scribano D, Petrucca A, Pompili M*, et al.* (2014) Polar localization of PhoN2, a periplasmic virulence-associated factor of *Shigella flexneri*, is required for proper IcsA exposition at the old bacterial pole. *PloS One* 9(2): e90230

Snapper SB, Takeshima F, Anton I*, et al.* (2001) N-WASP deficiency reveals distinct pathways for cell surface projections and microbial actin-based motility. *Nat Cell Biol* 3: 897-904.

Steinhauer J, Agha R, Pham T, Varga AW & Goldberg MB (1999) The unipolar *Shigella* surface protein IcsA is targeted directly to the bacterial old pole: IcsP cleavage of IcsA occurs over the entire bacterial surface. *Mol Microbiol* 32: 367-377.

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.

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.

Teh MY, Tran EN & Morona R (2012) Absence of O antigen suppresses *Shigella flexneri* IcsA autochaperone region mutations. *Microbiology* 158: 2835-2850.

Tran EN, Doyle MT & Morona R (2013) LPS unmasking of *Shigella flexneri* reveals preferential localisation of tagged outer membrane protease IcsP to septa and new poles. *PloS One* 8(7): e70508.

Tran ENH, Papadopoulos M & Morona R (2014) Relationship between O-antigen chain length and resistance to colicin E2 in *Shigella flexneri*. *Microbiol-SGM* 160: 589-601.

Treuner-Lange A & Sogaard-Andersen L (2014) Regulation of cell polarity in bacteria. *J Cell Biol* 206: 7-17.

Valencia-Gallardo CM, Carayol N & Tran Van Nhieu G (2014) Cytoskeletal mechanics during *Shigella* invasion and dissemination in epithelial cells. *Cell Microbiol* 10.1111/cmi.12400.

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

Wei J, Goldberg MB, Burland V*, et al.* (2003) Complete genome sequence and comparative genomics of *Shigella flexneri* serotype 2a strain 2457T. *Infect Immun* 71: 2775-2786.

Chapter Five

ARTICLE 3:

The passenger-associated transport repeat promotes virulence factor secretion efficiency and delineates a distinct autotransporter subtype.

Matthew Thomas Doyle, Elizabeth Ngoc Hoa Tran and Renato Morona

Chapter 5: *Research Article 3:* The passenger-associated transport repeat promotes virulence factor secretion efficiency and delineates a distinct autotransporter subtype.

5.1. Statement of Authorship

Title of Paper	The passenger-associated transport repeat promotes virulence factor secretion efficiency and delineates a distinct autotransporter subtype.
Status	Published
Citation	Doyle MT, Tran EN & Morona R (2015) The passenger-associated transport repeat promotes virulence factor secretion efficiency and delineates a distinct autotransporter subtype. *Mol Microbiol* 97(2):315-329. doi: 10.1111/mmi.13027.

Author Contributions: By signing the Statement of Authorship, each author certifies that: (i) the candidate's stated contribution to the publication is accurate, (ii) permission is granted for the candidate in include the publication in the thesis; and (iii) the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

Author	Matthew Thomas Doyle		
Contribution	Conducted all experiments pertaining to all figures, conceived manuscript theme, constructed all figures, tables, and supplementary, and wrote the manuscript.		
Signature		Date	28/07/15

Author	Elizabeth Ngoc Hoa Tran		
Contribution	Constructed *S. flexneri* 2457T *icsA/icsP* double knock-out, manuscript editing.		
Signature		Date	28/07/15

Author	Renato Morona		
Contribution	Supervised development of work, provision of laboratory and materials, manuscript evaluation and editing, acting corresponding author.		
Signature		Date	28/07/15

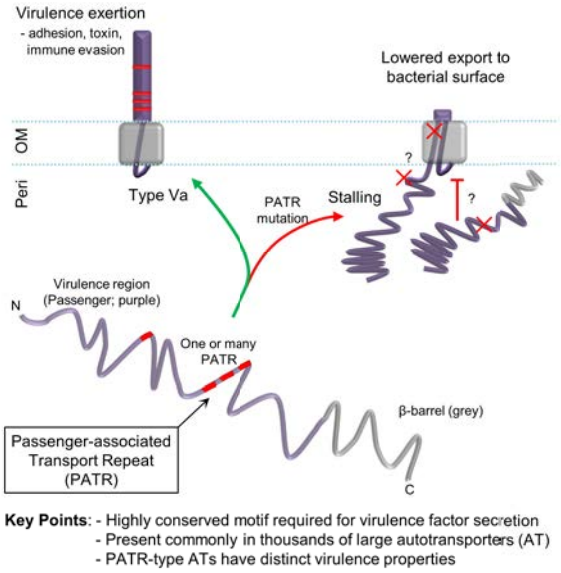
5.2. Purpose of the Article

As described in the Introduction, conserved autotransporter passenger elements (such as the Pertactin-like regions) can play key roles in the overall efficiency of secretion. During my studies I discovered that the IcsA passenger amino acid sequence contains a very highly conserved repeat motif which was designated by various protein domain databases as an 'autotransporter β-repeat'. The importance of this so-called autotransporter β-repeat for IcsA was completely unknown and the function of the repeat was completely uncharacterised for any other protein in the literature. As such, this article reflects a concerted effort to characterise the function of this repeat in autotransporter biogenesis (using IcsA as the model), and to uncover the wider use of the repeat in other autotransporter virulence factors.

5.2.1. Thesis Research Aims Addressed:

- **To investigate the role of conserved passenger elements in IcsA autotransport.**
  - To determine the function of the uncharacterised 'autotransporter β-repeat' sequence $IcsA_{526-557}$.
  - To explore the wider importance of the repeat sequence in other autotransporters.

5.3. Article Abbreviated Summary and Graphical Abstract



In this work we uncover the importance of a conserved but uncharacterized Autotransporter (AT) motif that we term the Passenger-associated Transport Repeat (PATR). The PATR is required for efficient AT secretion and is present in thousands of large ATs. PATR-type ATs have distinct architectures and virulence traits making them an important subgroup of ATs secreted by Gram-negative pathogens.

5.4. Article Summary

Autotransporters are a superfamily of virulence factors secreted by Gram negative bacteria. They are comprised of an N-terminal passenger domain that is translocated across the outer membrane, and a C-terminal domain that inserts into the outer membrane forming a β-barrel anchor. It is still poorly understood how the passenger is efficiently translocated in the absence of external energy inputs. Several mechanisms have been proposed in solution of this problem, yet due to the vast diversity of size, sequence, and function of the passenger, it is not clear how widely these mechanisms are employed. In this study we functionally characterize a conserved repeat found in many passengers which we designate the Passenger-associated Transport Repeat (PATR). Using the autotransporter IcsA from the enteropathogen *Shigella flexneri*, we identified conserved PATR residues that are required for efficient export of the passenger during growth and infection. Furthermore, PATR-containing autotransporters are significantly larger than non-PATR autotransporters, with PATR copy number correlating with passenger size. We also show that PATR-containing autotransporters delineate a subgroup that associates with specific virulence traits and architectures. These results advance our understanding of autotransporter composition, and indicate that an additional transport mechanism is important for thousands of these proteins.

5.5. Article Introduction

Gram negative bacteria coordinate infection and disease via a synergy of secreted virulence factors. As such, the secretion of these virulence factors across the double-membrane cell wall must be efficient for pathogenic fitness. The autotransporter (AT, or Type Va) secretion pathway is the most common solution to this problem (Henderson *et al*, 2004, Dautin & Bernstein, 2007, Leyton *et al*, 2012, Grijpstra *et al*, 2013). AT superfamily proteins have distinctive domain architecture; (i) an N-terminal signal sequence for Sec mediated passage of the inner membrane, (ii) a central passenger domain harboring the virulence properties of the protein, and (iii) a transmembrane β-barrel at the C-terminus. The β-barrel is required for the stages of outer membrane (OM) insertion, and the translocation of the passenger to the extracellular space. Depending on the virulence function of the AT, the passenger can remain attached to the bacterial surface or be released into the extra-bacterial milieu. Also, due to its functional diversity the passenger varies widely in its sequence and size (Celik *et al*, 2012).

Despite rigorous investigation, the exact mechanism of OM β-barrel insertion and passenger translocation remains incompletely understood. In general, the β-barrel and passenger are inserted and translocated sequentially by a series of intricate events coordinated by the essential Barrel Assembly Machinery (BAM) (Jain & Goldberg, 2007, Rossiter *et al*, 2011, Roman-Hernandez *et al*, 2014) and by the β-barrel itself (Pavlova *et al*, 2013, Leyton *et al*, 2014). The BAM (a complex of integral and lipoproteins) interacts with both the β-barrel (Ieva *et al*, 2011) and passenger (Ieva & Bernstein, 2009) portions of the AT to assist β-barrel insertion. Structural analysis suggests that the local distortion of the OM at the lateral gate of the BamA component facilitates the seeding of the nascent β-barrel into the membrane (Noinaj *et al*, 2013, Noinaj *et al*, 2014). The passenger is then translocated in a C- to N-terminal manner though the nascent β-barrel pore where it commonly folds into a β-helical stalk (Kajava & Steven, 2006, Junker *et al*, 2009).

The energy source for passenger translocation also remains under debate. In the absence of external energy sources such as ATP, it has been proposed that C- to N-terminal β-helical folding could drive passenger translocation (Junker *et al*, 2006, Braselmann & Clark, 2012). In this model, the initially translocated C-terminal portion of the passenger has high stability and folds with high efficiency (Junker *et al*, 2006, Renn & Clark, 2008, Peterson *et al*, 2010). This folding action itself pulls the rest of the less stable N-terminal portion of the passenger through the β-barrel pore in a vectorial fashion. This type of sequential folding,

although attractive in its simplicity, has only been implicated in a few model ATs (such as Petactin, EspP, and Pet (Renn & Clark, 2008, Junker *et al*, 2009, Peterson *et al*, 2010, Kang'ethe & Bernstein, 2013a) and it is unknown how widely this model is adhered to by the AT superfamily which is large and diverse. Indeed, the AT YapV varies from the model as it lacks a C-terminal (Pertactin-like, PL) stable region but employs an unstable region at the extreme N-terminus for its secretion (Besingi *et al*, 2013). Furthermore, it has been recognized that the vast majority of ATs do not contain a PL region (Drobnak *et al*, 2015b) and the model may not apply to ATs that possess globular passengers (for instance EstA (Van den Berg, 2010)). It has also been realized that passenger domains frequently have a net negative charge that may act to facilitate translocation, possibly by charge repulsion between the passenger and LPS molecules (Kang'ethe & Bernstein, 2013b). Consequently, it remains possible that mechanisms of passenger secretion are 'mixed-and-matched' depending on the size, fold-type, or function of the AT in question.

To further our understanding of AT secretion, we investigated a highly conserved, AT-associated, 32 amino acid repeat. Although models of this repeat had been deposited into both TIGRFAM (Haft *et al*, 2003) (TIGR02601) and Pfam (Finn *et al*, 2014b) (PF12951) databases, its function remained completely uncharacterized. This study aimed to investigate the importance of this motif within the AT family. We reveal that the repeat is: (i) associated with the passengers of a large and distinct group of ATs, (ii) required for efficient transport of the passenger to the bacterial surface, and (iii) connected with certain passenger domain architectures and functions. As such, in this study we refer to this repeat as the Passenger-associated Transport Repeat (PATR).

5.6.1. PATR mutation disrupts steady-state passenger surface presentation.

To investigate the function of the PATR in ATs, the well-studied AT IcsA was used as a model protein. IcsA is an essential virulence factor for the human enteric pathogen *Shigella flexneri* that enables spreading and lesion formation in infected intestinal epithelia (Makino *et al*, 1986, Bernardini *et al*, 1989, Lett *et al*, 1989, Goldberg & Theriot, 1995, Suzuki *et al*, 2002). It also has an uninvestigated single copy of the PATR in its passenger domain at IcsA$^{526\text{-}557}$ (see schematic Figure 5.1A) (Dai *et al*, 2010). On inspection of the Pfam (PF12951) PATR model (see Figure 5.1B), the occurrence of four highly conserved glycines was observed at positions G^{6}, G^{8}, G^{20}, and G^{27} that are present in IcsA (IcsA^{G531}, IcsA^{G533}, IcsA^{G545}, and IcsA^{G552} respectively). We hypothesized that the PATR glycine residues are important in biogenesis, and accordingly, glycine to alanine substitutions at these four sites were constructed in IcsA, along with a complete 32 amino acid deletion of the PATR. These mutants were constructed in a plasmid-borne *icsA* with a native P_{IcsA} promoter (Morona & Van Den Bosch, 2003a) and expressed in an IcsA and O-antigen deficient strain of *S. flexneri* (Van den Bosch & Morona, 2003) (see Table 5.1[S]). This allowed unhindered detection of IcsA surface levels (Morona *et al*, 2003, Morona & Van Den Bosch, 2003a, Morona & Van Den Bosch, 2003b) making strains suitable for the following experiments (see further).

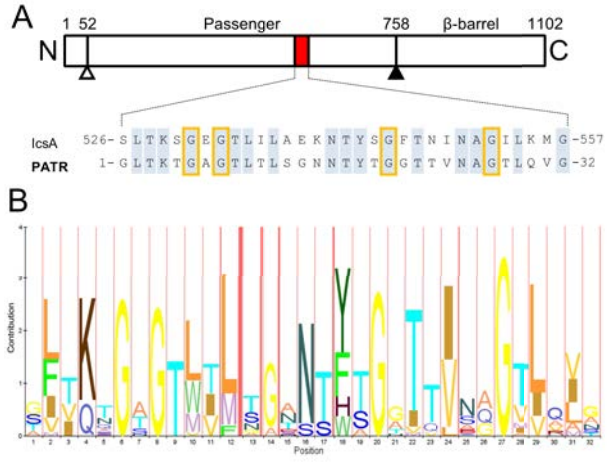


Figure 5.1: The passenger of IcsA has a single conserved PATR.

A scaled schematic of the AT IcsA (Q7BCK4) is shown **(A)** indicating the signal sequence (IcsA$^{1\text{-}52}$) cleaved at the open arrow, the passenger (IcsA$^{53\text{-}758}$), and the β-barrel (IcsA$^{759\text{-}1102}$). The black arrow indicates the site of specific low efficiency cleavage by IcsP. The passenger has a single copy of the PATR (IcsA$^{526\text{-}557}$, red) shown aligned with the PATR sequence (PF12951 consensus) with the PATR Hidden Markov Model in **(B)** (pfam.xfam.org/family/PF12951) (Schuster-Bockler *et al*, 2004). The positions of four conserved glycines are outlined (orange) and are completely conserved in the model (yellow).

First, to assess overall IcsA PATR mutant expression levels, total *S. flexneri* protein samples were analyzed by anti-IcsA (passenger) Western immunoblotting (Figure 5.2A). No substantial difference in total IcsA protein expression was observed, regardless of PATR mutation. Anti-IcsA immunofluorescence (IF) staining of these bacteria was also conducted (Figure 5.2B). Contrary to total protein levels, detectable passenger surface levels were visually reduced for IcsA PATR mutants relative to the wild-type protein. This was particularly true for $IcsA^{G531A}$, $IcsA^{G545A}$, $IcsA^{G552A}$ substitutions, and $IcsA^{\Delta PATR}$. IcsA fluorescence intensities of 250 bacteria (n = 5; 50 bacteria per experiment) were then measured for each PATR mutant and the wild-type protein (Figure 5.2C). This confirmed that $IcsA^{G531A}$, $IcsA^{G545A}$, $IcsA^{G552A}$, and $IcsA^{\Delta PATR}$ displayed significantly ($\alpha = 0.05$) reduced surface levels (mean fractions of wild-type 0.58 ±0.20, 0.61 ±0.31, 0.61 ±0.17, and 0.62 ±0.08 respectively). Although this experiment could not establish significance for $IcsA^{G533A}$ ($P = 0.588$), the trend of reduced surface detection was still observable.

The reduction of detectable surface exposed passenger due to PATR glycine substitutions or deletion may have been due to impaired transfer to the OM. To test this, OM protein (OMP) samples were extracted from *S. flexneri* expressing IcsA and IcsA PATR mutants using differential detergent treatment (see *Experimental Procedures*). Coomassie Blue staining of SDS-PAGE separated OMP samples showed equivalent loading and excellent enrichment of major OMPs relative to a total cell protein control (Figure 5.2D). OMP enrichment was further shown by anti-BamA immunoblotting. Fractionation purity controls showed that periplasmic SurA, inner membrane Wzz, and cytoplasmic DnaK did not contaminate the OMP samples. Interestingly, there was no discernible difference in the amount of full length IcsA in the OM for any of the IcsA PATR mutants relative to the wild-type protein. What was noticeable however, was an increase in OM-associated IcsA degradation products due to PATR mutation. This was observed for all IcsA PATR substitutions and deletion, but most clearly observed for $IcsA^{G545A}$.

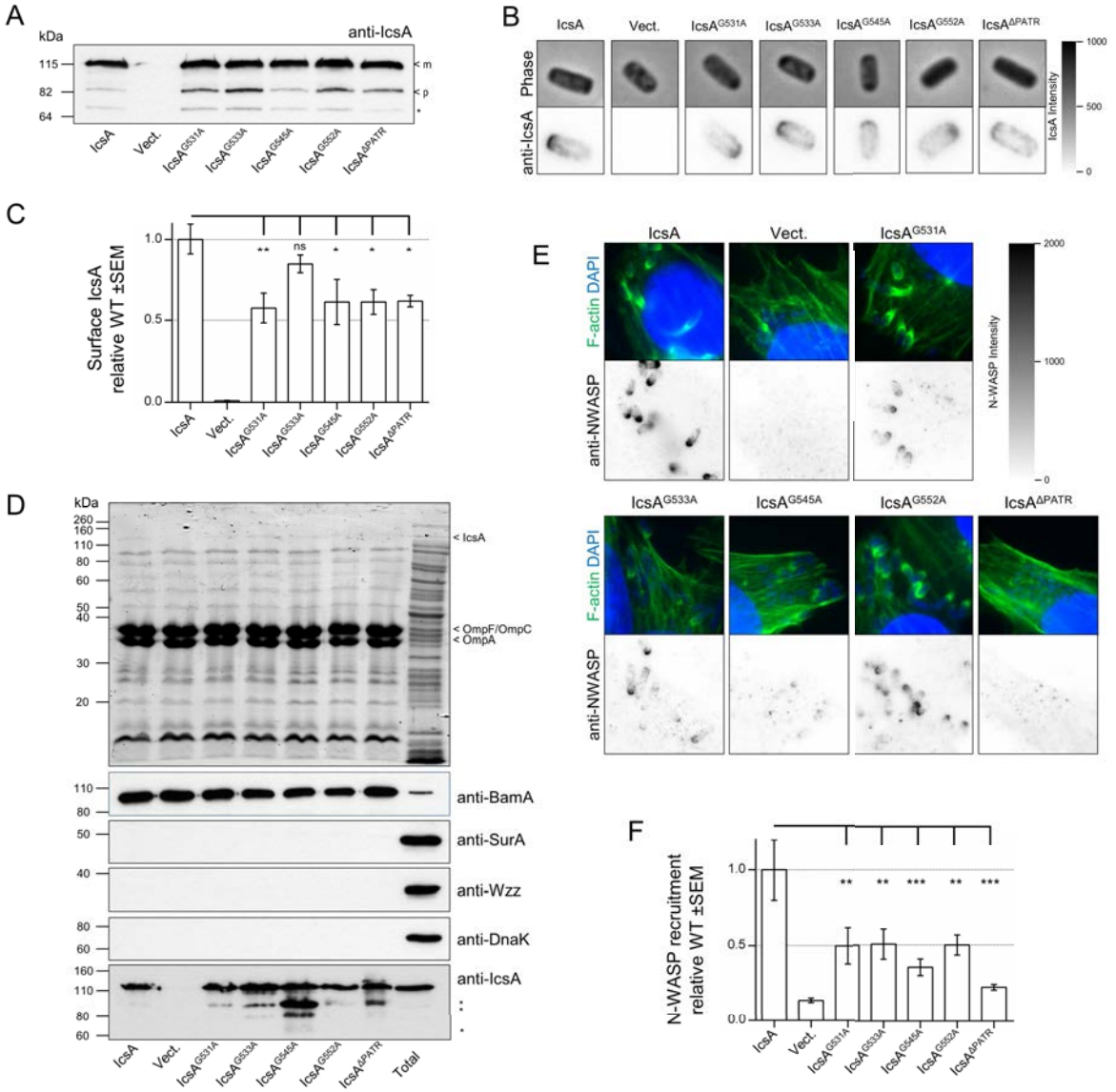


Figure 5.2: PATR mutations decrease surface presentation of IcsA.

(A) An anti-IcsA immunoblot of total cell samples from log-phase *S. flexneri* expressing IcsA and IcsA-PATR mutants. Mature (m), cleaved passenger (p), and common degraded form (*) are indicated. **(B)** The same bacteria subjected to anti-IcsA immunofluorescence microscopy (IFM). Representative bacteria are shown in phase (top) and fluorescence (bottom) images (4 x 4 μm). IcsA fluorescence was quantitated in **(C)** for n = 5 (where 50 bacteria were measured for each experiment) and analyzed by ordinary one-way ANOVA (Dunnett's, $\alpha = 0.05$). **(D)** OM protein was also extracted from these bacteria using sarkosyl and analyzed by Coomassie Blue staining and immunoblotting. Coomassie staining shows equivalent loading and enrichment of major OMPs. BamA serves as both a positive control for OMP enrichment and a loading control. SurA, Wzz, and DnaK serve as periplasmic, inner membrane and cytoplasmic controls respectively. Total = total bacterial protein sample of *S. flexneri* expressing IcsA. * = degraded IcsA products. **(E)** To indirectly assess intracellular IcsA surface levels, N-WASP recruitment and F-actin accumulation was also tested in infected HeLa cells by IFM. Overlay images are shown (top) for bacterial nucleoids and eukaryotic nuclei detected with DAPI (blue) and actin labelled with phalloidin (green). N-WASP fluorescence images are shown below (20 x 20 μm). N-WASP levels were also quantitated in **(F)** for n = 5 (where all bacteria were measured per infected cell for each experiment) and analyzed by ordinary one-way ANOVA (Dunnett's, $\alpha = 0.05$). All experiments were conducted using an IcsA and O-antigen deficient strain of *S. flexneri* (RMA2043) expressing IcsA and IcsA-PATR mutants from P_{IcsA} (see Table 5.1[S]). All fluorescence images are

scaled equally relative to each other. WT = wild type, SEM = standard error of the mean, ns = not significant, * = $P < 0.05$, ** = $P < 0.01$, *** = $P < 0.001$.

The reduction in detectable IcsA on the *S. flexneri* surface due to either PATR glycine substitutions or deletion (Figure 5.2B and C), despite equivalent cellular expression and OM localization (Figure 5.2A and D), indicated a defect in the appropriate assembly steps required for efficient transport of the passenger to the surface. This may also explain the increase in the degraded forms of IcsA in the OM (Figure 5.2D). If maturation of the IcsA passenger to the surface was reduced due to PATR glycine substitutions or deletion, then this would also be observed by a reduction of N-WASP recruitment by the bacterium upon infection of epithelial cells. Recruitment of host N-WASP is the intracellular function of IcsA that results in the formation of actin-based tails required for motility and pathogenicity (Egile *et al*, 1999, Suzuki *et al*, 2002). To test this, cultured HeLa cells were infected with *S. flexneri* expressing either the wild-type or PATR mutant forms of IcsA and stained for N-WASP, filamentous-actin, and DNA. As expected, wild-type IcsA recruited high levels of N-WASP resulting in commonly observed tail filaments (Figure 5.2E). However, dramatic reductions in N-WASP recruitment was observed for substitutions $IcsA^{G531A}$, $IcsA^{G533A}$, $IcsA^{G545A}$, and $IcsA^{G552A}$, and was completely abolished for $IcsA^{\Delta PATR}$. To confirm these observations, the N-WASP fluorescence intensity was measured for all bacteria per infected cell (n = 5) for each strain expressing the PATR mutants and the wild-type protein (Figure 5.2F). This confirmed that $IcsA^{G531A}$, $IcsA^{G533A}$, $IcsA^{G545A}$, $IcsA^{G552A}$, and $IcsA^{\Delta PATR}$ all resulted in significant ($\alpha = 0.05$) reductions in N-WASP recruitment (mean fractions of wild-type 0.50 ±0.12, 0.51 ±0.10, 0.35 ±0.06, 0.50 ±0.07, and 0.22 ±0.02 respectively). Also, N-WASP recruitment for $IcsA^{G545A}$ and $IcsA^{\Delta PATR}$ was not significantly different to the vector control ($P = 0.174$ and $P = 0.873$ respectively).

5.6.2. PATR mutation decreases the efficiency of passenger transport to the bacterial surface.

Thus far, a role for the PATR in steady-state passenger surface presentation, in the context of both bacterial culture and infection, had been established. To more closely investigate the defects in passenger transport to the cell surface, the same IcsA PATR glycine substitutions and deletion were constructed in the plasmid pBADIcsA which has the *icsA* gene controlled by the P_{BAD} promoter (Guzman *et al*, 1995). Expression of these constructs in an IcsP protease deficient strain of *S. flexneri* (see Table 5.1[S]) allowed reduced endogenous

proteolysis permitting informative limited protease shaving and pulse-chase protease accessibility assays (see further) to be performed (Figure 5.3).

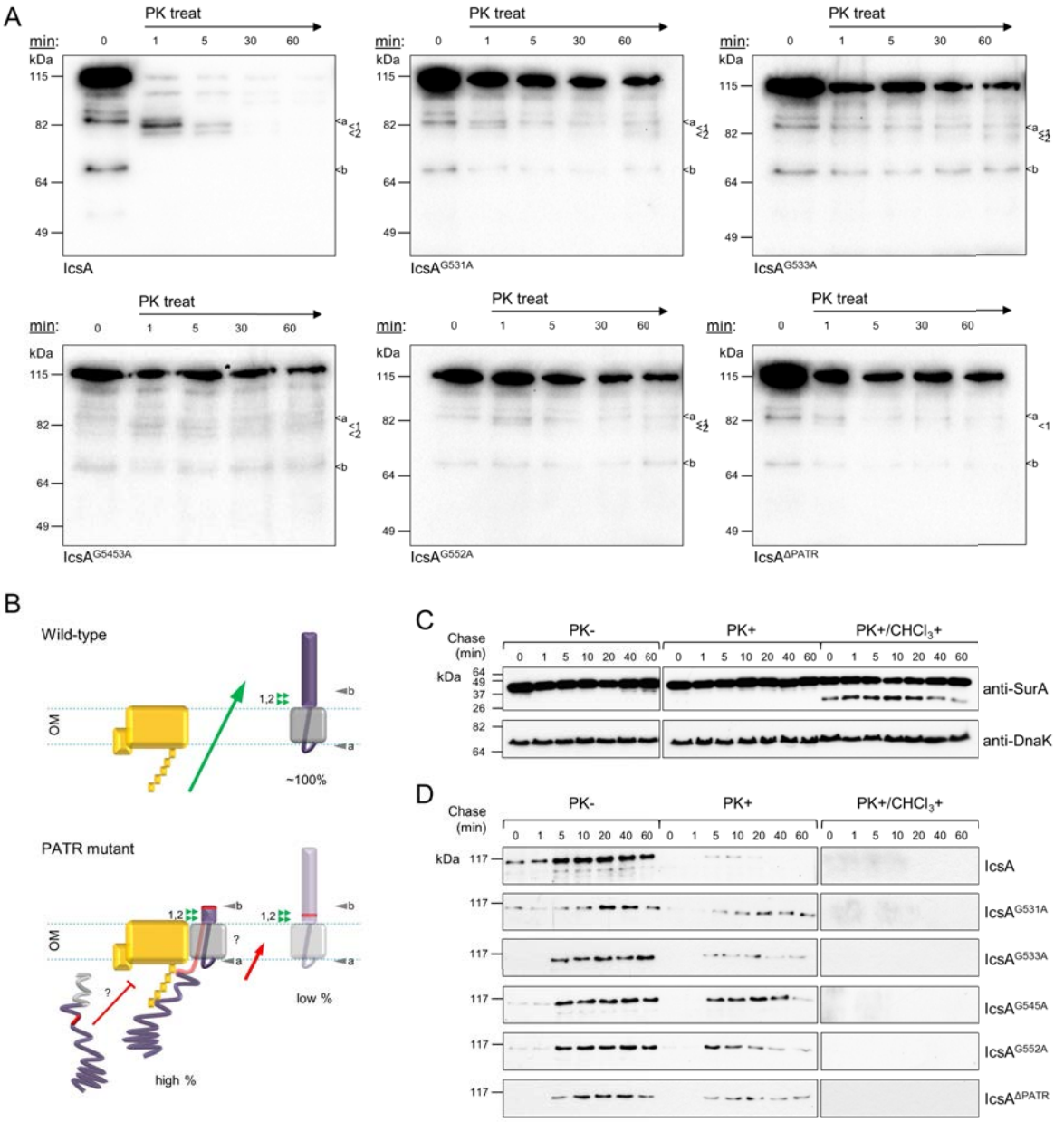


Figure 5.3: PATR mutations decrease the efficiency of IcsA passenger export.

Passenger translocation was investigated by limited protease shaving **(A)** and pulse-chase proteolysis assays **(D)** on live IcsA / IcsP deficient *S. flexneri* (RMA4378) expressing IcsA and IcsA-PATR mutants from an arabinose / glucose P_{BAD} switch (see Table 5.1S). **(A)** Bacteria with overexpressed IcsA and IcsA PATR mutant protein were treated with 1 μg mL^{-1} of Proteinase K (PK) over an hour time-course (n = 2). Positions of fragments resulting from PK shaving are marked as '1' and '2' and pre-treatment IcsA degradation fragments are marked as 'a' and 'b'. **(B)** A model showing possible effects in explanation of the results in **A,** and also previous experiments. The majority of wild-type IcsA is efficiently transported and folded (indicated by green arrow) to the cell surface and is accessible by PK. IcsA PATR mutants stall during secretion due to inappropriate assembly with, or processing by, export machinery required for translocation initiation (for instance BAM; yellow) and/or due to slow passenger transport across the OM. This inefficient transport is represented by the small red arrow. Stalling results in a high proportion of passenger 'stuck' in the periplasmic topology and a low

proportion of surface exposed passenger accessible to PK. Plausible sites of degradation that can result in fragments observed in **A** are similarly marked as '1' and '2' for PK fragments and 'a' and 'b' for pre-treatment fragments. Red = mutated PATR motif, purple = passenger, and grey = β-barrel. **(C)** A mock chase with pBAD30 only. 60 minute chase time-courses are shown where bacteria were treated with Proteinase K (PK+), PK and chloroform (PK+/$CHCl_3$+), or not treated (PK-). Immunoblot of periplasmic SurA shows proteolysis occurring only after OM permeabilization by $CHCl_3$ treatment. Immunoblot of cytoplasmic DnaK indicates treatments did not result in cytoplasmic protein proteolysis. **(D)** Passenger export was chased for IcsA and IcsA-PATR mutants (n = 2). Samples were immunoblotted using anti-IcsA (passenger). Estimations of band intensities and statistical analysis are presented in Figure 5.7[S].

For limited protease shaving experiments, IcsA and IcsA PATR mutants were overexpressed and then bacteria were subjected to low concentrations of Proteinase K (PK) over an hour time-course (see *Experimental Procedures*) (Figure 5.3A). The ~115 kDa band corresponding to wild-type IcsA passenger and β-barrel was highly accessible to PK degradation with the majority of the protein being digested by the first minute and appears as two detectable fragments (<80 kDa) as marked (Figure 5.3A, '1' and '2'). This indicates that the majority of IcsA is efficiently translocated to the surface as expected. However, $IcsA^{G531A}$, $IcsA^{G533A}$, $IcsA^{G545A}$, $IcsA^{G552A}$, and $IcsA^{\Delta PATR}$ all displayed dramatically reduced protease accessibility compared to wild-type IcsA as evidenced by the persistence of the full protein bands (~115 kDa) over the hour time-course. This shows that the majority of these mutant proteins remain un-translocated. Still, a minority of the IcsA PATR mutant proteins are accessible to PK since there is small reduction in full length protein band intensity over time compared to the 0 minute time point, and similar PK fragments (<80 kDa) are produced corresponding to the sizes of the PK fragments produced for wild-type IcsA. These PK fragment bands are faint which can be expected if less translocated passenger is available at the surface for digestion. It is also interesting that overexpressed IcsA and IcsA PATR mutants both displayed low levels of degraded products of equivalent sizes (Figure 5.3A, 'a' and 'b') which were present pre-PK treatment. These fragments were accessible to PK in the wild-type scenario as they were not present after 1 minute of proteolysis. However, for the IcsA PATR mutants, these fragments remained inaccessible to PK digestion over the entire time-course. This experiment, along with experiments described in Figure 5.2, strongly suggests that residues of the conserved PATR motif are required for efficient transport of the passenger to the bacterial surface by some mechanism (Figure 5.3B, top). When the PATR is mutated (Figure 5.3B, bottom), it is plausible that appropriate AT transport is disrupted at one or more points essential for maintenance of export competence resulting in a high proportion of transiting ATs still in the periplasm, and a low proportion that have successfully crossed the membrane. Finally, the similarity in PK fragments between IcsA and IcsA PATR mutants

shows that mutant proteins have similar intrinsic protease susceptibilities and likely adopt a conformation similar to wild-type once translocated.

To further investigate the notion that mutations of the conserved residues of the PATR stalls efficient passenger secretion, IcsA expression was momentarily pulsed by the addition of arabinose in culture, and the newly synthesized protein was chased by sampling over an hour time-course (see *Experimental Procedures*). *S. flexneri* samples were either untreated (PK-) or treated with excess Proteinase K (PK+), allowing an assessment of the rate of passenger transport to the cell surface due to protease accessibility. As a further control, samples were treated with chloroform to permeabilize the OM and allow periplasmic access to PK. Correct topological proteolysis was established in a mock pulse-chase (Figure 5.3C). This confirmed that digestion of periplasmic SurA only occurred after OM permeabilization. Additionally, the cytoplasmic protein DnaK was not affected by any treatment indicating that the inner membrane remained impermeable to PK under the conditions of this experiment. The protease accessibility chase for wild-type IcsA revealed an extremely fast rate of passenger export (Figure 5.3D) with the majority of IcsA being digested within the first 1 to 5 minutes and the remainder still being transported (as indicated by complete digestion after OM permeabilization). Wild-type IcsA was totally translocated after 20 minutes. This rate of secretion is expected since it is consistent with the life-cycle of rapidly dividing *S. flexneri* that require sufficient surface exposed IcsA passenger in order to initiate motility and pathogenic viability in the host. The IcsA PATR glycine substitutions and deletion mutants however, displayed marked reductions in passenger secretion efficiency (Figure 5.3D). The mutants remained significantly inaccessible to PK treatment and were not completely processed over the 1 hour time-course. PK inaccessible mutant forms were all stalled in the periplasmic topology since complete digestion was observed when the OM was permeabilized. This also indicated that PATR mutations did not alter the intrinsic protease sensitivity of IcsA. The band intensities from this experiment were also estimated and compared (Figure 5.7[S]), further suggesting significant differences between IcsA and IcsA PATR mutants.

These results show that PATR glycine substitutions or deletion caused a marked decrease in the rate of passenger maturation to the bacterial surface. This provides an explanation for the reduced levels of surface exposed PATR mutant passengers observed microscopically and by limited protease shaving.

5.6.3. Analysis of the PATR within the AT family.

With knowledge of the function of the PATR established *in vivo*, we attempted to identify the wider importance of the PATR within the AT family. The PATR was identified in a large number of ATs, with examples shown aligned in Figure 5.4A. These include known subtilisin-type serine proteases such as the inflammatory EprS from *Pseudomonas aeruginosa* (Kida *et al*, 2013), and NalP, the processor of other ATs from *Neisseria meningitidis* (van Ulsen *et al*, 2003). Also shown aligned is PATR9 from the fibronectin-binding host colonization factor ShdA of *Salmonella enterica* (ShdA contains an array of PATRs (Kingsley *et al*, 2000, Kingsley *et al*, 2004a) (see Figure 5.8[S])). The glycine residues investigated in this study, as well as other PATR residues, are highly conserved on a level not previously observed in AT passengers, especially between those of varied function.

An analysis of the usage of domains and motifs in all ATs within the UniProt Knowledgebase was also conducted by first identifying unique ATs using InterPro (Hunter *et al*, 2012) AT β-barrel identifiers IPR005546 and/or IPR006315 (7659 proteins). Figure 5.4B shows the different virulence domain usage of ATs (at least one site per passenger). Remarkably, 29.2% of the unique representative ATs within the database (2240 proteins) had at least one copy of the PATR (Figure 5.4B, blue). This is similar to the abundance of the PL region which we found present in 37.4% of the unique representative ATs (2864 proteins). Yet, the actual usage of the PATR in ATs is much higher since it is normally present in multiple copies (see further).

The usage of conserved domains, and combinations of these domains, were further grouped and analyzed based on the presence or absence of at least one annotated copy of the PATR (IPR013425). There was a significant difference ($P < 0.0001$) in the representations of passenger domain virulence traits between PATR-positive and non-PATR ATs (Figure 5.4C). For example; lipase-like, as well as type S6 (SPATE-like (Rawlings *et al*, 2014, Ruiz-Perez & Nataro, 2014)) serine proteases, and vacuolating cytotoxins were only present in non-PATR ATs (Figure 5.4Ci). Conversely, ATs containing type 2 phosphatidic acid phosphatase (PAP2) domains, Polymorphic OM protein repeats (POMPs), and type S8 (subtilisin-type (Rawlings *et al*, 2014)) serine proteases were all highly represented in PATR-positive ATs (Figure 5.4Ci). Interestingly, ATs with a Pectin lyase-like (PLL) region (an indicator of further β-helical wrapping) as the only other identifying passenger feature were represented more than twice as high in PATR-type ATs relative non-PATR ATs (Figure 5.4Cii). The inverse of this was true for ATs that also contained a PL region (Figure 5.4Cii).

Furthermore, the PATR was never observed in the passenger with the PL as an exclusive partnership (i.e.: PATR plus PL only passenger). This minimal overlap between the PATR and the PL is further shown via Venn diagram (see Figure 5.9[S]). Also identified was a significant difference ($P < 0.0001$) in protein lengths, where the mean length of PATR-positive ATs was 503 a.a. (± 13 a.a.) longer than non-PATR ATs. This is seen as a positively skewed lengths distribution for ATs containing a PATR (Figure 5.4D). Moreover, within the PATR-positive ATs there was a significant ($P < 0.0001$) correlation between increasing AT length and PATR copy number (Figure 5.4E).

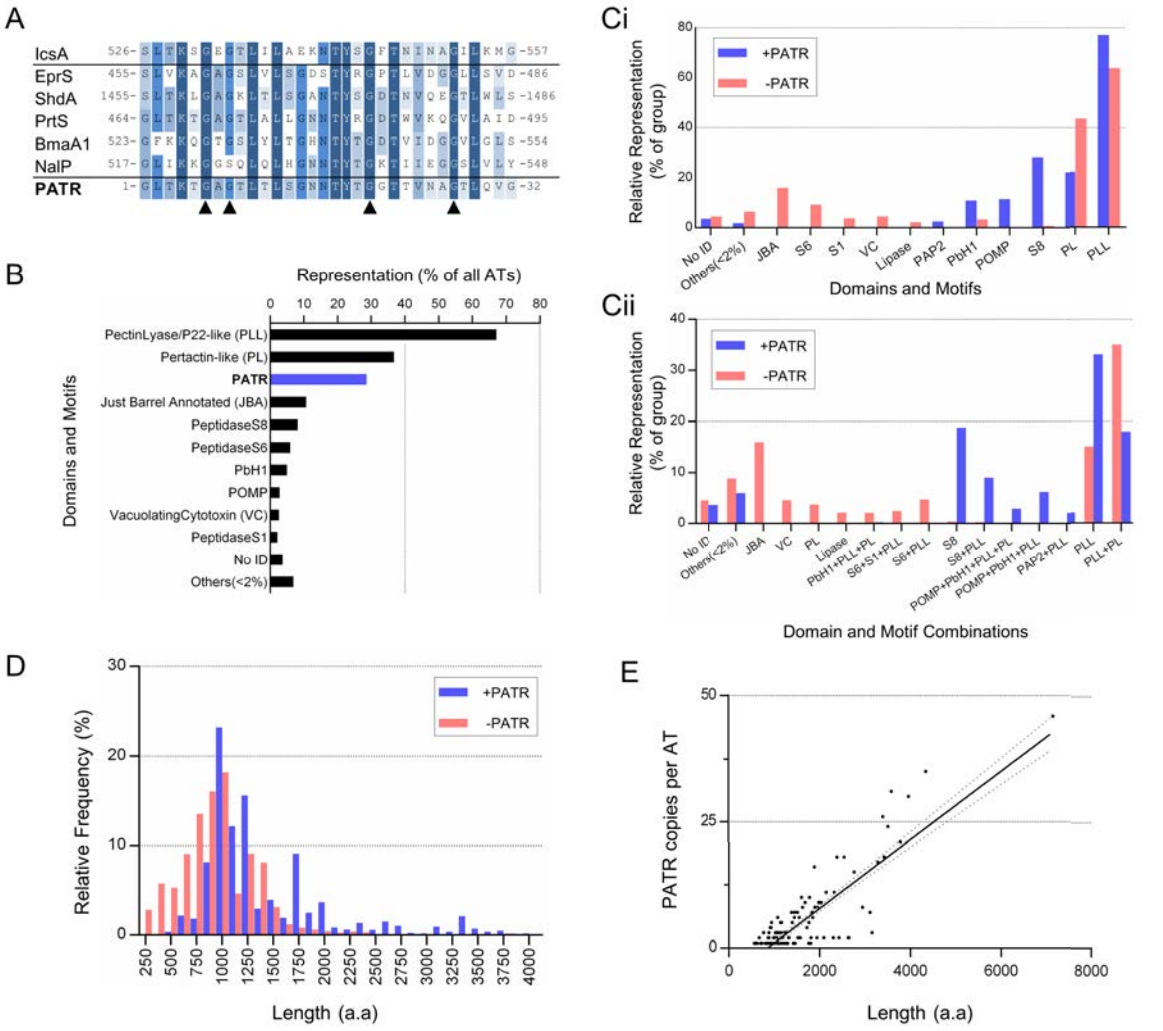


Figure 5.4: The importance of the PATR within the AT family.

(A) Alignment of the PATR sites of IcsA and five other ATs; EprS (Q9HY75), ShdA (Q9XCJ4), subtilisin-type serine protease PrtS (P09489) from *Serratia marcescens* (Shikata *et al*, 1992), subtilisin-type serine protease BmaA1 (H6T4K9) of *Haemophilus parasuis* (Pina-Pedrero *et al*, 2012), and NalP (Q8GKS5). Accessions are UniProtKB IDs. Black arrows indicate glycines investigated in this study. Additional PATR sites for ShdA are shown in Figure 5.8S. **(B)** ATs within the UniProt Knowledgebase were assessed based on their usage of different domains and motifs (at least one InterPro ID hit per sequence). Domain usage is shown as a percentage of the total AT sequences assessed (7659). Others = all domains / motifs that were < 2 % represented in ATs, No ID = entries

that are yet to be annotated. **(C)** ATs within the UniProt Knowledgebase were further grouped by the presence (+PATR) or absence (-PATR) of detectible PATR and further analyzed by InterPro (IPR) ID domain usage (single; **Ci**) and domain combinations (**Cii)**. Domain usage / combinations are shown as a percentage of the total AT sequences assessed per group (+PATR = 2240, -PATR = 5419). The dependency of domain usage / combination on the presence or absence of the PATR is significant in each case ($P < 0.0001$, chi-square). Note: the ordering of the domains in **Cii** does not indicate their position within the primary structure. Others = all combinations that were < 2 % represented in both groups. IPR IDs of domains presented = PectinLyase/P22-like (PLL); 012332, 011050, 012334, Pertactin-like (PL); 004899, 003991, 003992, Peptidase S8 (S8); 000209, 022398, 023828, 015500, 023827, 017318, Peptidase S6 (S6); 000710, parallel β-helix (PbH1); 006626, Polymorphic OM protein repeat (POMP); 003368, Vacuolating Cytotoxin (VC); 003842, 004311 Peptidase S1 (S1); 018114, 001254, type 2 phosphatidic acid phosphatase (PAP2); 000326, Lipase; 017186, 001887, 008265, 013831. The minimal overlap between the PATR and PL is shown further in Figure 5.9[S]. **(D)** Lengths frequency histogram. The mean lengths are significantly different (949.9±5.367 and 1453±15.24 for the -PATR and +PATR groups respectively) as tested by two-tailed t-test ($P < 0.0001$). **(E)** PATR copy number per AT correlates significantly with length (two-tailed, $P < 0.0001$, Pearson). N = 200 randomly chosen PATR-containing ATs.

Together, these data suggests that the presence or absence of a PATR strongly impacts the probability of containing certain passenger virulence functions, the potential size of the protein, and delineates an important sub-group of PATR-type ATs. Nevertheless, the positions of the PATR motifs within different passengers, and whether there are defined patterns of PATR usage dependent on the presence of other functionally important regions, remained unknown. To investigate the architecture of PATR-type AT passengers, and the position of PATR sites relative other virulence domains, a random sampling was taken of at least 100 representative AT sequences from each of the seven common PATR-type domain combination groups (identified in Figure 5.4Cii). The common positions of each annotated domain, and their occupancy over the passenger, was calculated (see *Experimental Procedures*) resulting in Figure 5.5. Most PATR-type AT architecture groups display a striking consistency in the placement of PATR sequences along the passenger. For instance, architectures containing POMPs (Figure 5.5, groups 5 and 6) always contained three PATRs (placed at approximately 0.35, 0.6, and 0.7 proportional distance respectively from the N-terminus of the passenger) with defined amino acid spacing between these and little variance. This consistency in placement is also seen for S8 type proteases and PAP2 type phosphatases (Figure 5.5, groups 2, 4, and 7). Architecture groups 1 and 3 have the most variance in PATR placement and number. When only the PATR and PLL are present (Figure 5.5, group 1) the number of PATRs per passenger varies greatly (±8.68) around a mean of 8.79. However, the likelihood of a passenger having multiple PATRs drops substantially (1.93 ±1.44) when a PL domain is also present. Other notable patterns are that: (i) PATRs almost always associate with a PLL region, (ii) PATRs are always located C-terminal to catalytic domains, and (iii) the majority of common PATR sites are located in the C-terminal 50 % of the passenger.

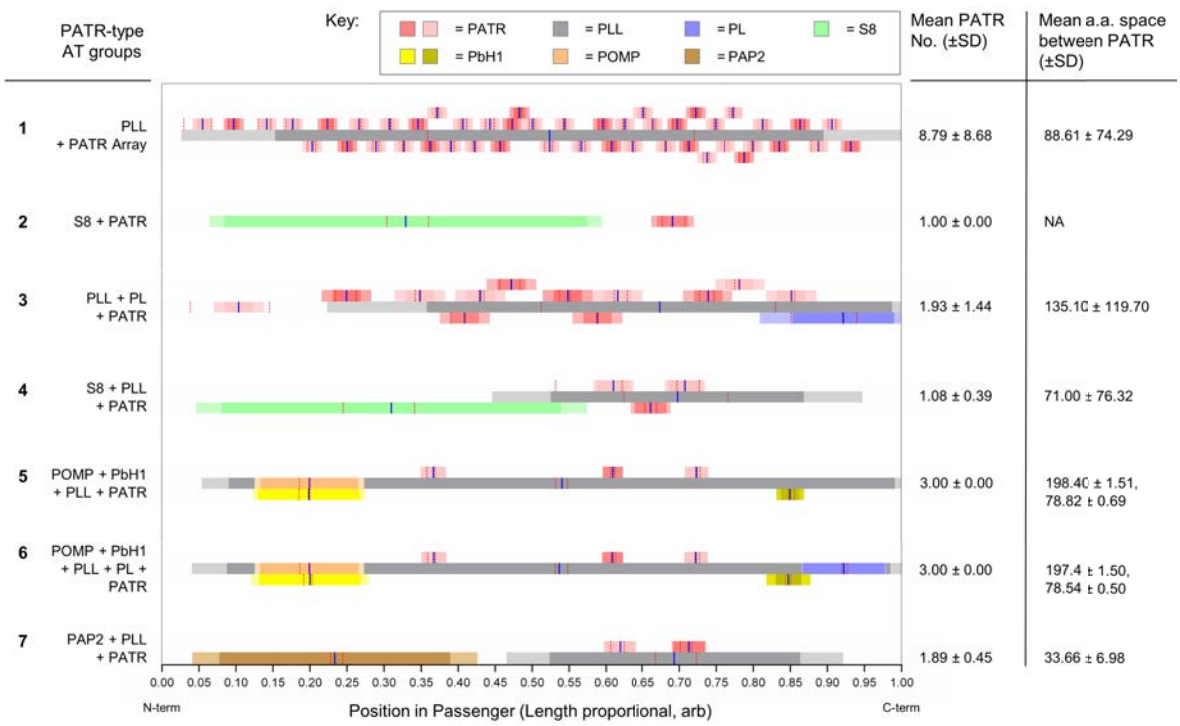


Figure 5.5: PATR-type AT architecture groups have distinct PATR placements.

Analysis of the functional architectures of the seven common PATR-type AT domain combination groups (identified in Figure 5.4C). Each architecture is based on at least 100 representative sequences (see *Experimental Procedures*). Group numbering is in descending number of total representative sequences per group. Blue vertical lines = median common midpoint of the domain or motif on the passenger, red-dotted vertical lines = interquartile range of the midpoint about the median, dark colour shaded box = mean area of the passenger occupied by the domain or motif about the midpoint, light color shaded box = plus one standard deviation (SD) of the mean area of the passenger occupied by the domain or motif. The mean number of PATR, and the space between PATR (amino acids; a.a.), is also shown for each architecture group. Note that the different shades of PbH1 (yellow) are to help distinguish between the common N-terminal cluster (light yellow) and the common C-terminal site (dark yellow). The different shades of PATR (red) sites are to allow easier differentiation between these sites. Sites are offset to avoid overlapping of boxes. PLL = PectinLyase/P22-like, PL = Pertactin-like, S8 = Peptidase S8, PbH1 = parallel β-helix, POMP = Polymorphic OM protein repeat, and PAP2 = type 2 phosphatidic acid phosphatase.

Although PATR-containing ATs appear to be a very important sub-type of ATs, none of the ATs with solved passenger structures appeared to contain the PATR when scrutinized by sequence analyses. Therefore, the tertiary structure of the PATR consensus sequence was modeled using I-TASSER (Roy *et al*, 2010, Xu *et al*, 2011) and then structurally aligned to all the solved AT passengers using TM-Align (Zhang & Skolnick, 2005). The PATR was predicted to form a right-handed β-helical triangular wedge with all PATR glycines clustered at each vertex (Figure 5.6A). Remarkably, upon alignment to passenger structures, we identified PATR-like sites in Ag43, Hap, and IgA1P (see Figure 5.6B, Figure 5.10[S], and Table 5.2[S] for additional sites). The alignment shows the positions of the conserved glycines

characterized in this study. The 'velcro-like' Ag43 (Heras *et al*, 2014) of pathogenic *E. coli* had the highest alignment score with clustering of the glycines at the $G^{6}/G^{8}/G^{27}$ PATR apex (Ag43 glycines G^{499}, G^{501}, G^{520} respectively). Identification of these putative degenerate PATR sequences indicates an even wider distribution of this site in ATs than can be identified by sequence motif recognition alone.

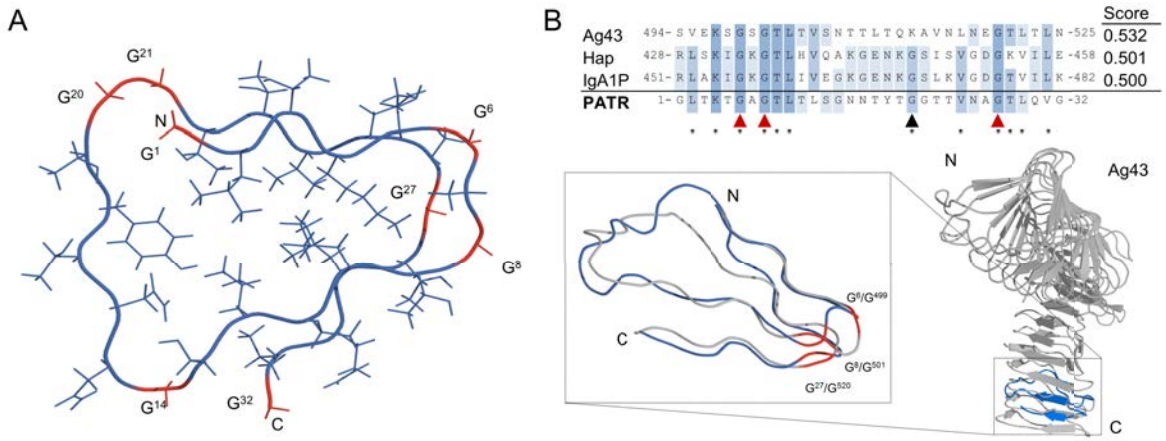


Figure 5.6: Identification of PATR-like sequences in structurally solved ATs.

(A) I-TASSER generated tertiary structure of the PATR. Orientation is a top-down cross-section from N- to C-terminus. The PATR is a predicted triangular wedge with all glycines (red) clustered at the three apexes. To find degenerate PATR this model was spatially aligned to all the solved AT passenger structures using TM-align. Identified degenerate PATR are aligned in **(B)** with the highest scoring site from Ag43 (Q8CVR0) shown below. Spatially conserved glycines (red) between the PATR (blue) and Ag43 (grey) are indicated. Positions of conservation between PATR-like sequences and PATR model (Figure 5.1B) are indicated by asterisks. For full lists and structure alignments see Table 5.2[S] and Figure 5.10[S].

5.7. Article Discussion

This study reveals the importance of a previously underappreciated and uncharacterized passenger feature which we have termed here, the Passenger-associated Transport Repeat (PATR). Strikingly, ~30 % of the unique AT representatives analyzed contain a PATR. This is comparable to the abundance of the Pertactin-like (PL) region which we found here to be ~37 % (slightly more than previous estimates (Drobnak *et al*, 2015b)). However the actual usage of the PATR is much higher than 30% since it is commonly present in many copies per passenger. Through alanine substitution of conserved PATR glycines (as well as a complete PATR deletion) within the PATR-type AT IcsA, we have shown *in vivo* that the PATR is required for efficient secretion of the passenger to the bacterial surface. Significant reductions in the transport of the passenger to the cell surface were observed for PATR mutants. This resulted in significantly decreased (approximately half that of wild-type) steady-state levels of surface exposed IcsA passenger which was further observed as a significant decrease in N-WASP recruitment levels by intracellular *S. flexneri* and by a lack of accessibility upon protease treatment. Furthermore, the substantial lags observed in the amount of translocated passenger over an extended time, also shows that the passenger was exposed to the periplasmic topology for a prolonged period where it may remain unfolded (Drobnak *et al*, 2015a). This would increase the likelihood of proteolysis by known proteases (for instance, DegP (Jong *et al*, 2007, Purdy *et al*, 2007)) and may explain the observed increase in degraded forms of PATR mutants in the OM and the Proteinase K resistant fragments.

It is clear that the PATR is required for some aspect of the efficient traffic and export of the AT. Yet this begs the question – by what mechanism(s)? Structural modelling suggests that the PATR prescribes a minimal right-handed triangular β-helix with the conserved glycines clustered spatially at the three vertices. We found an export deficiency effect arising from glycine-alanine substitutions at two of these three putative vertices. It is likely that these glycines are required for the flexibility of stable PATR folding at the corners, and that substitution disrupts the space requirements for this folding. Sequential folding as a mode of passenger translocation is a well-studied notion in other ATs (Junker *et al*, 2006, Renn & Clark, 2008, Junker *et al*, 2009, Peterson *et al*, 2010). Drawing from this, it is conceivable that proper sequential folding of the PATR is also required for the translocation of PATR-type ATs, similarly to what is observed for Pertactin and PL ATs. Certainly, it is intriguing that the presence of a PL region corresponds with a lower potential for the presence of a PATR and *vice versa*. Further, the PATR was never observed with the PL as an exclusive partnership in the passenger – an additional region (such as the Pectin lyase-like region) must also be

present for the PATR and PL to be in the same AT. Although care must be taken when attributing a folding function to PL regions (Drobnak *et al*, 2015b), it appears that both these regions have overlapping (but not equivalent) conserved functions in passenger biogenesis. This rationale needs to be investigated further through *in vitro* biophysical and structural characterization of the PATR before a direct effect in passenger translocation is confirmed. However, we have observed that PATR-type ATs are significantly larger (by 503±13 a.a.) than non-PATR ATs, and that the size of the AT correlates with PATR copy number. It is possible that the PATR acts as a dispersed module for folding stability where some larger ATs (for instance ShdA) may require multiple PATR modules for the efficient translocation of larger passengers. Indeed, the notion that longer passengers may rely more on folding as a driving force for translocation, has recently been postulated by others (Drobnak *et al*, 2015a). While a direct effect of the PATR on passenger translocation due to folding is a simple and attractive model, there remains the possibility that the PATR is required at an earlier stage in AT transport. It is plausible that the PATR is a site of recognition for specific chaperoning to maintain translocation competence. Alternatively, it may be required for appropriate and timely interactions with other components of the assembly pathway (such as the BAM complex) allowing initiation of translocation. Whether the PATR acts directly in translocation, or more indirectly in allowing translocation initiation, remains an exciting question worthy of further inquiry.

Besides the lack of PATR association with the PL region, there was also a striking association between passenger virulence attributes and the presence or absence of the PATR. The PATR was not found in lipase-like passengers, which is consistent with our modeling of the PATR as a minimal β-helical fold (since lipases are not β-helix based (Van den Berg, 2010)). Subtilisin-type (S8) serine proteases were also commonly PATR-type ATs, whereas SPATE (serine protease ATs of the enterobacteriacea)-like (S6) serine proteases generally excluded the presence of a PATR. The reasons behind these observations are unknown, but we speculate that the export requirements of a passenger with S8-type regions may be more amenable to a PATR-type mechanism as opposed to the requirements for S6-type passengers. It is also possible that the difference in representation reflects variances in the usage of the PATR between the enterobacteriaceae and other families. In support of this we also observed an increased association of the PATR with passengers containing the POMP repeat which is highly conserved in the Pmp adhesin ATs of the *Chlamydiaceae* (Henderson & Lam, 2001). Nevertheless, the association between the PATR and certain virulence domains allowed the PATR-type ATs to be grouped into seven common architectures. These display distinct

patterns of PATR placement within the passenger depending on the other functional regions present. Together, these results implicate the PATR as a convenient building block in passengers, providing scaffolding for other functional regions. Indeed, it has been previously proposed that small sequences have been incorporated into passenger architectures to enable niche specialization of ATs during evolution (Celik *et al*, 2012).

In conclusion, this study has uncovered the importance of a previously uncharacterized repeat that plays a role in AT secretion. The PATR is present in many passengers and delineates a further subtype of ATs that have defined architectures. These results stimulate the need for further investigation to expose the exact mechanism of PATR mediated export, whether it either directly effects passenger translocation or translocation initiation through specific interactions, to establish biophysical characteristics of this repeat, and to uncover its phylogenic origin and diversity within Gram-negative bacteria. Finally, it should be noted here that we identified that the PATR is also present in ~700 unique proteins that did not contain an identifiable AT β-barrel. Although it must be further established, this may indicate a role for the PATR in other Type V secretion systems – namely, the Two-partner pathway. Nevertheless, this work has highlighted the notion that passenger compositions are 'mixed-and-matched' to suit precise secretion or function requirements.

5.8. Article Experimental Procedures

5.8.1. Bacterial strains and plasmids.

Lists of strains and plasmids utilized in this study are included in the Supplementary Information (see Table 5.1[S]) which includes details of their construction (see *Supplementary Methods*) and oligonucleotides used (see Table 5.3*S*). *S. flexneri* colonies were grown on Congo Red agar for confirmation of virulence plasmid presence before routine growth in Luria-Bertani (LB) media at 37 °C with shaking. For all experiments, bacteria were sub-cultured (1:50 or 1:100) to a log-phase OD_{600} reading of 0.5 before use. When required, broths were supplemented with the following additives at respective concentrations; 0.2 % (w/v) glucose, 0.2 % (w/v) arabinose, tetracycline (10 μg mL^{-1}), kanamycin (50 μg mL^{-1}), chloramphenicol (25 μg mL^{-1}) and ampicillin (50 μg mL^{-1}).

5.8.2. Antibodies.

Polyclonal rabbit anti-IcsA (passenger) and polyclonal rabbit anti-N-WASP were produced and validated as described previously (Van Den Bosch *et al*, 1997, May & Morona, 2008). Polyclonal rabbit anti-SurA was a generous gift from Carol Gross (University of California, USA). Polyclonal rabbit anti-Wzz was produced as described previously (Daniels & Morona, 1999). Polyclonal rabbit anti-BamA was a generous gift from Thomas Silhavy (Princeton University, USA). Mouse anti-DnaK monoclonal antibody was from Enzo Life Sciences. HRP conjugated goat anti-mouse or goat anti-rabbit were used for western immunoblot and detected using Chemiluminescence (Sigma).

5.8.3. Total bacterial protein samples.

5 x 10^8 of log-phase bacteria were collected by centrifugation (16,000 x *g*, 1 min, 4 °C), resuspended in 100 μL of SDS-PAGE loading buffer (Lugtenberg *et al*, 1975), and heated to 100 °C for 10 min. Replicate total cell samples were pooled 1:1 before analysis.

5.8.4. Bacterial IcsA labelling.

Immunofluorescence microscopy and fluorescence quantitation was conducted as described previously (Tran *et al*, 2013). All solutions used were filtered through a 0.2 μm nitrocellulose filter. 10^8 of log-phase bacteria were harvested from a 1:50 sub-culture by centrifugation (16,000 x *g*, 2 min, 20 °C), resuspended in 3.7 % (v/v) formaldehyde solution

(Sigma) in phosphate buffered saline (PBS), and incubated at 20 °C for 20 min. Fixed bacteria were washed twice in PBS before resuspension in 100 µL of PBS. 5 µL of the bacteria were spotted onto sterile round coverslips (at the bottom of a 24-well tray) that were previously treated with 10 % (v/v) poly-L-lysine solution (Sigma) in PBS. Bacteria were centrifuged (775 x *g*, 5 min, 20 °C) and then incubated for 2 h with anti-IcsA diluted 1:100 in PBS containing 10 % (v/v) fetal calf serum (FCS). Bacteria were washed three times with PBS and then incubated for 30 min at 37 °C with donkey anti-rabbit Alexa Fluor 488 antibody (Invitrogen) diluted 1:100 in PBS containing 10 % (v/v) fetal calf serum (FCS). Bacteria were washed three times with PBS before mounting with 20 % Mowiol 4-88 (Calbiochem), 4 mg ml^{-1} *p*-phenylenediamine.

5.8.5. Cell infection and N-WASP/F-actin/DNA labelling.

Infection of semi-confluent HeLa cell monolayers with *S. flexneri* was conducted as described (Teh *et al*, 2012a). HeLa cells were grown on sterile round coverslips at the bottom of 24-well trays. Log-phase bacteria were harvested from a 1:50 sub-culture by centrifugation (16,000 x *g*, 2 min, 20 °C) and diluted to 3 x 10^8 bacteria/mL in Dulbecco's PBS (D-PBS). HeLa Cells were washed with 10 % (v/v) FCS in minimal essential medium (MEM) and then 80 µL of bacteria were added before centrifugation (500 x *g*, 5 min, 20 °C) to assist invasion. After incubation at 37 °C with 5 % CO_2 for 1 h, cells were washed three times with D-PBS, and incubated a further 1.5 h with 500 µL of MEM supplemented with 10 % (v/v) FCS and 40 µg mL^{-1} gentamycin. Cells were then washed three times with D-PBS, fixed for 15 min with 3.7 % (v/v) formaldehyde solution (Sigma) in PBS, and washed twice with PBS. Before staining, cells were incubated with 50 mM NH_4Cl in D-PBS for 10 min, washed with PBS, permeablized with 0.1 % (v/v) Triton X-100 in PBS for 5 min, and washed with PBS. Cells were blocked with 10 % (v/v) FCS in PBS for 20 min, before aspiration and addition of anti-N-WASP diluted at 1:100 in PBS containing 10 % (v/v) FCS for 30 min at 37 °C. Cells were washed three times with PBS, and then incubated for 1 h at 37 °C with donkey anti-rabbit AlexaFluor 594 antibody (Invitrogen) and AlexaFluor 488 phalloidin (Invitrogen) diluted to 1:100 and 1:200 respectively in PBS containing 10 % (v/v) FCS. After three washes with PBS, DNA was stained with 10 µg mL^{-1} DAPI for 1 min, washed three times with PBS, and mounted for microscopy as described above.

5.8.6. Microscopy.

All images of stained bacteria or infected HeLa cells were captured using an Olympus IX-70 Microscope and MetaMorph software (Version 7.7.1.0 Molecular Devices) with a phase contrast 100 x oil immersion objective, a 1.5 x enlarger, and focused with the aid of an MFC-2000 digital z-axis controller (Applied Scientific Instruments). For fluorescence imaging an X-Cite 120Q lamp was used set at high intensity. All bacterial IcsA fluorescence images were acquired with 100 millisecond exposures. All N-WASP fluorescence images were acquired with 500 millisecond exposures. Fluorescence images for background correction were taken for each experiment. IcsA and N-WASP fluorescence images for presentation were recolored using the LUT options in ImageJ. MetaMorph region measurement tools were used to quantitate fluorescence intensities for individual bacteria. For IcsA quantitation, 50 bacteria were routinely measured for each experiment. For N-WASP recruitment, all bacteria within an infected cell were measured for each experiment.

5.8.7. OMP extraction.

OMPs were isolated using differential Sarkosyl treatment (Hobb *et al*, 2009). 5 x 10^{10} log-phase bacteria were collected from a 1:50 sub-culture by centrifugation (3,000 x *g*, 20 min, 4 °C), resuspended in 15 mL of 10 mM HEPES, pH 7.5, and lysed by sonication. Debris was removed by centrifugation (10,000 x *g*, 10 min, 4 °C) and supernatant ultracentrifuged (149,000 x *g* average, 1 h, 4 °C). Whole membrane pellets were homogenized in 15 mL of 10 mM HEPES, pH 7.5, re-ultracentrifuged (as above), homogenized in 15 mL of 1 % (w/v) Sodium N-lauroylsarcosinate, 10 mM HEPES, pH 7.5, and incubated at 37 °C for 30 min with inversion. OMPs were collected by ultracentrifugation (as above), homogenization in 15 mL of 10 mM HEPES, pH 7.5, ultracentrifuged a final time (as above), and homogenization in 250 µL of 10 mM HEPES, pH 7.5. OMPs were diluted in 10 x SDS-PAGE loading buffer and heated to 100 °C for 10 min before analysis.

5.8.8. *In vivo* limited protease shaving.

1:50 sub-cultures of *S. flexneri* harboring pBADIcsA and derivatives (see Table S1) were grown to log-phase in 10 mL of LB supplemented with glucose. Bacteria were collected by centrifugation (3,000 x *g*, 5 min, 4 °C), washed with LB twice, resuspended in 10 mL of LB containing arabinose, and further cultured for 2 h at 37 °C. A total of 10^9 bacteria were then collected (16,000 x *g*, 1 min, 4 °C), washed twice with 1 mL of PBS, resuspended in 1 mL of

PBS, and placed on ice. A 100 µL aliquot was precipitated with 12 % (w/v) final concentration of trichloroacetic acid (TCA) on ice (0 min; untreated). Proteinase K was added to bacteria for a final concentration of 1 µg mL^{-1} and 100 µL aliquots were taken at 1, 5, 30, and 60 min time points, treated with 4 mM final concentration of phenylmethanesulfonyl fluoride (PMSF) to stop proteolysis, and immediately precipitated similarly to the 0 min sample. All precipitated samples were washed with acetone, dried, and resuspended with 25 µL of SDS-PAGE loading buffer.

5.8.9. *S. flexneri* pulse-chase proteolysis assay.

The arabinose/glucose (on/off) expression switch of the vector pBAD30 (Guzman *et al*, 1995) was utilized for controlled expression essentially as described by (Leyton *et al*, 2014). 1:100 sub-cultures of *S. flexneri* harboring pBADIcsA and derivatives (see Table S1) were grown to log-phase in 100 mL of LB supplemented with glucose before collection by centrifugation (4,000 x *g*, 4 min, 4 ˚C). Bacteria were washed with LB, resuspended in media containing arabinose, and incubated (5 min, 25 ˚C) for IcsA production (pulse). Bacteria were collected (as above), resuspended in 30 mL of media containing glucose, and placed on a 25 ˚C block for 60 min (chase). During the chase, four 1 mL aliquots were taken at times 0 (resuspension), 1, 5, 10, 20, 40, and 60 mins. The protein from the first aliquot was precipitated with 12 % (w/v) final concentration of TCA on ice. The second aliquot was treated with a final concentration of 10 µg mL^{-1} Proteinase K (PK) on ice for 10 min. Proteolysis was then stopped by addition of 4 mM final concentration of PMSF before TCA precipitation (as above). Bacteria from the third aliquot were collected (16,000 x *g*, 1 min, 4 ˚C), treated with 20 µL of chloroform to permeabilize the OM (Ames *et al*, 1984, Wagner *et al*, 2009), and then PK treated, stopped, and precipitated as described for the second aliquot. All precipitated samples were washed with acetone, dried, and resuspended with 50 µL of SDS-PAGE loading buffer per 1 OD_{600} unit (measured using the fourth aliquot). Samples were heated to 100 ˚C for 10 min before analysis. The relative surface exposure was estimated by the formula: Relative Export=100-100(PK+/PK-) after densitometric estimation using ImageJ.

5.8.10. Database analysis.

To generate AT annotation lists, the UniProt knowledgebase (Jain *et al*, 2009, Magrane & Consortium, 2011) (uniprot.org) was used due to its extensive architecture annotation of protein entries by InterPro (Hunter *et al*, 2012). A list of ATs that were PATR

exclusive (-PATR) was generated by searching for entries including InterPro cross-references for the AT β-barrel (IPR005546/IPR006315) but not the PATR (IPR013425). Lists were pruned to exclude fragments. Note that IPR013425 incorporates both PFAM (Finn *et al*, 2014b) (PF12951) and TIGRFAM (Haft *et al*, 2003) (TIGR02601) PATR models. A list of ATs that were PATR inclusive (+PATR) was generated by searching for entries including both the β-barrel and the PATR. This yielded 14518 -PATR entries and 4961 +PATR entries. To remove redundancy, each list was clustered into UniRef100 sequence clusters (Suzek *et al*, 2007) which groups identical sequences and represents each group with a representative entry. This produced non-redundant -PATR and +PATR lists of 5419 and 2240 representatives respectively. Note that this is a higher number of ATs than previous estimates (Celik *et al*, 2012) but is consistent with differing databases used and differing methods of AT identification. Each group was further analyzed based on other InterPro annotations, protein length, and PATR copy number.

To investigate PATR-type AT architecture, at least 100 representatives were randomly chosen from each common architecture group (seven groups). For groups that had less than 100 representatives, all sequences were analyzed. The passenger size was obtained by subtracting the InterPro identified AT β-barrel from each sequence. The midpoints and sizes of the annotated domains and motifs were calculated from InterPro annotated boundaries of the latter. These metrics were made relative to the size of the passenger that they were present in allowing normalization (proportional positions between 0 and 1). The median and interquartile range of the midpoints of each domain or motif was found along the normalized passenger and the median midpoint used to place the mean passenger occupancy of that domain or motif.

5.8.11. PATR structural modelling and identification of degenerate PATR.

Modeling of the PATR sequence (PF12951 consensus) was achieved using I-TASSER which uses threading and *ab initio* modeling (zhanglab.ccmb.med.umich.edu/I-TASSER) (Roy *et al*, 2010, Xu *et al*, 2011). The model had a high C-score of -0.43 (possible range is -5 to 2 were higher values signifies higher confidence in the model), a TM-score of 0.66±0.13 (>0.5 indicates correct topology, <0.17 indicates random similarity. The following templates contributed to the model: Hap (PDB 3SYJ), Ag43 (PDB 4KH3), WlbB (PDB 3MQG), PCSK9 (PDB 2QTW), and KalataB1 (PDB 1JJZ). TM-Align was used to identify degenerate PATR from solved passenger structures (zhanglab.ccmb.med.umich.edu/TM-

align) (Zhang & Skolnick, 2005). SCOP/CATH protein folds are shared when TM-score > 0.5. All alignment sites are detailed in Table 5.2S.

5.9. Article Acknowledgments

This work is supported by a Program Grant (No.: 565526) from the National Health and Medical Research Council (NHMRC) of Australia. MTD is the recipient of a Doctor of Philosophy scholarship from the University of Adelaide. We also thank the Research Centre for Infectious Diseases (RCID) for support during this work.

We greatly appreciate Alistair Standish for critical reading of the manuscript. We also appreciate the donations of antisera from both Carol Gross and Thomas Silhavy, and for bioinformatics advice from Dan Kortschak.

5.10. Article References

Ames, G.F., Prody, C. & Kustu, S., (1984) Simple, rapid, and quantitative release of periplasmic proteins by chloroform. *J Bacteriol* 160: 1181-1183.

Bernardini, M.L., Mounier, J., d'Hauteville, H., Coquis-Rondon, M. & Sansonetti, P.J., (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.

Besingi, R.N., Chaney, J.L. & Clark, P.L., (2013) An alternative outer membrane secretion mechanism for an autotransporter protein lacking a C-terminal stable core. *Mol Microbiol* 90: 1028-1045.

Braselmann, E. & Clark, P.L., (2012) Autotransporters: The cellular environment reshapes a folding mechanism to promote protein transport. *J Phys Chem Lett* 3: 1063-1071.

Celik, N., Webb, C.T., Leyton, D.L., Holt, K.E., Heinz, E., Gorrell, R., Kwok, T., Naderer, T., Strugnell, R.A., Speed, T.P., Teasdale, R.D., Likic, V.A. & Lithgow, T., (2012) A bioinformatic strategy for the detection, classification and analysis of bacterial autotransporters. *PloS One* 7: e43245.

Dai, J.J., Wang, S.H., Guerlebeck, D., Laturnus, C., Guenther, S., Shi, Z.Y., Lu, C.P. & Ewers, C., (2010) Suppression subtractive hybridization identifies an autotransporter adhesin gene of *E. coli* IMT5155 specifically associated with avian pathogenic *Escherichia coli* (APEC). *BMC Microbiol* 10.1186/1471-2180-10-236.

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.

Dautin, N. & Bernstein, H.D., (2007) Protein secretion in gram-negative bacteria via the autotransporter pathway. *Annu Rev Microbiol* 61: 89-112.

Drobnak, I., Braselmann, E., Chaney, J.L., Leyton, D.L., Bernstein, H.D., Lithgow, T., Luirink, J., Nataro, J.P. & Clark, P.L., (2015) Of linkers and autochaperones: an unambiguous nomenclature to identify common and uncommon themes for autotransporter secretion. *Mol Microbiol* 10.1111/mmi.12838.

Drobnak, I., Braselmann, E. & Clark, P.L., (2015) Multiple Driving Forces Required for Efficient Secretion of Autotransporter Virulence Proteins. *J Biol Chem* 10.1074/jbc.M114.629170.

Egile, C., Loisel, T.P., Laurent, V., Li, R., Pantaloni, D., Sansonetti, P.J. & Carlier, M.F., (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.

Finn, R.D., Bateman, A., Clements, J., Coggill, P., Eberhardt, R.Y., Eddy, S.R., Heger, A., Hetherington, K., Holm, L., Mistry, J., Sonnhammer, E.L.L., Tate, J. & Punta, M., (2014) Pfam: the protein families database. *Nucleic Acids Res* 42: D222-D230.

Goldberg, M.B. & Theriot, J.A., (1995) *Shigella flexneri* surface protein IcsA is sufficient to direct actin-based motility. *Proc Natl Acad Sci U S A* 92: 6572-6576.

Grijpstra, J., Arenas, J., Rutten, L. & Tommassen, J., (2013) Autotransporter secretion: varying on a theme. *Res Microbiol* 164: 562-582.

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

Haft, D.H., Selengut, J.D. & White, O., (2003) The TIGRFAMs database of protein families. *Nucleic Acids Res* 31: 371-373.

Henderson, I.R. & Lam, A.C., (2001) Polymorphic proteins of *Chlamydia spp.*-autotransporters beyond the Proteobacteria. *Trends Microbiol* 9: 573-578.

Henderson, I.R., Navarro-Garcia, F., Desvaux, M., Fernandez, R.C. & Ala'Aldeen, D., (2004) Type V protein secretion pathway: the autotransporter story. *Microbiol Mol Biol Rev* 68: 692-744.

Heras, B., Totsika, M., Peters, K.M., Paxman, J.J., Gee, C.L., Jarrott, R.J., Perugini, M.A., Whitten, A.E. & Schembri, M.A., (2014) The antigen 43 structure reveals a molecular Velcro-like mechanism of autotransporter-mediated bacterial clumping. *Proc Natl Acad Sci U S A* 111: 457-462.

Hobb, R.I., Fields, J.A., Burns, C.M. & Thompson, S.A., (2009) Evaluation of procedures for outer membrane isolation from *Campylobacter jejuni*. *Microbiology-SGM* 155: 979-988.

Hunter, S., Jones, P., Mitchell, A., Apweiler, R., Attwood, T.K., Bateman, A., Bernard, T., Binns, D., Bork, P., Burge, S., de Castro, E., Coggill, P., Corbett, M., Das, U., Daugherty, L., Duquenne, L., Finn, R.D., Fraser, M., Gough, J., Haft, D., Hulo, N., Kahn, D., Kelly, E., Letunic, I., Lonsdale, D., Lopez, R., Madera, M., Maslen, J., McAnulla, C., McDowall, J., McMenamin, C., Mi, H., Mutowo-Muellenet, P., Mulder, N., Natale, D., Orengo, C., Pesseat, S., Punta, M., Quinn, A.F., Rivoire, C., Sangrador-Vegas, A., Selengut, J.D., Sigrist, C.J., Scheremetjew, M., Tate, J., Thimmajanarthanan, M., Thomas, P.D., Wu, C.H., Yeats, C. & Yong, S.Y., (2012) InterPro in 2011: new developments in the family and domain prediction database. *Nucleic Acids Res* 40: D306-312.

Ieva, R. & Bernstein, H.D., (2009) Interaction of an autotransporter passenger domain with BamA during its translocation across the bacterial outer membrane. *Proc Natl Acad Sci U S A* 106: 19120-19125.

Ieva, R., Tian, P., Peterson, J.H. & Bernstein, H.D., (2011) Sequential and spatially restricted interactions of assembly factors with an autotransporter beta domain. *Proc Natl Acad Sci U S A* 108: E383-391.

Jain, E., Bairoch, A., Duvaud, S., Phan, I., Redaschi, N., Suzek, B.E., Martin, M.J., McGarvey, P. & Gasteiger, E., (2009) Infrastructure for the life sciences: design and implementation of the UniProt website. *BMC Bioinform* 10: 136.

Jain, S. & Goldberg, M.B., (2007) Requirement for YaeT in the outer membrane assembly of autotransporter proteins. *J Bacteriol* 189: 5393-5398.

Jong, W.S., ten Hagen-Jongman, C.M., den Blaauwen, T., Slotboom, D.J., Tame, J.R., Wickstrom, D., de Gier, J.W., Otto, B.R. & Luirink, J., (2007) Limited tolerance towards folded elements during secretion of the autotransporter Hbp. *Mol Microbiol* 63: 1524-1536.

Junker, M., Besingi, R.N. & Clark, P.L., (2009) Vectorial transport and folding of an autotransporter virulence protein during outer membrane secretion. *Mol Microbiol* 71: 1323-1332.

Junker, M., Schuster, C.C., McDonnell, A.V., Sorg, K.A., Finn, M.C., Berger, B. & Clark, P.L., (2006) Pertactin beta-helix folding mechanism suggests common themes for the secretion and folding of autotransporter proteins. *Proc Natl Acad Sci U S A* 103: 4918-4923.

Kajava, A.V. & Steven, A.C., (2006) The turn of the screw: variations of the abundant beta-solenoid motif in passenger domains of Type V secretory proteins. *J Struct Biol* 155: 306-315.

Kang'ethe, W. & Bernstein, H.D., (2013a) Charge-dependent secretion of an intrinsically disordered protein via the autotransporter pathway. *Proc Natl Acad Sci U S A* 110: E4246-4255.

Kang'ethe, W. & Bernstein, H.D., (2013b) Stepwise folding of an autotransporter passenger domain is not essential for its secretion. *J Biol Chem* 288: 35028-35038.

Kida, Y., Taira, J., Yamamoto, T., Higashimoto, Y. & Kuwano, K., (2013) EprS, an autotransporter protein of *Pseudomonas aeruginosa,* possessing serine protease activity induces inflammatory responses through protease-activated receptors. *Cell Microbiol* 15: 1168-1181.

Kingsley, R.A., Keestra, A.M., de Zoete, M.R. & Baumler, A.J., (2004) The ShdA adhesin binds to the cationic cradle of the fibronectin 13FnIII repeat module: evidence for molecular mimicry of heparin binding. *Mol Microbiol* 52: 345-355.

Kingsley, R.A., van Amsterdam, K., Kramer, N. & Baumler, A.J., (2000) The *shdA* gene is restricted to serotypes of *Salmonella enterica* subspecies I and contributes to efficient and prolonged fecal shedding. *Infect Immun* 68: 2720-2727.

Lett, M.C., 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.

Leyton, D.L., Johnson, M.D., Thapa, R., Huysmans, G.H., Dunstan, R.A., Celik, N., Shen, H.H., Loo, D., Belousoff, M.J., Purcell, A.W., Henderson, I.R., Beddoe, T., Rossjohn, J., Martin, L.L., Strugnell, R.A. & Lithgow, T., (2014) A mortise-tenon joint in the transmembrane domain modulates autotransporter assembly into bacterial outer membranes. *Nature Commun* 5: 4239.

Leyton, D.L., Rossiter, A.E. & Henderson, I.R., (2012) From self sufficiency to dependence: mechanisms and factors important for autotransporter biogenesis. *Nat Rev Microbiol* 10: 213-225.

Lugtenberg, B., Meijers, J., Peters, R., van der Hoek, P. & van Alphen, L., (1975) Electrophoretic resolution of the "major outer membrane protein" of *Escherichia coli* K12 into four bands. *FEBS letters* 58: 254-258.

Magrane, M. & Consortium, U., (2011) UniProt Knowledgebase: a hub of integrated protein data. *Database* 2011: bar009.

Makino, S., Sasakawa, C., Kamata, K., Kurata, T. & Yoshikawa, M., (1986) A genetic determinant required for continuous reinfection of adjacent cells on large plasmid in *S. flexneri* 2a. *Cell* 46: 551-555.

May, K.L. & 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.

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. & Van Den Bosch, L., (2003a) Lipopolysaccharide O antigen chains mask IcsA (VirG) in *Shigella flexneri*. *FEMS Microbiol Lett* 221: 173-180.

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

Noinaj, N., Kuszak, A.J., Balusek, C., Gumbart, J.C. & Buchanan, S.K., (2014) Lateral opening and exit pore formation are required for BamA function. *Structure* 22: 1055-1062.

Noinaj, N., Kuszak, A.J., Gumbart, J.C., Lukacik, P., Chang, H.S., Easley, N.C., Lithgow, T. & Buchanan, S.K., (2013) Structural insight into the biogenesis of beta-barrel membrane proteins. *Nature* 501: 385-390.

Pavlova, O., Peterson, J.H., Ieva, R. & Bernstein, H.D., (2013) Mechanistic link between beta barrel assembly and the initiation of autotransporter secretion. *Proc Natl Acad Sci U S A* 110: E938-947.

Peterson, J.H., Tian, P., Ieva, R., Dautin, N. & Bernstein, H.D., (2010) Secretion of a bacterial virulence factor is driven by the folding of a C-terminal segment. *Proc Natl Acad Sci U S A* 107: 17739-17744.

Pina-Pedrero, S., Olvera, A., Perez-Simo, M. & Bensaid, A., (2012) Genomic and antigenic characterization of monomeric autotransporters of *Haemophilus parasuis*: an ongoing process of reductive evolution. *Microbiol-SGM*158: 436-447.

Purdy, G.E., Fisher, C.R. & Payne, S.M., (2007) IcsA surface presentation in *Shigella flexneri* requires the periplasmic chaperones DegP, Skp, and SurA. *J Bacteriol* 189: 5566-5573.

Rawlings, N.D., Waller, M., Barrett, A.J. & Bateman, A., (2014) MEROPS: the database of proteolytic enzymes, their substrates and inhibitors. *Nucleic Acids Res* 42: D503-509.

Renn, J.P. & Clark, P.L., (2008) A conserved stable core structure in the passenger domain beta-helix of autotransporter virulence proteins. *Biopolymers* 89: 420-427.

Roman-Hernandez, G., Peterson, J.H. & Bernstein, H.D., (2014) Reconstitution of bacterial autotransporter assembly using purified components. *eLife* 3: e04234.

Rossiter, A.E., Leyton, D.L., Tveen-Jensen, K., Browning, D.F., Sevastsyanovich, Y., Knowles, T.J., Nichols, K.B., Cunningham, A.F., Overduin, M., Schembri, M.A. & Henderson, I.R., (2011) The essential beta-barrel assembly machinery complex components BamD and BamA are required for autotransporter biogenesis. *J Bacteriol* 193: 4250-4253.

Roy, A., Kucukural, A. & Zhang, Y., (2010) I-TASSER: a unified platform for automated protein structure and function prediction. *Nat Prot* 5: 725-738.

Ruiz-Perez, F. & Nataro, J.P., (2014) Bacterial serine proteases secreted by the autotransporter pathway: classification, specificity, and role in virulence. *Cell Mol Life Sci* 71: 745-770.

Schuster-Bockler, B., Schultz, J. & Rahmann, S., (2004) HMM Logos for visualization of protein families. *BMC Bioinform* 5: 7.

Shikata, S., Shimada, K., Kataoka, H., Horinouchi, S. & Beppu, T., (1992) Detection of large COOH-terminal domains processed from the precursor of *Serratia marcescens* serine protease in the outer membrane of *Escherichia coli*. *J Biochem* 111: 627-632.

Suzek, B.E., Huang, H., McGarvey, P., Mazumder, R. & Wu, C.H., (2007) UniRef: comprehensive and non-redundant UniProt reference clusters. *Bioinformatics* 23: 1282-1288.

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.

Teh, M.Y., Tran, E.N. & Morona, R., (2012) Absence of O antigen suppresses *Shigella flexneri* IcsA autochaperone region mutations. *Microbiol-SGM* 158: 2835-2850.

Tran, E.N., Doyle, M.T. & Morona, R., (2013) LPS unmasking of *Shigella flexneri* reveals preferential localisation of tagged outer membrane protease IcsP to septa and new poles. *PloS One* 8: e70508.

van den Berg, B., (2010) Crystal structure of a full-length autotransporter. *J Mol Biol* 396: 627-633.

Van den Bosch, L., Manning, P.A. & 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 Ulsen, P., van Alphen, L., ten Hove, J., Fransen, F., van der Ley, P. & Tommassen, J., (2003) A Neisserial autotransporter NalP modulating the processing of other autotransporters. *Mol Microbiol* 50: 1017-1030.

Wagner, J.K., Heindl, J.E., Gray, A.N., Jain, S. & Goldberg, M.B., (2009) Contribution of the periplasmic chaperone Skp to efficient presentation of the autotransporter IcsA on the surface of *Shigella flexneri*. *J Bacteriol* 191: 815-821.

Xu, D., Zhang, J., Roy, A. & Zhang, Y., (2011) Automated protein structure modeling in CASP9 by I-TASSER pipeline combined with QUARK-based ab initio folding and FG-MD-based structure refinement. *Proteins* 79 Suppl 10: 147-160.

Zhang, Y. & Skolnick, J., (2005) TM-align: a protein structure alignment algorithm based on the TM-score. *Nucleic Acids Res* 33: 2302-2309.

5.11. Article Supplementary Information

5.11.1. Supplementary Figures

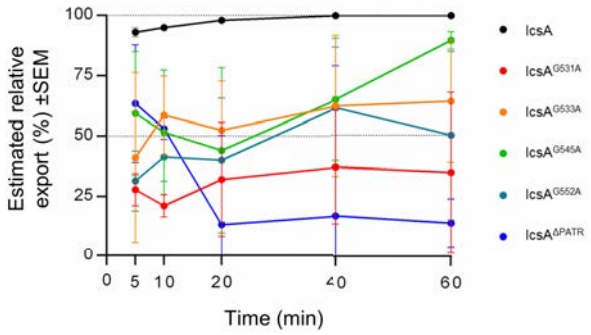


Figure 5.7[S]: Estimation of export from pulse-chases.

Densitometry estimation (see *Experimental Procedures*) of the pulse-chase bands in Figure 5.3 yields the above graph. All time points between 5 and 60 minutes were analyzed by repeated measures ANOVA (Dunnett's, $\alpha = 0.05$). 0 min and 1 min time points were not quantitated as bands were beyond the limits of detection. All IcsA PATR mutants had significantly lower estimated export compared to IcsA wild-type; IcsA^{G531A} $P < 0.0001$, IcsA^{G533A} $P = 0.0007$, IcsAG^{545A} $P = 0.0256$, IcsA^{G552A} $P = 0.0007$, IcsA$^{\Delta PATR}$ $P = 0.0186$. SEM = standard error of the mean, n = 2.

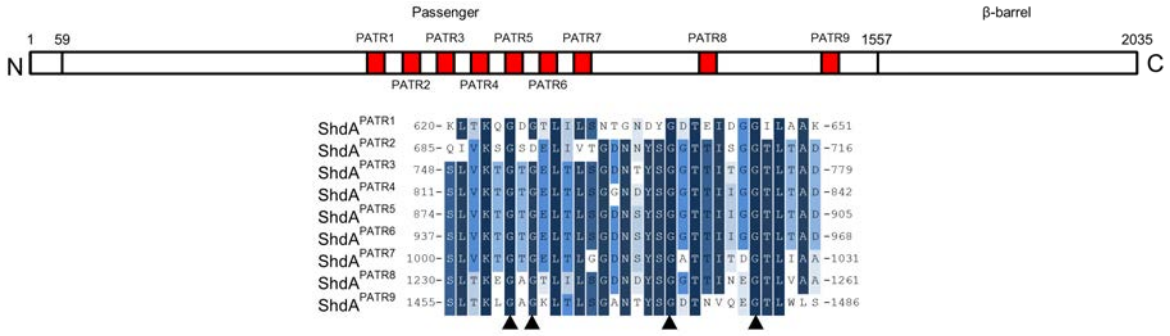


Figure 5.8[S]: The ShdA PATR array.

A scaled schematic of the AT ShdA (Q9XCJ4) is shown indicating the signal sequence (ShdA$^{1\text{-}59}$), the passenger (ShdA$^{60\text{-}1556}$), and the β-barrel (ShdA$^{1557\text{-}2035}$). The passenger has nine PATR (red) which are also aligned below. These were previously termed 'B regions' (Kingsley *et al*, 2000, Kingsley *et al*, 2004a). Conserved glycine residues investigated in this study are indicated by black arrows. Figure relates to Figure 5.4.

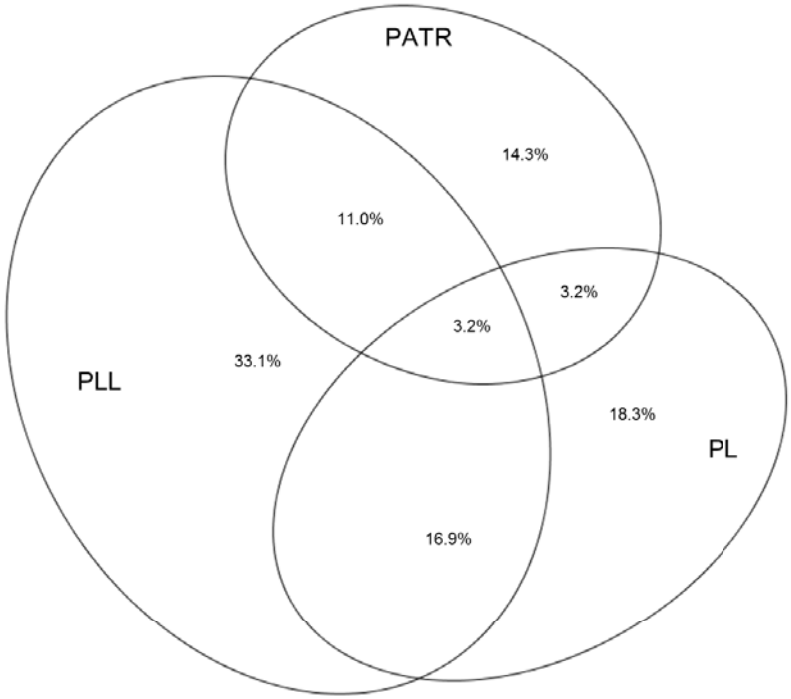


Figure 5.9[S]: Minimal overlap of PATR and Pertactin-like usage in passengers.

All unique representative ATs that contained a PATR or a Pertactin-like (PL) region or a Pectin lyase/P22-like (PLL) region (total of 3940 from 7659 ATs; 51.4 %) were plotted as a proportional Venn diagram to show overlap of usage of these passenger domain features (inclusive). Percentages are based on the PATR/PL/PLL group total. Figure relates to Figure 5.4.

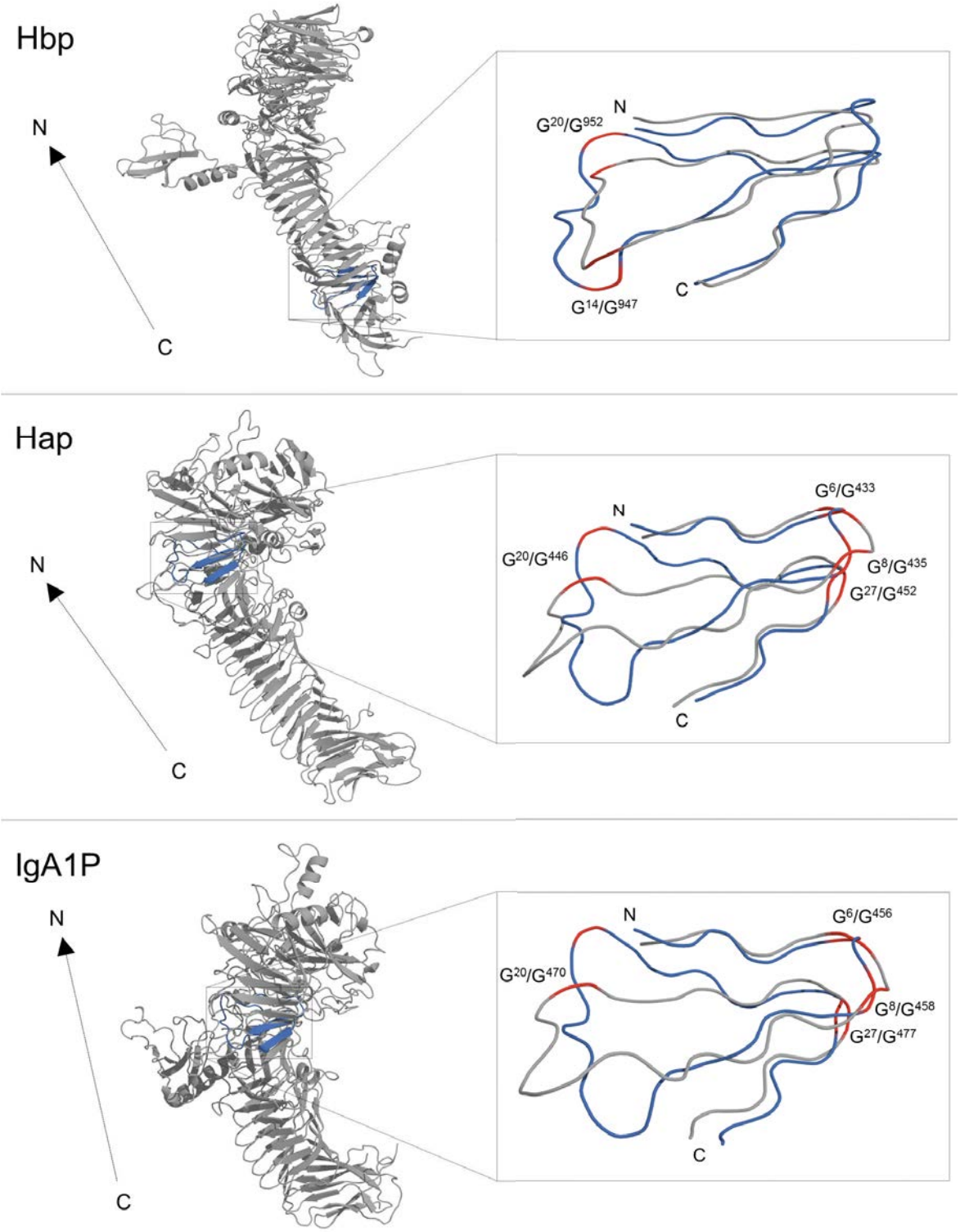


Figure 5.10[S]: Degenerate PATR sites in Hbp, Hap, and IgA1P.

The I-TASSER generated tertiary structure of the PATR was spatially aligned to all the solved AT passenger structures using TM-align. Identified degenerate PATR are shown in the hemoglobin-binding protease Hbp (O88093) from pathogenic *E. coli* (Otto *et al*, 2005), the fibronectin binding protease Hap (P45387) (Meng *et al*, 2011) and Immunoglobulin A1 Protease (P44969) (Johnson *et al*, 2009) from *Haemophilus influenzae*. Spatially conserved glycines (red) between the PATR (blue) and passenger (grey) are indicated. Figure relates to Figure 5.6.

5.11.2. Supplementary Tables

Table 5.1[S]: Strains and plasmids used in this study.

Strain or Plasmid	Description	Source
Strain		
2457T	Wild-type *Shigella flexneri* 2a	Laboratory collection
RMA2043	2457T ⊿*icsA::Tc*^{R} ⊿*rmlD::Km*^{R}	(Van den Bosch & Morona, 2003)
RMA2041	2457T ⊿*icsA::Tc*^{R}	(Van den Bosch & Morona, 2003)
RMA4378	2457T ⊿*icsA::Tc*^{R} *icsP::Km*^{R}	This study
Plasmids		
pCACTUS-*icsP*::*kan*^{R}	suicide vector, Cml^R, pCACTUS containing *icsP::kan*^{R} gene	(Tran *et al*, 2013)
pBR322	medium copy number, colE1 *ori*, Ap^R, Tc^R	(Bolivar *et al*, 1977)
pIcsA	pBR322 derivative containing *icsA* gene, P_{IcsA} promoter, Ap^R	(Morona & Van Den Bosch, 2003a)
pIcsA^{G531A}	pIcsA derivative, *icsA*^{G531A} allele	This study
pIcsA^{G533A}	pIcsA derivative, *icsA*^{G533A} allele	This study
pIcsA^{G545A}	pIcsA derivative, *icsA*^{G545A} allele	This study
pIcsA^{G552A}	pIcsA derivative, *icsA*^{G552A} allele	This study
pIcsA$^{\Delta PATR}$	pIcsA derivative, *icsA*$^{\sqcup 526\text{-}557}$ allele	This study
pBAD30	low copy number, pACYC184 *ori,* P_{BAD} promoter, arabinose inducible, glucose repressible, Ap^R	(Guzman *et al*, 1995)
pBADIcsA	pBAD30 derivative containing cloned *icsA* gene, P_{BAD} promoter, Ap^R	This study
pBADIcsA^{G531A}	pBADIcsA derivative, *icsA*^{G531A} allele	This study
pBADIcsA^{G533A}	pBADIcsA derivative, *icsA*^{G533A} allele	This study
pBADIcsA^{G545A}	pBADIcsA derivative, *icsA*^{G545A} allele	This study
pBADIcsA^{G552A}	pBADIcsA derivative, *icsA*^{G552A} allele	This study
pBADIcsA$^{\Delta PATR}$	pBADIcsA derivative, *icsA*$^{\sqcup 526\text{-}557}$ allele	This study

Tc^R = Tetracycline resistance, Km^R = Kanamycin resistance, Cml^R = Chloramphenicol resistance, Ap^R = Ampicillin resistance

Table 5.2[S]: Identification of degenerate PATR sites from solved passenger structures.

UniProtKB	PDB	Name	Start	Sequence[α]	End	TM-Score[β]
Q8CVR0	4KH3	Ag43	494	SVEKSGSGTLTVSNTTLTQKAVNLNEGTLTLN	525	**0.532**
O88093	1WXR	Hbp	935	TVSMTDTQWSMNGNSTAGNMKLNRTIVGFN	964	**0.504**
P45387	3SYJ	Hap	428	RLSKIGKGTLHVQAKGENKGSISVGDGKVILE	458	**0.501**
P44969	3H09	IgA1P	451	RLAKIGKGTLIVEGKGENKGSLKVGDGTVILK	482	**0.500**
P14283	1DAB	Pertactin	432	SLSIDNATWVMTDNSNVGALRLASDGSVDFQ	462	0.497
O68900	4OM9	Pet	734	ALTIHKGANVTASSSLFTTSNIKSGGDLTLT	764	0.470
Q7BSW5	3SZE	EspP	735	KATIENGADVTTQSGMFSTSDISISGNLSMT	765	0.462
Q9XD84	4Q1Q	TibA	214	DTVLNNTGRQFVSSGGSAAKTTINSGGMYLY	245	0.447
Q48245	2QV3	VacA	488	DISLGRFVNLKVDAHTANFKGIDTGNGGFNTLDFSG	523	0.440
		Consensus PATR		GLTKTGAGTLTLSGNNTYTGGTTVNAGTLQVG		

[α] Sites found using TM-align (Zhang & Skolnick, 2005).
[β]SCOP/CATH protein folds are shared when TM-score >0.5 (bold)

Table 5.3[S]: Oligonucleotides used in this study.

Code	Sequence (5' to 3')	Use
md17	GGTACCAGGAGGAAACGATGAATCAAATTCACAAATTTTTTTGTAATATGA	pBADIcsA construction
md18	GTCGACTCAGAAGGTATATTTCACACCC	pBADIcsA construction
md27	CCCTAACAAAATCAGGGGAGGCAACTCTCATTTTGGC	QuickChange G^{533}→A
md28	GCCAAAATGAGAGTTGCCTCCCCTGATTTTGTTAGGG	QuickChange G^{533}→A
md62	ACAGTTGAAGCTATGACACG	ΔPATR
md63	TTTTCCGTCCCAGCTATTTATA	ΔPATR
md74	CCCTAACAAAATCAGCGGAGGGAACTCTCATTTTGGC	QuickChange G^{531}→A
md75	GCCAAAATGAGAGTTCCCTCCGCTGATTTTGTTAGGG	QuickChange G^{531}→A
md76	CCTACTCTGCTTTCACCAACATCAATGCAGGC	QuickChange G^{545}→A
md77	GCCTGCATTGATGTTGGTGAAAGCAGAGTAGG	QuickChange G^{545}→A
md78	CTCTGGTTTCACCAACATCAATGCAGCCATT	QuickChange G^{552}→A
md79	GAATGGCTGCATTGATGTTGGTGAAACCAGAG	QuickChange G^{552}→A

5.11.3. Supplementary Methods

5.11.3.1. Construction of S. flexneri icsA- icsP-.

The suicide vector pCACTUS-*icsP*::*kan*^{R} (see Table 5.1[S]) containing *icsP* gene disrupted by a *kan*^{R} cassette, was transformed into an *S. flexneri* 2457T ⊿*icsA*::*Tc*^{R} strain (see Table 5.1[S]) by electroporation. Allelic exchange mutagenesis of the *icsP* gene was conducted as previously described (Morona *et al*, 1995, Tran *et al*, 2013).

5.11.3.2. Construction of pBADIcsA.

The *icsA* gene was placed under the control of and an arabinose/glucose switch present in the pBAD30 vector. First the *icsA* ORF was PCR amplified from pIcsA (see Table 5.1[S]) using primers md17 and md18 (see Table 5.3[S]) introducing a 5' RBS and flanking KpnI/SalI sites. This amplicon was cloned into pGEM-T Easy (Promega) creating intermediate plasmid pGEM-T::*RBS-icsA*. This was then digested with EcoRI and SalI and fragment containing *icsA* was ligated into similarly digested pBAD30 creating pBADIcsA.

5.11.3.3. Site-directed Mutagenesis of pIcsA and pBADIcsA.

To generate pIcsA derivatives pIcsA^{G531A}, pIcsA^{G533A}, pIcsA^{G545A}, pIcsA^{G552A}, and pBADIcsA derivatives pBADIcsA^{G531A}, pBADIcsA^{G533A}, pBADIcsA^{G545A}, and pBADIcsA^{G552A}, site-directed mutagenesis was performed according to the QuikChange Lightening Site-directed Mutagenesis kit (Agilent Technologies) protocol using QuikChange oligonucleotides described in Table 5.3[S]. All derivatives were confirmed by DNA sequencing.

5.11.3.4. Deletion of PATR sites in pIcsA and pBADIcsA.

To generate pIcsA$^{\Delta PATR}$ and pBADIcsA$^{\Delta PATR}$ pIcsA and pBADIcsA were subjected to inverse PCR amplification using primers md62 and md63 (see Table 5.3[S]) producing blunt and linearized plasmids lacking the PATR encoding sequence. These linear plasmids were then re-circularized by blunt-end self-ligation producing pIcsA$^{\Delta PATr}$ and pBADIcsA$^{\Delta PATr}$ with in frame deletions of the PATR. These derivatives were confirmed by DNA sequencing.

5.11.4. Supplementary References

Bolivar, F., Rodriguez, R.L., Greene, P.J., Betlach, M.C., Heyneker, H.L., Boyer, H.W., et al. (1977) Construction and characterization of new cloning vehicles. II. A multipurpose cloning system. *Gene* 2: 95-113.

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

Johnson, T.A., Qiu, J., Plaut, A.G. and Holyoak, T. (2009) Active-site gating regulates substrate selectivity in a chymotrypsin-like serine protease the structure of *Haemophilus influenzae* immunoglobulin A1 protease. *J Mol Biol* 389: 559-574.

Kingsley, R.A., Keestra, A.M., de Zoete, M.R. and Baumler, A.J. (2004) The ShdA adhesin binds to the cationic cradle of the fibronectin 13FnIII repeat module: evidence for molecular mimicry of heparin binding. *Mol Microbiol* 52: 345-355.

Kingsley, R.A., van Amsterdam, K., Kramer, N. and Baumler, A.J. (2000) The *shdA* gene is restricted to serotypes of *Salmonella enterica* subspecies I and contributes to efficient and prolonged fecal shedding. *Infect and Immun* 68: 2720-2727.

Meng, G., Spahich, N., Kenjale, R., Waksman, G. and St Geme, J.W., 3rd (2011) Crystal structure of the *Haemophilus influenzae* Hap adhesin reveals an intercellular oligomerization mechanism for bacterial aggregation. *Embo J* 30: 3864-3874.

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

Otto, B.R., Sijbrandi, R., Luirink, J., Oudega, B., Heddle, J.G., Mizutani, K., et al. (2005) Crystal structure of hemoglobin protease, a heme binding autotransporter protein from pathogenic *Escherichia coli*. *J Biol Chem* 280: 17339-17345.

Tran, E.N., Doyle, M.T. and Morona, R. (2013) LPS unmasking of *Shigella flexneri* reveals preferential localisation of tagged outer membrane protease IcsP to septa and new poles. *PloS One* 8: e70508.

Van den Bosch, L. and 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.

Zhang, Y. and Skolnick, J. (2005) TM-align: a protein structure alignment algorithm based on the TM-score. *Nucleic Acids Res* 33: 2302-2309

5.12. Additional Results – 1

Results presented in this chapter thus far were all included in the final accepted publication. The following additional results (Additional Results – 1 [section 5.12] and Additional Results – 2 [section 5.13] were not included in the publication, but nevertheless continue on the topics of the PATR motif and the importance of the passenger domain in general.

5.12.1. Analysis of the Bacterial Representation of the PATR.

Although individual important PATR-type AT proteins were identified in various important pathogens (Figure 5.4A), I remained interested in how widely the PATR motif is used in different bacteria. To gain an understanding of the broader use of PATR-type ATs in Gram negative bacteria, and to determine whether there is a difference in the usage of PATR-type and non-PATR ATs between different types of bacteria, the UniProt AT lists previously used for domain combination analyses were used to extract genus metadata for each entry resulting in Figure 5.11[A].

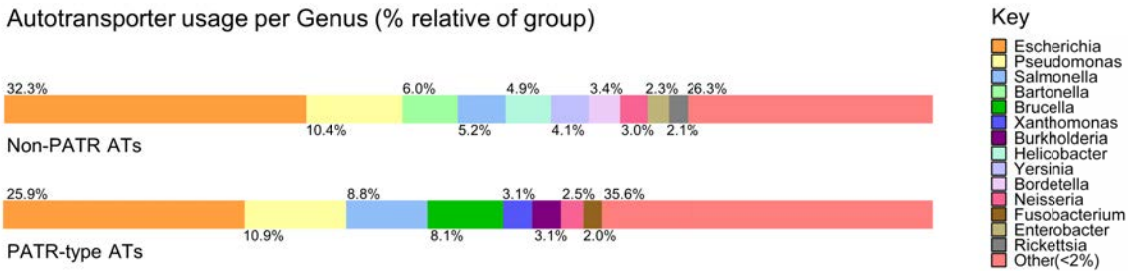


Figure 5.11[A]: PATR-type / non-PATR AT usage in Gram negatives.

ATs within the UniProt Knowledgebase were assessed on their usage by different Gram negative genera and grouped based on the presence of at least one PATR motif (IPR013425). Each coloured bar represents the whole group. Other = genera that were < 2 % represented in either group. The dependency of the presence or absence of the PATR on the usage by different genera is significant ($P < 0.0001$, χ^2).

It is clear that the use of PATR-type AT is quite wide-spread and appears to be more diversely used than non-PATR AT ('others' being 35.6 % and 26.3 % respectively; Figure 5.11[A]). PATR-type and non-PATR ATs are similarly used by the *Pseudomonas* and *Neisseria* genera, however PATR-type ATs are less common than non-PATR ATs in the *Escherichia* genus. The largest differences were the high use of PATR-type ATs in *Brucella* (8.1 %), *Xanthomonas* (3.1 %), *Burkholderia* (3.1 %), and *Fusobacterium* (2 %) genera (compared to only 0.2 %, 0.1 %, 0.7 %, and 0.8% respectively for the non-PATR ATs). Conversely, non-

PATR ATs were more common in *Bartonella* (6 %), *Helicobacter* (4.9 %), *Yersinia* (4.1%), *Bordetella* (3.4 %), *Enterobacter* (2.3 %), and *Rickettsia* (2.1 %) genera (compared to 0.04 %, 0.04 %, 0.5 %, 1.9 %, 1 %, and 0 % respectively for the PATR-type ATs). Importantly, the differences in the presence/absence of the PATR on the usage by different genera is significant when subjected to contingency analysis ($P < 0.0001$, χ^2).

5.12.2. Propensity for Higher Isoelectric Point for PATR-type ATs.

It has been proposed that AT passengers generally have a net negative charge (low isoelectric point) and that this contributes to the passenger translocation stage of AT secretion possibly by some form of charge repulsion (Kang'ethe & Bernstein, 2013b). However, this trend was determined from only a small selection of AT sequences.

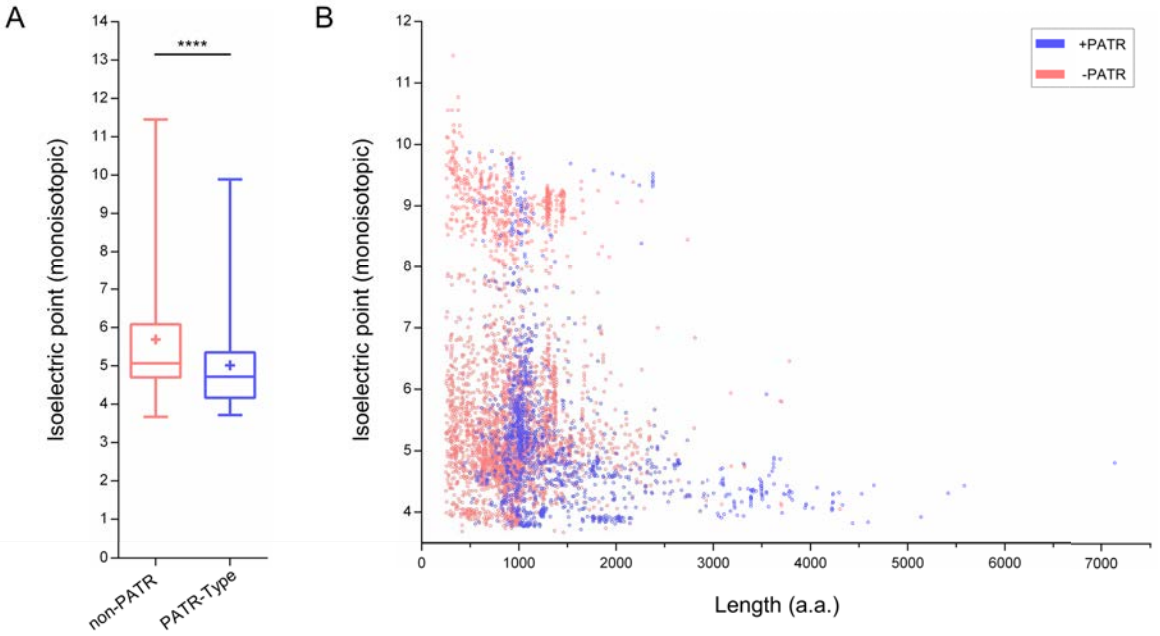


Figure 5.12[A]: PATR influence on isoelectric point.

(A) ATs within the UniProt Knowledgebase were grouped based on the presence of at least one PATR motif (IPR013425) and their monoisotopic isoelectric points calculated and compared. Box plots show median pI and interquartile range. Bars = pI data range, plus sign = mean pI. The difference in pI between PATR-type and non-PATR ATs is significant ($P < 0.0001$, Mann Whitney two-tailed test). **(B)** Isoelectric points (7659 data points) spread by protein length parameter.

The previously generated UniProt AT lists provided an opportunity to more thoroughly investigate this and to observe if there is a difference in isoelectric point (pI) for PATR-type and non-PATR ATs. Entry ID list were input into ExPASy-Compute pI/Mw tool to calculate monoisotopic pIs and the output data further analysed resulting in Figure 5.12[A]. While the ATs showed a wide range of isoelectric points (~3.5 to 11.5), the distribution of

isoelectric points is not Gaussian ($P =< 0.0001$, D'Agostino and Pearson omnibus normality test), and the vast majority were below a pI of 7 with an overall mean pI of 5.49 ±1.51. Further, there was a significant (although small) difference in isoelectric points when grouped into PATR-type and non-PATR ATs (PATR-type ATs having slightly lower pI; (Figure 5.12[A]A).As previously shown, PATR-type ATs are significantly larger than non-PATR ATs (Figure 5.4D). This is again seen in the scatter plot Figure 5.12[A]B, but intriguingly, the larger PATR-type entries commonly have a very low overall pI. It is also interesting that the scatter forms two distinct isoelectric point clusters – a high cluster ranging from pI ~8 – 10, and a low cluster ranging from pI ~7 – 3.5 (Figure 5.12[A]B). The high cluster is mostly comprised of non-PATR ATs, whereas the low cluster has high numbers of PATR-type ATs which is further skewed by length.

5.12.3. Summary of Additional Results – 1

Provided that there are limited biases (with respect to organism type or genus) in the entries in the UniProt database, ATs containing a PATR are present in a diverse range of Gram negative genera (Figure 5.11[A]). The under- and over-representations in PATR-type AT usage indicates that certain genera may be more amenable to a PATR-type secretion mechanism. Alternatively, I have shown that PATR-type ATs have a propensity to contain certain conserved functional elements (Figure 5.4 and Figure 5.5) – the differences in genera representations may reflect the specific function requirements for these different bacteria to survive in their environmental or pathological niche. Furthermore, it is not understood how very large AT passengers are efficiently translocated across the OM. If the charge of the passenger can be utilised in the translocation process (perhaps by repulsion from the negative charges of the inner core and lipid A moieties of lipopolysaccharide), then this might by exploited by very large PATR-type ATs during translocation and provide and explanation for the skew in Figure 5.12[A]B. Although I have not been able to test these assertions during the time of candidature, these results provide a basis for further bioinformatics analysis on PATR-type and non-PATR ATs. Specifically, I predict that it would be advantageous to further sort the UniProt AT entries by pathogenic function and pathogenic niche of the bacterium. The graph presented in Figure 5.12[A]B could then be coloured based on these groupings. This may further reveal the basis of both the clusters observed in Figure 5.12[A]B and the divergent roles of PATR-type and non-PATR ATs.

5.13. Additional Results – 2

The results presented in this chapter have accentuated the importance of the passenger in the process of Type V secretion. Although the AT β-barrel and the BAM complex are essential in the passenger translocation stage, certain portions of the passenger can significantly augment the efficiency of this process. Intrigued by this, I wished to observe the behaviour of the passenger in isolation of the β-barrel – again using IcsA as the model (IcsA passenger; IcsAp), as expression of the passenger alone may increase the likelihood of detecting interactions during nascent AT transit that are usually very fast.

5.13.1. Interaction of the Passenger and the Outer Membrane.

A C-terminally tagged (MycHis) linker insertion mutant IcsA passenger ($IcsAp_{i87}MycHis_6$) without the β-barrel had been expressed from pBAD/MycHisA in previous attempts at efficient purification from the periplasm (Teh 2012; and Morona laboratory unpublished) Interestingly, after expression of $IcsAp_{i87}MycHis_6$ in *E. coli* TOP10F', it could not be extracted from the periplasm but was found to be tightly associated with the OM (Teh, 2012). This phenomenon could represent a specific interaction between the passenger and components of the OM, or an artificial effect due to the insertion mutation, the tags, the added vector derived residues, or heterologous expression.

To further investigate this effect, a new construct was made by amplifying the IcsAp coding region for amino acids 1 – 758 (signal sequence plus passenger up to the IcsP cleavage site) using oligonucleotides md57/58 (Appendix A) to include a 5' KpnI site and ribosome binding site, and a 3' SalI site and stop codon. This amplicon was cloned into pGEMT-Easy (see 2.4.13) resulting in strain MDRM113 (DH5α[pGEMT-E::*icsAp*]) (see Appendix C). The pGEMT-E::*icsAp* construct was digested with KpnI/SalI and resulting fragment ligated into similarly digested pBAD30 vector resulting in pBAD30::*icsAp* (MDRM116; DH5α[pBAD30::*icsAp*]). This construct was transformed into a virulence plasmid positive *S. flexneri* ⊿*icsA::Tc*^{R} strain (RMA2041; see Appendix B) resulting in MDRM119 (RMA2041[pBAD30::*icsAp*]; see Appendix C) which was subsequently used for experimentation.

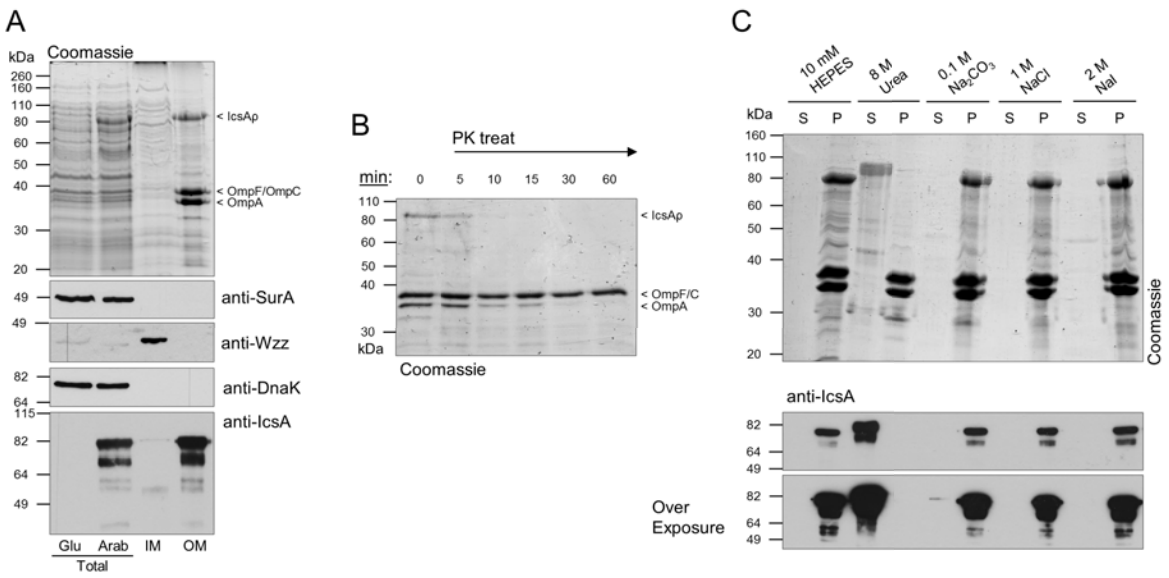


Figure 5.13[A]: Strong peripheral interaction of IcsA passenger with the outer membrane.

(A) Strain MDRM119 (*S. flexneri ⊿icsA::Tc*^{R} [pBAD30::*icsAp*]) was grown to log-phase in the presence of 0.2 % glucose (Glu), and then expression of the IcsA passenger (IcsAp = $IcsA_{1\text{-}758}$) was induced with 0.2 % arabinose (Arab) for 2 hours. Whole membrane was obtained and inner (IM) and outer (OM) membrane material was isolated by 2 % Triton X-100 treatment (see *Materials and Methods* sub-section 2.6.3). IcsAp retains association with the OM material. Major OMPs are indicated. SurA, Wzz, and DnaK serve as periplasmic, inner membrane and cytoplasmic controls respectively. Total = total bacterial protein sample. **(B)** Time-course of *in vitro* Proteinase K (PK) digestion of OM sample from panel **A**. OM samples was treated with 10 μg mL^{-1} of PK for 60 min with samples stopped by precipitation at time-points shown (see *Materials and Methods* sub-section 2.6.5). **(C)** Peripheral OMP disassociation assay (see *Materials and Methods* sub-section 2.6.4). Attempts at isolation of IcsAp from OM material using pH 7.5 HEPES buffer (control), denaturant urea solution, pH 11.3 Na_2CO_3 alkaline buffer, high concentration salt buffer, and chaotrophic ion buffer. IcsA western immunoblot over exposure shows low level dissociation with alkaline buffer treatment. S = supernatant, P = pellet.

Despite removal of extraneous tags and mutations, expression of IcsAp and subsequent isolation of the OM material by differential detergent treatment (see sub-section 2.6.3) revealed that the passenger is tightly associated with the OM (Figure 5.13[A]A). This shows that the direct or indirect interaction of the passenger to the OM occurs due to the IcsAp sequence itself. As expected, this interaction is peripheral and not integral since IcsAp is efficiently digested upon Proteinase K treatment (see sub-section 2.6.5) as opposed to the major integral OMPs which remain comparatively protease stable as they are protected by their integration into the OM (Figure 5.13[A]B).

In an attempt to disassociate IcsAp from the OM material, OM samples were treated with a variety of solutions (Figure 5.13[A]C). Although a very small population of IcsAp could be isolated upon alkaline treatment (anti-IcsA western immunoblot over exposure), the denaturing urea solution was the only treatment able to totally isolate IcsAp

from the OM (Figure 5.13[A]C). This indicates that the interaction between the OM and IcsAp is quite strong. Further, that buffers containing high salt or chaotrophic ions could not disassociate IcsAp indicates that electrostatic bonds are not responsible for this interaction.

5.13.2. Putative IcsAp Complexes with Outer Membrane Proteins.

To investigate if the IcsAp-OM interactions could be detected by chemical cross-linking, *S. flexneri ⊿icsA::Tc*^{R} [pBAD30::*icsAp*] was again grown, IcsAp expression induced, and then whole bacteria were chemically cross-linked *in situ* using DSP (see sub-section 2.6.8). Inner and outer membranes were then extracted and subjected to anti-IcsA western immunoblot analysis (Figure 5.14[A]). Multiple cross-linked products were visible, some of which were difficult to reverse even when heated to 100 °C in sample buffer containing β-mercaptoethanol indicating very strong interactions were involved. There were also very high molecular weight products in the cross-linked OM sample that were much greater than the 180 kDa marker.

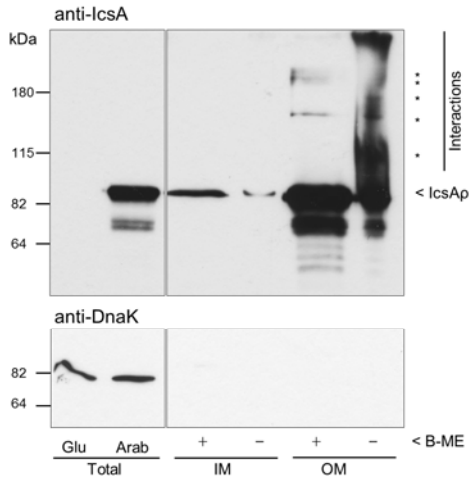


Figure 5.14[A]: Cross-linking reveals interactions of IcsA passenger with itself or other factors in the outer membrane.

Strain MDRM119 (*S. flexneri ⊿icsA::Tc*^{R} [pBAD30::*icsAp*]) was grown to log-phase in the presence of 0.2 % glucose (Glu), and then expression of the IcsA passenger (IcsAp) was induced with 0.2 % arabinose (Arab) for 2 hours. Bacteria were cross-linked *in situ* using 0.2 mM DSP, whole membrane was obtained, and inner (IM) and outer (OM) membrane was isolated by 2 % Triton X100 treatment (see *Materials and Methods* sub-sections 2.6.3 and 2.6.8). Samples were suspended in sample buffer with (+) or without (-) β-mercaptoethanol (β-ME) which reverses cross-links. Total = total bacterial protein samples, * = identified cross-linked products. DnaK serves as a cytoplasmic contamination control.

5.13.3. Summary of Additional Results – 2

It is possible that the higher order bands observed in Figure 5.14[A] represent complexes of IcsAp with periplasmic chaperones and OM translocation machineries. Regardless, results presented in Additional Results – 2 reveal that, in isolation from the β-barrel, the passenger is sufficient at localisation to the periphery of the OM (Figure 5.13[A]). This may be indicative of the early stages of transport such as chaperoning to the BAM complex. Although these claims could not be confirmed within the time of candidature, the results lay the foundation for further work to ascertain the identities of these interactors. For this, site specific cross-linking strategies could be used. It would also be valuable to introduce the PATR mutations investigated in this chapter into the pBAD30::*icsAp* construct. The mutants could then be investigated for effects on OM interaction and/or cross-linking pattern.

It should also be noted that the isolation of IcsAp from the OM using urea extraction (Figure 5.13[A]C) produced a sample that was approximately 85% pure for IcsAp and the other proteins have smaller molecular weights. Therefore, this is an ideal starting point for a method of non-tagged IcsAp purification. The IcsA passenger could be easily isolated from this sample using size exclusion chromatography and then refolded by dialysis. Although IcsAp purification was not one of my doctoral aims – this method is now being utilised in the Morona laboratory in attempts at purifying soluble IcsA passenger prior to refolding.

Chapter Six

ARTICLE 4:

A small conserved motif supports polarity augmentation of *Shigella flexneri* IcsA.

Matthew Thomas Doyle, Marcin Grabowicz and Renato Morona

Chapter 6: *Research Article 4*: A small conserved motif supports polarity augmentation of *Shigella flexneri* IcsA.

6.1. Statement of Authorship

Title of Paper	A small conserved motif supports polarity augmentation of *Shigella flexneri* IcsA.
Status	Submitted, in review (at time of Thesis submission). Now Accepted
Citation	Doyle MT, Grabowicz M, & Morona R (2015) A small conserved motif supports polarity augmentation of *Shigella flexneri* IcsA. *Microbiol-SGM* doi: 10.1099/mic.0.000165.

Author Contributions: By signing the Statement of Authorship, each author certifies that: (i) the candidate's stated contribution to the publication is accurate, (ii) permission is granted for the candidate in include the publication in the thesis; and (iii) the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

Author	Matthew Thomas Doyle		
Contribution	Conducted all experiments except those conducted by Marcin Grabowicz, conceived manuscript theme, constructed all figures, tables, and supplementary, and wrote the manuscript.		
Signature		Date	28/07/15

Author	Marcin Grabowicz		
Contribution	Constructed all IcsA-GFP fusions, conducted all live bacterial microscopy used in Figures 6.1 and 6.4, manuscript editing.		
Signature		Date	28/07/15

Author	Renato Morona		
Contribution	Supervised development of work, provision of laboratory and materials, manuscript evaluation and editing, acting corresponding author.		
Signature		Date	28/07/15

6.2. Purpose of Article

As established in the Introduction, IcsA itself plays a role in its direction to the *S. flexneri* pole. Two intramolecular polar targeting (PT) regions (one at the N-terminus and one central in the primary structure) have been attributed polar targeting functionality. These are thought to act in the cytoplasm before secretion. However, reports from our laboratory have cast doubt on the necessity of the N-terminal PT region on polar targeting. Furthermore, there was no knowledge as to the importance of specific PT region amino acid residues in the exertion of the polar targeting effect. Thus, this article describes a combination of mutagenesis, reporter protein, and quantitative microscopy methods used to focus in on the key functional portions of the PT regions.

6.2.1. Thesis Research Aims Addressed:

- **To investigate IcsA-dependent regulation of IcsA polarity.**

  - To re-examine the importance of the N-terminal and central Polar targeting regions (nPT; $IcsA_{1-104}$ and cPT; $IcsA_{506-620}$ respectively) in IcsA-dependent polarisation.

  - To identify amino acid residues required for appropriate IcsA-dependent polar targeting.

6.3. Article Summary

The rod-shaped enteric intracellular pathogen *Shigella flexneri*, and other *Shigellae* species, are the causative agents of bacillary dysentery. *S. flexneri* are able to spread within the epithelial lining of the gut resulting in lesion formation, cramps, and bloody stools. The outer membrane protein IcsA is essential for this spreading process. It is the initiator of an actin-based form of motility whereby it allows the formation of a filamentous actin 'tail' at the bacterial pole. Importantly, IcsA is specifically positioned at the bacterial pole such that this process occurs asymmetrically. The mechanism of IcsA polarity is not completely understood, but it appears to be a multifactorial process involving IcsA-dependent and IcsA-independent factors. In this study, we further investigate IcsA-dependent IcsA polarization by its intramolecular N-terminal and central polar targeting (PT) regions (nPT and cPT). The results obtained support a role in polar localisation for the cPT region and contend the role of nPT. We identify single residues that have measureable impacts on IcsA polarity augmentation, resulting in decreased *S. flexneri* spreading efficiency. Intriguingly, regions and residues involved in polar targeting clustered around a highly conserved motif which may provide a functional scaffold for polarity augmenting residues. How these results fit with current thoughts and models of IcsA-dependent polarity determination, are discussed.

6.4. Article Introduction

Shigella flexneri and other *Shigella* species are Gram negative obligate human pathogens and are the causative agents of shigellosis. Shigellosis can manifest itself as watery diarrhoea, or as painful dysentery with symptoms including: cramps, vomiting, and seizures (Niyogi, 2005, von Seidlein *et al*, 2006, Lima *et al*, 2015). Within this process, the outer membrane protein IcsA is a critical *Shigella* virulence factor required for adhesion, spreading and lesion formation (Makino *et al*, 1986, Bernardini *et al*, 1989, Lett *et al*, 1989, Suzuki *et al*, 2002, Brotcke Zumsteg *et al*, 2014). IcsA is an adhesin which contributes to initial bacteria-host epithelial cell interactions prior to invasion (Brotcke Zumsteg et al., 2014). Post-invasion, IcsA recruits host N-WASP protein and other actin polymerizing machinery to initiate actin based motility (Egile *et al*, 1999, Steinhauer *et al*, 1999, Snapper *et al*, 2001, Suzuki *et al*, 2002, Valencia-Gallardo *et al*, 2015).

Secreted by the type Va autotransporter pathway, IcsA has conserved primary structural features required for its biogenesis including an N-terminal signal peptide, a C-terminal β-barrel domain, and a central passenger domain (Leyton *et al*, 2012). Amino acid region $IcsA_{1-52}$ is an extended signal peptide for targeting to the periplasm and $IcsA_{759-1102}$ includes the β-barrel which inserts into the outer membrane and plays a key role in passenger secretion to the bacterial surface (Fukuda *et al*, 1995, Suzuki *et al*, 1995). $IcsA_{53-758}$ comprises the passenger domain and contains all information necessary for N-WASP binding (Suzuki *et al*, 2002). The passenger also contains a region of Pertactin-like homology ($IcsA_{650-736}$) and the recently characterised conserved PATR motif at $IcsA_{526-557}$ involved in passenger secretion efficiency (Doyle *et al*, 2015a).

Importantly, IcsA is localised to the surface of the old bacterial pole (pole which is not derived from the septum of the mother cell) allowing the asymmetric polymerisation and formation of filamentous (F)-actin 'comet tails' which enable intra- and intercellular spreading of *S. flexneri* in infected intestinal epithelia (Steinhauer *et al*, 1999, Robbins *et al*, 2001, Schroeder & Hilbi, 2008). Hence, all *Shigella* species require fine regulation of IcsA spatial asymmetry for appropriate functioning in motility. We have previously reported that the specific outer membrane protease of IcsA, IcsP, is targeted to the new cell pole and the septa of dividing *S. flexneri* (Tran *et al*, 2013); this results in asymmetric IcsA cleavage to refine IcsA surface polarity post secretion. Yet, while bacteria that are deficient in IcsP have increased IcsA along the lateral edge, they still have considerable levels of polar IcsA (Robbins

et al, 2001, Tran *et al*, 2013). Hence, IcsA polarity regulation is multifactorial with initial targeting to the pole coupled with continued maintenance of polarity post secretion.

The mechanism governing the initial set-up of IcsA polarity is not understood. It occurs before secretion with the help of the IcsA passenger itself which contains N-terminal polar targeting (nPT) and central PT (cPT) regions, $IcsA_{1-104}$ and $IcsA_{507-620}$ respectively (see Figure 6.1a) (Charles *et al*, 2001). These regions have been observed to direct fluorescent fusion proteins to the cell pole in *S. flexneri*, as well as various other Gram negatives including: *Escherichia coli*, *Salmonella typhimurium*, *Yersinia pseudotuberculosis*, and *Vibrio cholera* (Charles *et al*, 2001, Janakiraman *et al*, 2009). Indeed, heterologous expression of $IcsA_{506-620}$-GFP fusion proteins has frequently been used to study the nature of spatial protein targeting in bacteria (Janakiraman *et al*, 2009, Fixen *et al*, 2012). However, in previous mutagenesis studies of full-length IcsA, we have found that insertions (5-aa linkers) within the nPT region (positions i56, i81, and i87) were unable to perturb polar targeting, while insertions within the cPT region (positions i532 and i563) dramatically perturbed polar targeting resulting in circumferential distribution on the surface of bacteria (May & Morona, 2008). This data indicated that the cPT plays a more significant role in IcsA polar targeting.

In this study we investigate the importance of the IcsA PT regions in IcsA-dependent polarity determination using fusion protein analysis and directed mutagenesis. The results obtained support a role in polar localisation for the cPT region and contend the role of nPT. We identify residues within the cPT that, when substituted, dramatically affect the localisation of full length IcsA and the resulting associated virulence traits of the bacterium. We then further refine the region within the cPT that exerts polar targeting and find that the cPT residues cluster around a conserved element.

6.5. Article Methods

6.5.1. Bacterial strains, plasmids, and culture conditions.

Strains and plasmids used in this study are listed in Table 6.1. *S. flexneri* were routinely grown on Congo Red tryptic soy agar to confirm virulence plasmid presence before culture in Lysogeny–Broth (LB) media at 37°C with shaking. For all experiments, bacteria were sub-cultured to an OD_{600} = 0.5 before use. When required, broths were supplemented with antibiotics at the following concentrations; ampicillin (100 μg mL^{-1}), tetracycline (10 μg mL^{-1}), and kanamycin (50 μg mL^{-1}).

Table 6.1: Strains and plasmids used in this study.

Strain or Plasmid	Description	Source
Strain		
RMA2041	*S. flexneri* 2a 2457T ⊿*icsA*::Tc^R	(Van den Bosch & Morona, 2003)
DH5α	*E. coli* K-12 lab strain	Gibco-BRL
Plasmids		
pBR322	medium copy number, colE1 *ori*, Ap^R, Tc^R	(Bolivar *et al*, 1977)
pIcsA	pBR322 derivative containing *icsA* gene, P_{IcsA}, Ap^R	(Van den Bosch & Morona, 2003)
$pIcsA^{T563A}$	pIcsA derivative, $icsA^{T563A}$ allele	This study
$pIcsA^{R564A}$	pIcsA derivative, $icsA^{R564A}$ allele	This study
$pIcsA^{M562R}$	pIcsA derivative, $icsA^{M562R}$ allele	This study
$pIcsA^{W522R}$	pIcsA derivative, $icsA^{W522R}$ allele	This study
$pIcsA^{G531W}$	pIcsA derivative, $icsA^{G531W}$ allele	This study
pBBR1MCS2	pBBR1 replicon, *mob*+, Km^R	(Kovach *et al*, 1995)
$pWH1012gfp^+$	pBR322 derivative containing gfp^+, P_{Tet}, Ap^R	(Scholz *et al*, 2000)
pRMA2611	pBBR1MCS2::$P_{Tet}gfp^+$ C-terminal fusions, Km^R	This study
pRMA2611::$icsA_{1-104}$	$icsA_{1-104}$-gfp^+ fusion	This study
pRMA2611::$icsA_{53-104}$	$icsA_{53-104}$-gfp^+ fusion	This study
pRMA2611::$icsA_{53-505}$	$icsA_{53-505}$-gfp^+ fusion	This study
pRMA2611::$icsA_{105-505}$	$icsA_{105-505}$-gfp^+ fusion	This study
pRMA2611::$icsA_{105-620}$	$icsA_{105-620}$-gfp^+ fusion	This study
pRMA2611::$icsA_{506-620}$	$icsA_{506-620}$-gfp^+ fusion	This study
pRMA2611::$icsA_{532-620}$	$icsA_{532-620}$-gfp^+ fusion	This study
pRMA2611::$icsA_{538-620}$	$icsA_{538-620}$-gfp^+ fusion	This study
pRMA2611::$icsA_{506-562}$	$icsA_{506-562}$-gfp^+ fusion	This study
pRMA2611::$icsA_{506-570}$	$icsA_{506-570}$-gfp^+ fusion	This study
pRMA2611::$icsA_{532-570}$	$icsA_{532-570}$-gfp^+ fusion	This study
pRMA2611::$icsA_{538-562}$	$icsA_{538-562}$-gfp^+ fusion	This study

Tc^R = tetracycline resistance, Ap^R = ampicillin resistance, Km^R = kanamycin resistance

6.5.2. DNA methods and mutagenesis.

The C-terminal GFP$^+$ fusion vector pRMA2611 was constructed by ligating the *gfp*$^+$ bearing EcoRV-EcoRI fragment of pWH1012*gfp*$^+$ into similarly digested pBBR1MCS2. Amplification of *icsA* fragments from pIcsA using oligonucleotides (see Table 6.2) that introduced 5' NheI and 3' XbaI allowed in-frame cloning at the NheI site at nucleotide 5 of *gfp*$^+$ in pRMA2611. Fragments spanning the natural *icsA* XbaI site at nucleotide 304 were amplified with reverse oligonucleotides that introduced a 3' NheI site. To generate pIcsA derivatives pIcsA^{T563A} and pIcsA^{R564A}, site-directed mutagenesis was performed according to the QuikChange Lightening Site-directed Mutagenesis kit (Agilent Technologies) protocol using primer sets md29/30 and md31/32 (Table 6.2). To generate random mutants specifically over the PT region encoding portion of pIcsA, domain-targeted random mutagenesis was performed according to the GeneMorph II EZclone Domain Mutagenesis kit (Agilent Technologies) protocol using md5/6 primers (Table 6.2). The resultant random mutant pIcsA pool was transformed into RMA2041, colonies were manually screened by immunofluorescence (IF) microscopy (see below) for aberrant surface polarity, and resultant alleles identified by DNA sequencing. See Figure 6.5[S] for sample images from the screening process.

Table 6.2: Oligonucleotides used in this study.

Oligo	Sequence 5'-3'	Description / use
Mutagenesis		
md5	gagttcaatctgaaagaaaacgctgaacttact	Anneals immediately before *icsA* codon 506
md6	ccgacatttgacgaactattattgagaataacatg	Anneals immediately after *icsA* codon 620
md29	gacagttgaagctatggcacgtaccgctggtgttattg	QuickChange T^{563}→A
md30	caataacaccagcggtacgtgccatagcttcaactgtc	QuickChange T^{563}→A
md31	ggacagttgaagctatgacagctaccgctggtgttattg	QuickChange R^{564}→A
md32	caataacaccagcggtagctgtcatagcttcaactgtcc	QuickChange R^{564}→A
Fusions		
IcsA_1_Nhe	cacttactgataatatagtgcatggctagcattcac	IcsA$_{1-104}$
IcsA_53_Nhe	gctgctagcctttcgggtactcaagaacttcatttttc	IcsA$_{53-104}$, IcsA$_{53-505}$
IcsA_104_Nhe	cgtagcttgctagcagatgcatgagagg	IcsA$_{1-104}$, IcsA$_{53-104}$
IcsA_105_Nhe	gcatctagagctagcctacggattaac	IcsA$_{105-505}$, IcsA$_{105-620}$
IcsA_505_Nhe	ccagaatagtactaacagtaaggctagcgttttctttc	IcsA$_{53-505}$, IcsA$_{105-505}$
IcsA_505_Xba	ccagaatagtactaacagtaagtctagagttttctttc	IcsA$_{105-505}$
IcsA_506_Nhe	cgctgaagctagcgttagtactattctggc	IcsA$_{506-620}$, IcsA$_{506-562}$
IcsA_620_Xba	gaattctagaaccgtttttctccagag	IcsA$_{105-620}$, IcsA$_{506-620}$, IcsA$_{538-620}$, IcsA$_{532-620}$
IcsA_538_Nhe	ctctcgctagcgcggaaaaaaatacctactctgg	IcsA$_{538-620}$, IcsA$_{538-562}$
IcsA_562_Xba	gcggttctagacatagcttcaactgtccccat	IcsA$_{506-562}$, IcsA$_{538-562}$
IcsA_532_Nhe	caaaagctagcgagggaactctcattttggc	IcsA$_{532-620}$, IcsA$_{532-570}$
IcsA_570_Xba	gcacctctagaaacaataacaccagcggtacg	IcsA$_{506-570}$, IcsA$_{532-570}$

6.5.3. Antibodies.

Polyclonal rabbit anti-N-WASP and polyclonal rabbit anti-IcsA were produced and validated as described previously (Van Den Bosch *et al*, 1997, May & Morona, 2008). Monoclonal mouse anti-DnaK antibody was from Enzo Life Sciences. Polyclonal donkey anti-rabbit AlexaFluor 488 and donkey anti-rabbit AlexaFluor 594 antibodies were from Invitrogen. Horse radish peroxidase conjugated goat anti-mouse or goat anti-rabbit antibodies were used for western immunobloting and detected using chemiluminescence (Sigma).

6.5.4. Total bacterial protein samples.

5×10^8 bacteria from a log-phase culture were collected by centrifugation (16,000 × *g*, 1 min, 4 °C), resuspended in 100 µL of SDS-PAGE loading buffer (Lugtenberg *et al*, 1975) and heated to 100 °C for 10 min before SDS-PAGE and immunoblot analysis.

6.5.5. Bacterial immunofluorescence IcsA labelling.

Bacterial IF labelling was conducted as previously described (Doyle *et al*, 2015b). All solutions to be used were filtered through 0.22 µm nitrocellulose. 10^8 log-phase bacteria were harvested by centrifugation (16,000 x g, 2 min, 20 °C), fixed in 3.7 % (v/v) formaldehyde solution in phosphate buffered saline (PBS) (20 min, 20 °C), washed twice in PBS, and resuspended in 100 µL of PBS. Sterile round coverslips (at the bottom of a 24-well tray) that were pre-treated with 10 % poly-L-lysine solution (Sigma) were spotted with 5 µL of the fixed bacteria, centrifuged (775 x g, 5 min, 20 °C), and then incubated for 2 h at 20°C with anti-IcsA antibodies diluted 1:100 in PBS containing 10 % (v/v) fetal calf serum (FCS). Bacteria were then washed three times with PBS, incubated for 30 min at 37 °C with anti-rabbit Alexa Fluor 488 diluted 1:100 in PBS containing 10% (v/v) FCS, and washed three more times with PBS before mounting with 20 % Mowiol 4–88 (Calbiochem), 4 mg mL^{-1} *p*-phenylenediamine.

6.5.6. Cell infection and N-WASP/F-actin/DNA labelling.

Infection of semi-confluent HeLa cell monolayers with *S. flexneri* was conducted as described previously (Teh *et al*, 2012a, Doyle *et al*, 2015a). HeLa cells were grown on sterile round coverslips (at the bottom of 24 well trays). Log-phase bacteria were harvested by centrifugation (16,000 × *g*, 2 min, 20 °C) and diluted to 3×10^8 bacteria/ml in Dulbecco's PBS (D-PBS). HeLa Cells were washed with 10% (v/v) FCS in minimal essential medium (MEM), and then 80 µl of bacteria was added before centrifugation (500 × *g*, 5 min, 20 °C) to

assist adhesion and invasion. After incubation at 37 °C with 5 % CO2 for 1 h, cells were washed three times with D-PBS, and incubated a further 1.5 h with 500 µl of MEM supplemented with 10 % (v/v) FCS and 40 µg ml^{-1} gentamycin. Cells were then washed three times with D-PBS, treated for 15 min with 3.7 % (v/v) formaldehyde solution (Sigma) in PBS, and washed twice with PBS. Cells were incubated with 50 mM NH_4Cl in D-PBS for 10 min, washed with PBS, permeablized with 0.1 % (v/v) Triton X-100 in PBS for 5 min and washed with PBS. Cells were blocked with 10 % (v/v) FCS in PBS for 20 min, before incubation with anti-N-WASP diluted at 1:100 in PBS containing 10 % (v/v) FCS for 30 min at 37 °C. Cells were washed three times with PBS, and then incubated for 1 h at 37 °C with donkey anti-rabbit Alexa Flour 594 and Alexa Flour 488 phalloidin (Invitrogen) diluted to 1:100 and 1:200 respectively in PBS containing 10 % (v/v) FCS. These were then washed three times with PBS, DNA was stained with 10 µg ml^{-1} DAPI for 1 min, washed three times with PBS again, and mounted for microscopy as described above for bacterial IcsA labelling.

6.5.7. Microscopy and quantitation.

All microscopy images were captured using an Olympus IX-70 Microscope and MetaMorph software (v7.7.1.0, Molecular Devices) with a phase contrast 100 × oil immersion objective and a 1.5 × enlarger. For fluorescence imaging, an X-Cite 120Q lamp was used set at high intensity. Live bacterial imaging was conducted at 37 °C on custom made 1 % (w/v) LB-agarose solid media mounts before fluorescence imaging and polar foci counts. All bacterial IcsA and N-WASP fluorescence images were acquired with 100 and 500 millisecond exposures respectively. Fluorescence images for background correction were also taken using the same exposure times. MetaMorph line-scan measurement tools were used as previously described (Doyle *et al*, 2015b) to quantitate IcsA fluorescence intensities across the perpendicular axis of a point-to-point scan. Scans were conducted from pole-to-pole starting from intense pole, with scan width (perpendicular axis) equal to the bacterium. For each strain under investigation, cumulative scans were conducted on many bacteria (50 bacteria from each independent experiment '*n*') that were without a visible septum, resulting in distribution profiles representative of the population. Polarity was measured as the quotient of the old pole (first 25 % of the bacterium) IcsA intensity and the new pole (last 25 % of the bacterium) IcsA intensity. Representative IcsA images for qualitative presentation were recoloured using the ImageJ look up tables.

6.5.8. Plaque formation assay.

Plaque formation assays were conducted as described previously (Oaks *et al*, 1985, Teh & Morona, 2013). HeLa cells were grown overnight to a confluent monolayer in MEM with 10 % (v/v) FCS in a 6-well tray. Prior to infection monolayers were washed twice with D-PBS and once in Dulbecco's modified Eagle medium (DMEM). 10^8 log-phase bacteria were harvested (16,000 × *g*, 2 min, 20 °C), and resuspended in 1 mL of DMEM and serial diluted to 10^5 cell mL^{-1} in DMEM before 200 µL was added to each well. Trays were incubated (37 °C, 5 % CO_2) with gentle rocking every 15 min. At 120 min post-infection, the wells were aspirated and 3 mL of DMEM, 5 % (v/v) FCS, 20 µg mL^{-1} gentamycin, 0.5 % (w/v) agarose (Seakem ME) overlay mixture was added to each well before further incubation (37 °C, 5 % CO_2) to set. The second overlay (DMEM, 5 % (v/v) FCS, 20 µg mL^{-1} gentamycin, 0.5 % (w/v) agarose, 0.1 % (w/v) Neutral Red [MP Biomedicals]) was added at 48 h post-infection before further incubation at 37 °C with CO_2 (5 %). Wells were photographed at 70 h post infection from a standardized height and plaque diameter quantitated using MetaMorph image analysis software.

6.5.9. IcsA passenger structural modelling.

Modelling of IcsA was achieved using the I-TASSER server for protein structure and function prediction (zhanglab.ccmb.med.umich.edu/I-TASSER) which uses threading and *ab initio* modelling (Roy *et al*, 2010, Xu *et al*, 2011). In order to model only the passenger domain of IcsA, the amino acid input range was $IcsA_{52\text{-}735}$. The model had a high C-score of 0.15 (possible range is -5 to 2 where higher values signifies higher confidence in the model), and a TM-score of 0.73 ± 0.11 (> 0.5 indicates correct topology, < 0.17 indicates random similarity). The following templates contributed to the model: Hap (PDB 3SYJ), Pertactin (PDB 1DAB), IcsA autochaperone (PDB 3ML3), Hbp (PDB 1WXR), and EspP (PDB 3SZE).

6.6.1. Relative contribution of nPT and cPT to IcsA polar targeting

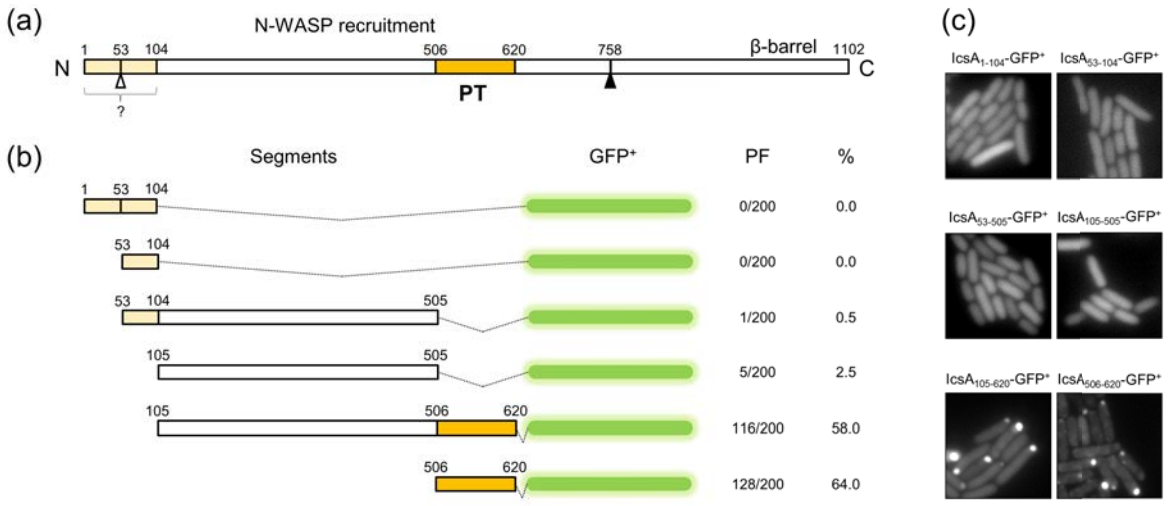


Figure 6.1: Analysis of nPT and cPT regions in IcsA.

(a) Scaled schematic of IcsA primary structure. The region $IcsA_{1-52}$ is a signal sequence with open arrow head indicating signal peptidase cleavage site, $IcsA_{53-758}$ is the passenger domain containing N-WASP recruitment functionality, the passenger can be cleaved by IcsP from the rest of the protein that contains the β-barrel between positions R^{758} and R^{759}. The polar targeting (PT) region $IcsA_{506-620}$ (cPT) is indicated in orange. nPT = an additional region also suggested to have polar targeting properties. (b) Various IcsA-GFP$^+$ fusions and resulting polar foci (PF) frequencies (n = 200) during expression in *E. coli* with representative fluorescence micrographs shown in (c). Each micrograph is 6 µm by 6 µm.

A previous study had fused nPT or cPT to GFP and shown that either region was sufficient for directing polar localisation of the fusion protein. More recently, the nPT was shown to be remarkably permissive for insertion mutations that did not affect polar localisation of full-length IcsA (May & Morona, 2008). Conversely, insertions in cPT could disrupt the polar localisation of IcsA, despite the presence of an intact nPT (May & Morona, 2008). Collectively, these data questioned the contribution of nPT for determining IcsA polar localisation. To advance the investigation of IcsA polarity and the determinants that drive polar targeting, these inconsistencies needed to be resolved. To this end, we designed a series of IcsA fusions to improved GFP (GFP$^+$) (Scholz *et al*, 2000) and evaluated their cell polarity upon expression in *E. coli* (Figure 6.1). $IcsA_{1-104}$-GFP$^+$ (nPT) did not form polar foci in our experimental system (Figure 6.1b and c). We also tested $IcsA_{53-104}$-GFP$^+$ and $IcsA_{53-505}$-GFP$^+$ reasoning that the nPT may act in synergy with a larger portion of the passenger domain. However, no bacteria exhibited polar foci upon expression of $IcsA_{53-104}$-GFP$^+$, and only 0.5 % of bacteria exhibited polar foci upon expression of $IcsA_{53-505}$-GFP$^+$ (Figure 6.1b). Indeed, even the control fusion $IcsA_{105-505}$-GFP$^+$ lacking both PT regions produced a higher number of

polar foci (2.5 %) (Fig. 1b). Only when the fusion contained the cPT ($IcsA_{105-620}$-GFP^+) did we observe a propensity for polar foci formation (58 % of bacteria) (Figure 6.1b). The fusion of the cPT itself to GFP^+ ($IcsA_{506-620}$-GFP^+) produced the highest proportion of polar foci positive cells (64 %) from all the fluorescent fusions tested (Figure 6.1b).

We also note that the frequency of $IcsA_{506-620}$-GFP^+ foci formation observed here was lower than that previously reported for this region (Charles *et al*, 2001). These discrepancies may be explained by our use of *gfp*$^+$ for IcsA fusions, over the previously used *gfpmut2*, as GFP^+ has folding optimized mutations (Scholz *et al*, 2000). It has been reported that $IcsA_{507-620}$-GFPmut2 is only 30 % soluble (Rokney *et al*, 2009), therefore, the use of folding optimized GFP^+ may have limited the likelihood of nonspecific aggregation due to unfolding and increased the stringency of our experiments. These data, and that of previous studies (May & Morona, 2008), strongly support the cPT as the primary region of IcsA that mediates polarity.

6.6.2. Identification of cPT residues important for driving IcsA polarity

We next focused our attention on determining important polarity augmenting residues within the cPT region. We first conducted domain-targeted random mutagenesis on the cPT region of IcsA while retaining the fidelity of the rest of the encoded protein (see *Methods*). The plasmid pIcsA (Van den Bosch & Morona, 2003), which contains the *icsA* gene, was used as the target for this mutagenesis. Generated random pIcsA cPT mutant plasmids were then transformed into an *S. flexneri* 2a *icsA*$^-$ strain (RMA2041) and 100 IcsA cPT putative mutant transformants (PT1 to PT100) were screened individually via IF microscopy for defects in IcsA surface localisation. 9 of the 100 transformants were found to have altered IcsA localisation compared to wild-type IcsA ($IcsA^{WT}$) (See Figure 6.5[S] sample IF images obtained during screening). The pIcsA cPT mutant plasmids from these isolates were subsequently isolated and sequenced to identify encoded substitutions. These were then mapped on the amino acid sequence of the cPT region as a hit frequency plot (Figure 6.2a).

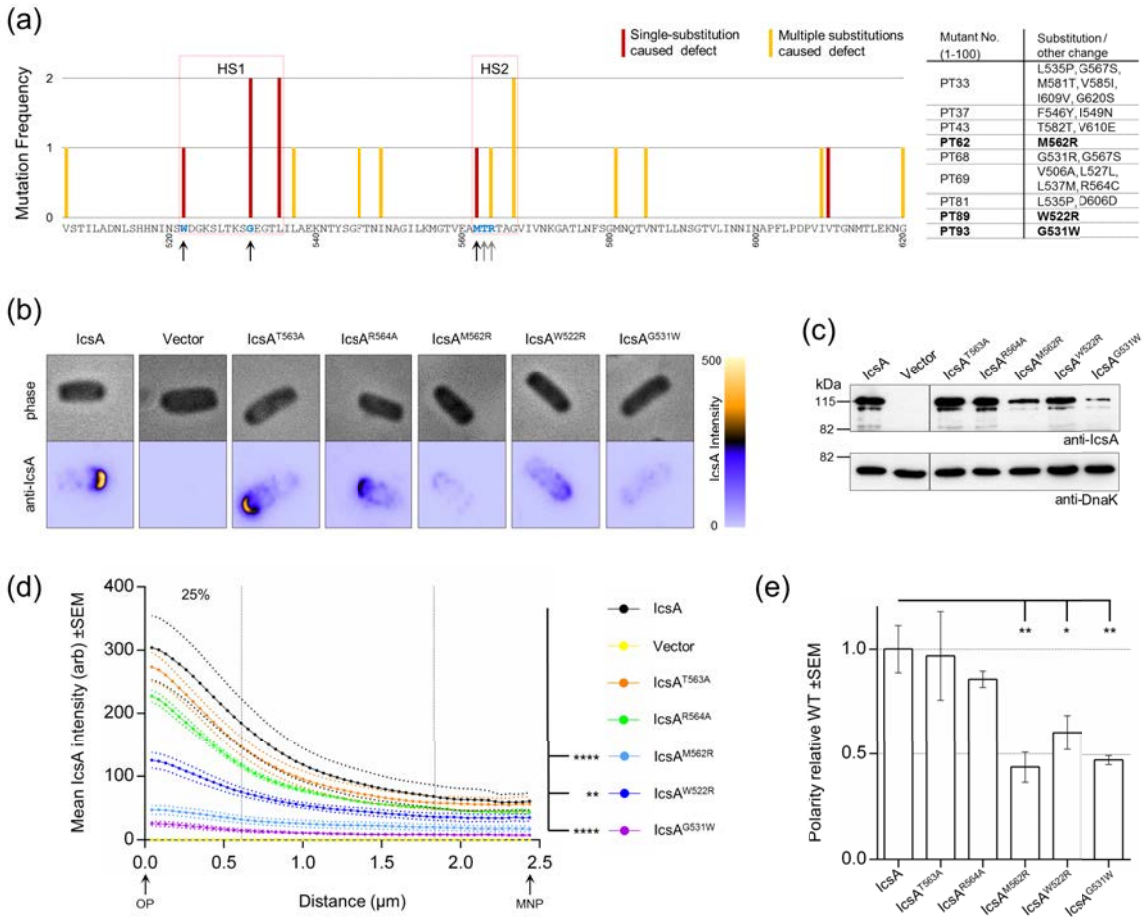


Figure 6.2: Identification of residues in the cPT effecting IcsA localisation.

All panels show data generated using *S. flexneri* ⊿*icsA* (RMA2041) complemented with base vector pBR322, pIcsA, or pIcsA cPT mutant derivatives. (a) Positions of important polar targeting residues identified by cPT region random mutagenesis and subsequent anti-IcsA IF screening (for representative images from screening stage see Figure S1 in *Supplementary Data*). Positions of proposed hot-spots (HS) with polar targeting function are also indicated. Derived mutant variants are shown tabulated on the right. Residues to be further characterised are indicated in blue, and black and grey arrows indicate point mutants derived from random and site-directed mutagenesis respectively. (b) Phase (top) and anti-IcsA IF micrographs (bottom) of representative bacteria with fluorescence intensities scaled equally relative to each other. Each micrograph is 4 μm by 4 μm. (c) anti-IcsA western immunoblots of total protein samples from bacteria expressing IcsA PT region point mutants. DnaK serves as a loading control, $n = 3$. IF experiments were repeated ($n = 3$–5) and IcsA surface distributions were measured for each strain on a population basis producing IcsA distribution profiles (d) that was further analysed to assess IcsA polarity values (e). Differences in polar IcsA intensity over the first 25 % of the cell (d) or polarity relative wild-type (WT) IcsA (e) were analysed by one-way ANOVA with Dunnett's tests (* = $P < 0.05$, ** = $P < 0.01$, **** = $P < 0.0001$). SEM = standard error, OP = old pole, MNP = mean new pole.

These mutations appeared to cluster as two 'hot-spots' (HS) spanning residues $IcsA_{522\text{-}535}$ (HS1), and residues $IcsA_{562\text{-}567}$ (HS2) that, coincidentally, overlapped 5-aa insertion sites (i532 and i563 respectively) previously found to disrupt polar localisation (May & Morona, 2008). Interestingly, single substitutions were identified that were able to cause dramatic effects on IcsA localisation. These included $IcsA^{M562R}$, $IcsA^{W522R}$, and $IcsA^{G531W}$ pictured in Figure 6.2b. As controls, we specifically substituted $IcsA^{T563}$ and $IcsA^{R564}$ to alanine

to assess involvement of residues close to IcsA^{M562}. IcsA^{T563A} had indistinguishable localisation compared to IcsA^{WT}, and IcsA^{R564A} exhibited only a minor decrease in polar IcsA (Figure 6.2b). In terms of total cellular protein levels IcsA^{T563A}, IcsA^{R564A}, and IcsA^{W522R} each had equivalent levels to IcsA^{WT} (Figure 6.2c) when examined by western immunobloting. IcsA^{M562R} and IcsA^{G531W}, however had decreased overall levels in the bacteria comparatively to IcsA^{WT} indicating a defect in protein stability (Figure 6.2c).

To quantitate the importance of each of these substitutions on IcsA's resultant surface localisation, we modelled IcsA surface distribution profiles on a population basis using anti-IcsA IF microscopy and multi-cell intensity scanning (see *Methods*). The resultant surface distributions for the IcsA PT mutants were varied (Figure 6.2d and e). The distribution profiles for IcsA^{WT} and IcsA^{T563A} were similar and were not significantly different with regards to IcsA at the pole (Figure 6.2d) and polarity value (Figure 6.2e). There was also no significant difference for IcsA^{R564A}, yet there was a trend of decreased IcsA at the pole (Figure 6.2d). IcsA^{M562R}, IcsA^{W522R}, and IcsA^{G531W} all had significant decreases (83.7 %, 59.5 %, and 91.7 % decreases respectively) in IcsA at the pole compared to IcsA^{WT} (Figure 6.2d). These also displayed polarity values that were generally half that of IcsA^{WT} (Figure 6.2e). Interestingly, mutations that caused decreases in total IcsA cellular levels (IcsA^{M562R} and IcsA^{G531W}; Figure 6.2c) generally appeared to coincide with larger changes in polarity (Figure 6.2e) indicating a concentration dependence.

We also tested the ability of the set of IcsA cPT mutants to form plaques on HeLa cell monolayers which is a strong correlate to *Shigella* spreading and lesion formation in colonic epithelia. We hypothesized that reductions in plaque size would be responsive to reductions in IcsA polarity. As expected, bacteria expressing IcsA^{T563A} produced plaques of equivalent size to IcsA^{WT} (Figure 6.3a and b). However, IcsA^{R564A} and IcsA^{W522R} each displayed significant reductions in spreading (24.7 % and 44.7 % reductions respectively) compared to IcsA^{WT} (Figure 6.3a and b). Over the period of this experiment, plaques were not visible for both IcsA^{M562R} and IcsA^{G531W} (Figure 6.3a and b) which were shown to have the most severe changes in surface localisation (Figure 6.2). Upon comparing the size of these plaques to our polarity quantitation, we were able to demonstrate that decreasing polar IcsA or polarity metrics correlates significantly to diminishing plaque diameter (Figure 6.3c). These correlations were well fitted by linear models with strong correlation coefficients. We also conducted N-WASP recruitment assays to determine if the functional differences were due to the effects on surface localisation and not an inability to bind N-WASP. All strains expressing IcsA PT mutants were able to bind N-WASP except IcsA^{M562R} (Figure 6.3, bottom panels). This

strongly suggests that the effects on bacterial spreading by all IcsA PT mutants (except IcsA^{M562R}) were due to the changes in surface localisation. Indeed, the formation of normal actin comet tails appeared to be most perturbed for IcsA PT mutants with lower polarity (Figure 6.3d, top panels).

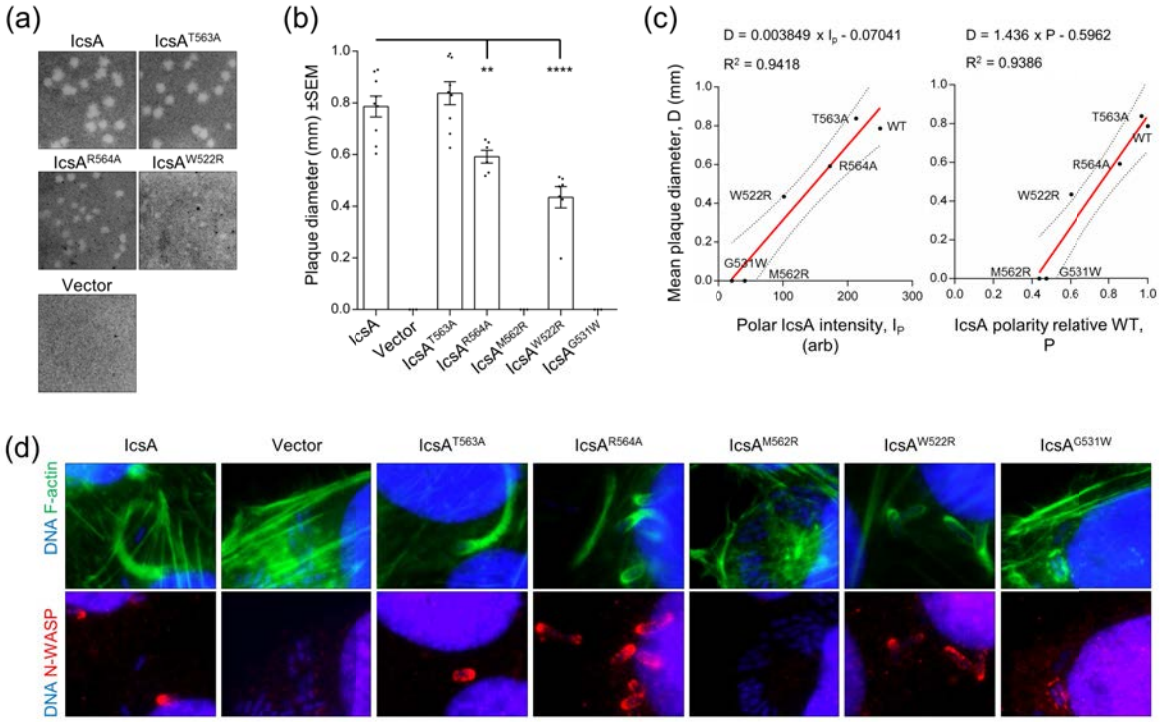


Figure 6.3: Disrupted IcsA polarity decreases *S. flexneri* virulence indicators.

All panels show data generated using *S. flexneri* ⊿*icsA* (RMA2041) complemented with base vector pBR322, pIcsA, or pIcsA cPT mutant derivatives. (a) Plaques in confluent HeLa cell monolayers due to cell-to-cell bacterial spreading. (b) Plaque diameters were measured (n = 6-9) and differences analysed by one-way ANOVA with Dunnett's tests (** = $P < 0.01$, **** = $P < 0.0001$). No plaques were observed for the vector control, IcsA^{M562R}, and IcsA^{G531W}. SEM = standard error. (c) Mean plaque diameters (D) for IcsA and IcsA cPT mutants correlates significantly with polar IcsA intensity (I_P; two-tailed P = 0.0013, Pearson) or IcsA polarity (P; two-tailed P = 0.0014, Pearson) (measured in Figure 2). Red line = fitted linear function, dotted line = 95 % CI. (d) IF microscopy of infected HeLa cells with bacteria expressing IcsA and IcsA cPT mutants. Overlay images are shown for bacterial nucleoids and eukaryotic nuclei detected with DAPI (blue) and either F-actin labelled with phalloidin (green) or N-WASP labelled with anti-N-WASP (red). Each micrograph is 18 μm by 15 μm.

The single-substitution mutations described here have dramatic effects on end-point IcsA surface polarity. This further supports our focus on the cPT region. The observed importance of HS1 and HS2, although based on limited data, indicates that a local structure may be present at this narrower sub-region of the cPT region which drives accumulation at the cell pole.

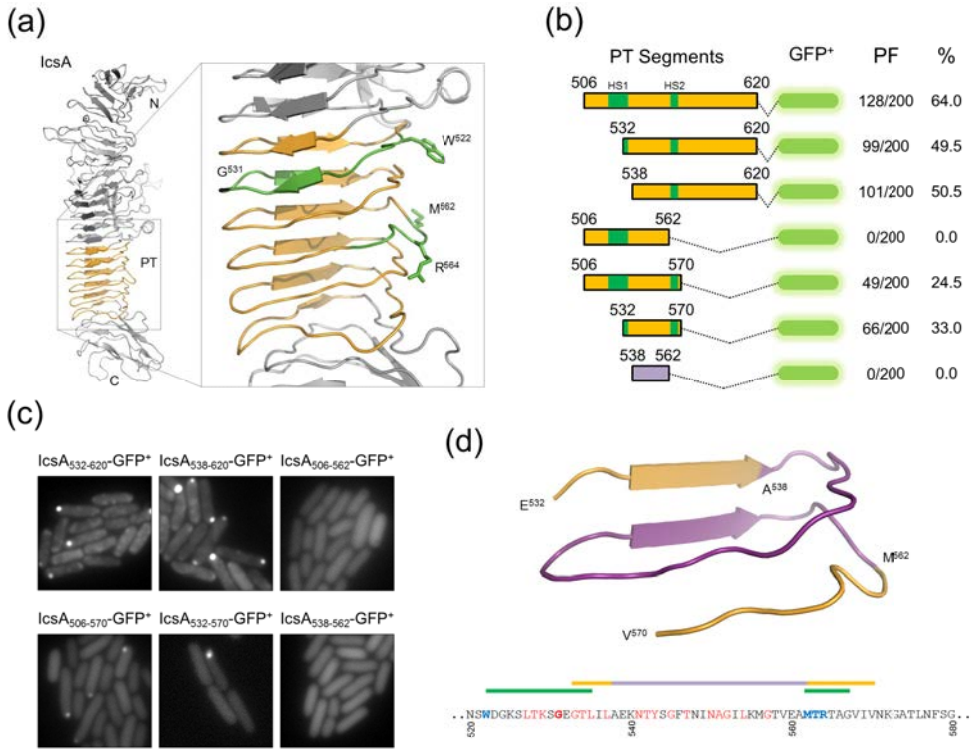


Figure 6.4: Analysis of cPT region fusions and structure.

(a) I-TASSER generated tertiary structure of the IcsA passenger ($IcsA_{53-758}$) with cPT region marked orange and zoomed to show hot-spots (green), important investigated residues, and the putative beta-helical structure. (b) Various IcsA-GFP^+ fusions and resulting polar foci (PF) frequencies (n = 200) during expression in *E. coli* with representative fluorescence micrographs shown in (c). Polarity hot-spots are coloured green. Each micrograph is 6 µm by 6 µm. (d) Minimally sufficient region for polar targeting $IcsA_{532-570}$ is pictured (based on I-TASSER structure). Region insufficient for polar targeting $IcsA_{538-562}$ is coloured purple. Information from mutagenesis (green hot-spots and blue/bold residues) and fusion protein analysis shows that polar targeting properties are clustered in and around a conserved PATR motif (motif ID: IPR013425, PF12951, TIGR02601). Conserved PATR residues are red.

The data suggested that the observed clustering of amino acids important for secretion to the polar surface to two hot-spots was a clue to a local intramolecular structure that enables polarity. We modelled the structure of the IcsA passenger using I-TASSER (Roy *et al*, 2010, Xu *et al*, 2011) and, as expected, the predicted model is primarily a β-helical stalk-like structure (Figure 6.4a). Curiously, the portion of IcsA that exhibited the most ordered core β-helical structure was the cPT region (Figure 6.4a, orange). Furthermore, the two polarity determining hot-spots HS1 and HS2 are separated by one β-rung (Figure 6.4a, inset). Therefore we constructed a series of IcsA cPT segment fusions to GFP+ to see if disrupting these putative rungs would perturb polarisation (Figure 6.4b and c). Fusions containing HS2, but also a partially disrupted HS1 ($IcsA_{532-620}$-GFP^+), or no HS1 ($IcsA_{538-620}$-GFP^+), had reduced polar foci formation frequency (49.5 % and 50.5 % respectively) (Figure 6.4b and c)

compared to that previously observed for $IcsA_{506-620}$-GFP^+ (Figure 6.1b and c, Figure 6.4b and c). Conversely, the fusion $IcsA_{506-562}$-GFP^+ containing HS1 and a disrupted HS2 was completely unable to form polar foci (Figure 6.4b and c). Upon incorporation of the HS2 as fusion $IcsA_{506-570}$-GFP^+ there was recovery of foci formation, although with reduced frequency (24.5 %) (Figure 6.4b and c). While we do not exclude the possibility that other small segments of the PT may be able to induce polar foci formation, we found that the fusion $IcsA_{532-570}$-GFP^+ was minimally sufficient to form polar foci (33 %) (Figure 6.4b and c) and its putative structure is shown in Fig. 4d. Finally, a fusion containing just the putative HS-separating β-rung $IcsA_{538-562}$-GFP^+ (Figure 6.4d, purple) did not form any polar foci (Figure 6.4b and c).

These data strongly suggest that a portion of the cPT forms a polar targeting competent local conformation. The minimal region $IcsA_{532-570}$ appears to be just enough to fold over into a β-wedge (Figure 6.4d). Indeed, residues such as $IcsA^{W522}$ and $IcsA^{M562}$ that were identified to effect IcsA surface localisation are closely spaced in the predicted model and are outward facing (Figure 6.4a) to enable potential intermolecular protein-protein interactions. Additionally, residue $IcsA^{G531}$ is putatively situated in the turn of a β-rung (Figure 6.4a). If these residues provide flexibility for local folding, then it is plausible that the substitution G531W causes enough local disruption to prevent correct polar targeting competent tertiary conformations. These notions are further supported by the observations described in this work that single-substitution mutations produced varied degrees of perturbed IcsA localisation and that larger segments, such as $IcsA_{506-620}$, are maximally efficient at directing GFP^+ to cell poles. Intriguingly, this region is overlapped by a highly conserved 32 amino acid motif called the Passenger-associated Transport Repeat (PATR) ($IcsA_{526-557}$) which we have recently characterized with regards to autotransporter transport across the outer membrane (Doyle *et al*, 2015a). This motif appears to be the focus for IcsA polarity determination with amino acid substitutions that perturbed IcsA polarity located within and flanking to this motif (Figure 6.4d).

6.7. Article Discussion

IcsA is an outer membrane autotransporter protein and a key *Shigella* virulence factor of the human pathogen *S. flexneri*. IcsA is localized to one of the cell poles and enables motility within infected cells and spreading into neighboring cells (Goldberg *et al*, 1993, Goldberg & Theriot, 1995, Kocks *et al*, 1995, Robbins *et al*, 2001). However, the process of IcsA biogenesis, and the mechanism by which IcsA is polarly localized to enable this motility, is still poorly understood.

In this study we attempted to refine our understanding of IcsA-dependent IcsA polar targeting. We focused our attention on the cPT (residues 506-620) rather than the nPT region (residues 1-104) since we were unable to observe polar aggregation of fusions $IcsA_{1-104}$-GFP^+, $IcsA_{53-104}$-GFP^+, and $IcsA_{53-505}$-GFP^+, but were able to observe polar aggregation of fusions $IcsA_{105-620}$-GFP^+ and $IcsA_{506-620}$-GFP^+ (Figure 6.1). Using both random and site-directed mutagenesis we generated substitution mutants within the PT region that caused significant changes in IcsA surface polarity (Figure 6.2). Importantly, we detected clustering of critical amino acid residues to sub-regions within the PT region (Figure 6.2) which suggested the hypothesis that a small tertiary structure, rather than specific residues, facilitates polar targeting of nascent IcsA. Subsequently, we analyzed the ability of smaller segments of the PT region to polarize GFP^+ and identified a minimally sufficient 38 amino acid region (532-570) (Figure 6.4). Coincidentally, this region overlaps a highly conserved 32 amino acid PATR motif (a.a 526-557) which appears to be the focus for IcsA polarity determination. Amino acid substitutions that perturbed IcsA polarity were located within and adjacent to this motif.

These results lead to further questions. For instance, it is certainly intriguing that residues important for polarity cluster around a conserved PATR motif (Pfam identifier; PF12951) (Finn *et al*, 2014a). We have previously observed that the PATR motif is required for the efficient transport of the passenger domain across the outer membrane, is present in many autotransporters, and likely prescribes a β-wedge structure (Doyle *et al*, 2015a). Yet, the PATR functions in a separate cellular topology post secretion, and IcsA-dependent polarity determination is thought to be set-up prior to secretion (Charles *et al*, 2001, Janakiraman *et al*, 2009). How might this motif augment IcsA polarity? Most of the polarity determining residues identified flank the motif, such as $IcsA^{W522}$ and $IcsA^{M562}$ (Figure 6.4d). The substitution $IcsA^{G531W}$ is at a conserved site but is a bulky change that may hinder the local folding of this region. Indeed, we have previously proposed that this and other glycines in the motif are required for folding flexibility (Doyle *et al*, 2015a). Therefore, it is seems likely that the motif

merely provides a convenient scaffold for the polarity determining residues to function in the context of a transient intramolecular structure in the cytosol.

The mechanism of polarity determination at this site might be remarkably simple. It has previously been observed that $IcsA_{PT}$-GFP fusions behave similarly to thermo-aggregating proteins whereby small randomly distributed aggregates form upon expression, which then coalesce into polar aggregates at the old pole (Rokney *et al*, 2009). The preference of aggregates to accumulate at the pole is likely due to the combined actions of nucleoid occlusion and asymmetric aggregate inheritance during population expansion (Lindner *et al*, 2008, Rokney *et al*, 2009, Sabate *et al*, 2010, Winkler *et al*, 2010). However, for this idea to be viable, the removal of IcsA from polar aggregates must be efficient. It has been shown that the formation of polar aggregates was abrogated in cells with the disrupted gene *dnaK* (Rokney *et al*, 2009). DnaK is a heat shock protein described as the central hub in the chaperone network that is responsible for efficient protein disaggregation, but also the initial depository of aggregates to the pole (Carrio & Villaverde, 2005, Calloni *et al*, 2012, Castanie-Cornet *et al*, 2013, Saibil, 2013). Therefore, the DnaK system may be the mediators of IcsA disaggregation out of these polar aggregates, at which point it can be secreted in an old polar biased manner. This links well with the knowledge that DnaK depletion results in circumferential localisation of IcsA on the bacterial surface (Janakiraman *et al*, 2009). Thus, a system involving aggregation and disaggregation of IcsA can be an efficient and dynamic process. The knowledge that other proteins (for instance *E. coli* MurG) can be deposited and removed dynamically from polar aggregates provides a precedent for the model (Michaelis & Gitai, 2010).

Our data fits well with this model as aggregation hot-spots are commonly found in proteins due to their necessity in the formation of β-sheet secondary structures (Vendruscolo *et al*, 2003, Sabate *et al*, 2010). Indeed, according to the Pfam database, our previous studies on this region, and analysis in this work, the cPT is likely to adopt a β-helical structure (Doyle *et al*, 2015a). Although further studies are required to completely understand this process, we suggest that polarity disrupting substitutions investigated in this work may inhibit the proper functioning of this region in aggregate nucleation and coalescence to the pole. We also do not preclude the possibility that this region may also be required for appropriate IcsA-IcsA or IcsA-DnaK interactions that drive this process. This model relies on there being sufficient IcsA concentration in the cytosol to allow coalescence into polar aggregates and predicts that decreases in total IcsA concentration will cause decreases in resultant surface polarity. This correlates exactly with what we observed for IcsA cPT mutants $IcsA^{M562R}$ and $IcsA^{G531W}$ which had reduced total cellular levels combined with reduced surface

polarity (Figure 6.2). Conversely, we have previously observed an effect of increased cellular levels of IcsA in increasing the tendency of bi-polar IcsA distributions (Doyle *et al*, 2015b). This again may be explained by an aggregative mechanism of IcsA polarity.

In summary, we investigated the protein regions involved in IcsA-dependent IcsA polarity determination. Several novel findings have been made; (i) we could not replicate a role for nPT in IcsA polarization, (ii) mutagenesis of the cPT has revealed polar targeting functional hot-spots (HS1 and HS2), (iii) single residues have a measurable (and variable) role in polar determination, and (iv) functional residues and regions cluster around a conserved PATR motif which may provide a functional scaffold for polarity augmenting residues. These observations fit well with previously proposed aggregative models of IcsA polarity and stimulate the need to identify the specific interactions of this region.

6.8. Article Acknowledgements

This work is supported by a Program Grant (No.: 565526) from the National Health and Medical Research Council (NHMRC) of Australia. MTD is the recipient of a Doctor of Philosophy scholarship from the University of Adelaide. We also thank the Research Centre for Infectious Diseases (RCID) for support during this work. We greatly appreciate Alistair Standish and Elizabeth Ngoc Hoa Tran for critical reading of the manuscript.

6.9. Article References

Bernardini, M. L., Mounier, J., d'Hauteville, H., Coquis-Rondon, M. & Sansonetti, P. J. (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.

Bolivar, F., Rodriguez, R. L., Greene, P. J., Betlach, M. C., Heyneker, H. L., Boyer, H. W., Crosa, J. H. & Falkow, S. (1977). Construction and characterization of new cloning vehicles. II. A multipurpose cloning system. *Gene* 2, 95-113.

Brotcke Zumsteg, A., Goosmann, C., Brinkmann, V., Morona, R. & Zychlinsky, A. (2014). IcsA is a *Shigella flexneri* adhesion regulated by the type III secretion system and required for pathogenesis. *Cell Host Microbe* 15, 435-445.

Calloni, G., Chen, T., Schermann, S. M., Chang, H. C., Genevaux, P., Agostini, F., Tartaglia, G. G., Hayer-Hartl, M. & Hartl, F. U. (2012). DnaK functions as a central hub in the *E. coli* chaperone network. *Cell Rep* 1, 251-264.

Carrio, M. M. & Villaverde, A. (2005). Localization of chaperones DnaK and GroEL in bacterial inclusion bodies. *J Bacteriol* 187, 3599-3601.

Castanie-Cornet, M. P., Bruel, N. & Genevaux, P. (2013). Chaperone networking facilitates protein targeting to the bacterial cytoplasmic membrane. *Biochim Biophys Acta* 10.1016/j.bbamcr.2013.11.007.

Charles, M., Perez, M., Kobil, J. H. & Goldberg, M. B. (2001). Polar targeting of *Shigella* virulence factor IcsA in *Enterobacteriacae* and *Vibrio*. *Proc Natl Acad Sci U S A* 98, 9871-9876.

Doyle, M. T., Tran, E. N. & Morona, R. (2015a). The passenger-associated transport repeat promotes virulence factor secretion efficiency and delineates a distinct autotransporter subtype. *Mol Microbiol* 97(2):315-329.

Doyle, M. T., Grabowicz, M., May, K. L. & Morona, R. (2015b). Lipopolysaccharide surface structure does not influence IcsA polarity. *FEMS Microbiol Lett* 362, fnv042.

Egile, C., Loisel, T. P., Laurent, V., Li, R., Pantaloni, D., Sansonetti, P. J. & Carlier, M. F. (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.

Finn, R. D., Bateman, A., Clements, J., Coggill, P., Eberhardt, R. Y., Eddy, S. R., Heger, A., Hetherington, K., Holm, L. & other authors (2014). Pfam: the protein families database. *Nucleic Acids Res* 42, D222-230.

Fixen, K. R., Janakiraman, A., Garrity, S., Slade, D. J., Gray, A. N., Karahan, N., Hochschild, A. & Goldberg, M. B. (2012). Genetic reporter system for positioning of proteins at the bacterial pole. *MBio* 3, e00251-00211. 00210.01128/mBio.00251-00211.

Fukuda, I., Suzuki, T., Munakata, H., Hayashi, N., Katayama, E., Yoshikawa, M. & Sasakawa, C. (1995). Cleavage of *Shigella* surface protein VirG occurs at a specific site, but the secretion is not essential for intracellular spreading. *J Bacteriol* 177, 1719-1726.

Goldberg, M. B. & Theriot, J. A. (1995). *Shigella flexneri* surface protein IcsA is sufficient to direct actin-based motility. *Proc Natl Acad Sci U S A* 92, 6572-6576.

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

Janakiraman, A., Fixen, K. R., Gray, A. N., Niki, H. & Goldberg, M. B. (2009). A genome-scale proteomic screen identifies a role for DnaK in chaperoning of polar autotransporters in *Shigella*. *J Bacteriol* 191, 6300-6311.

Kocks, C., Marchand, J. B., Gouin, E., d'Hauteville, H., Sansonetti, P. J., Carlier, M. F. & Cossart, P. (1995). The unrelated surface proteins ActA of *Listeria monocytogenes* and IcsA of *Shigella flexneri* are sufficient to confer actin-based motility on *Listeria innocua* and *Escherichia coli* respectively. *Mol Microbiol* 18, 413-423.

Kovach, M. E., Elzer, P. H., Hill, D. S., Robertson, G. T., Farris, M. A., Roop, R. M., 2nd & Peterson, K. M. (1995). Four new derivatives of the broad-host-range cloning vector pBBR1MCS, carrying different antibiotic-resistance cassettes. *Gene* 166, 175-176.

Lett, M. C., 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.

Leyton, D. L., Rossiter, A. E. & Henderson, I. R. (2012). From self sufficiency to dependence: mechanisms and factors important for autotransporter biogenesis. *Nat Rev Microbiol* 10, 213-225.

Lima, I. F., Havt, A. & Lima, A. A. (2015). Update on molecular epidemiology of *Shigella* infection. *Curr Opin Gastroen* 31, 30-37.

Lindner, A. B., Madden, R., Demarez, A., Stewart, E. J. & Taddei, F. (2008). Asymmetric segregation of protein aggregates is associated with cellular aging and rejuvenation. *Proc Natl Acad Sci U S A* 105, 3076-3081.

Lugtenberg, B., Meijers, J., Peters, R., van der Hoek, P. & van Alphen, L. (1975). Electrophoretic resolution of the "major outer membrane protein" of *Escherichia coli* K12 into four bands. *FEBS Lett* 58, 254-258.

Makino, S., Sasakawa, C., Kamata, K., Kurata, T. & Yoshikawa, M. (1986). A genetic determinant required for continuous reinfection of adjacent cells on large plasmid in *S. flexneri* 2a. *Cell* 46, 551-555.

May, K. L. & 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.

Michaelis, A. M. & Gitai, Z. (2010). Dynamic polar sequestration of excess MurG may regulate enzymatic function. *J Bacteriol* 192, 4597-4605.

Niyogi, S. K. (2005). Shigellosis. *J Microbiol* 43, 133-143.

Oaks, E. V., Wingfield, M. E. & Formal, S. B. (1985). Plaque formation by virulent *Shigella flexneri*. *Infect Immun* 48, 124-129.

Robbins, J. R., Monack, D., McCallum, S. J., Vegas, A., Pham, E., Goldberg, M. B. & Theriot, J. A. (2001). The making of a gradient: IcsA (VirG) polarity in *Shigella flexneri*. *Mol Microbiol* 41, 861-872.

Rokney, A., Shagan, M., Kessel, M., Smith, Y., Rosenshine, I. & Oppenheim, A. B. (2009). *E. coli* transports aggregated proteins to the poles by a specific and energy-dependent process. *J Mol Biol* 392, 589-601.

Roy, A., Kucukural, A. & Zhang, Y. (2010). I-TASSER: a unified platform for automated protein structure and function prediction. *Nat Prot* 5, 725-738.

Sabate, R., de Groot, N. S. & Ventura, S. (2010). Protein folding and aggregation in bacteria. *Cell Mol Life Sci* 67, 2695-2715.

Saibil, H. R. (2013). Machinery to reverse irreversible aggregates. *Science* 339, 1040-1041.

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.

Schroeder, G. N. & Hilbi, H. (2008). Molecular pathogenesis of *Shigella spp.*: controlling host cell signaling, invasion, and death by type III secretion. *Clin Microbiol Rev* 21, 134-156.

Snapper, S. B., Takeshima, F., Anton, I., Liu, C. H., Thomas, S. M., Nguyen, D., Dudley, D., Fraser, H., Purich, D. & other authors (2001). N-WASP deficiency reveals distinct pathways for cell surface projections and microbial actin-based motility. *Nat Cell Biol* 3, 897-904.

Steinhauer, J., Agha, R., Pham, T., Varga, A. W. & Goldberg, M. B. (1999). The unipolar *Shigella* surface protein IcsA is targeted directly to the bacterial old pole: IcsP cleavage of IcsA occurs over the entire bacterial surface. *Mol Microbiol* 32, 367-377.

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

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.

Teh, M. Y. & Morona, R. (2013). Identification of *Shigella flexneri* IcsA residues affecting interaction with N-WASP, and evidence for IcsA-IcsA co-operative interaction. *PloS One* 8, e55152.

Teh, M. Y., Tran, E. N. & Morona, R. (2012). Absence of O antigen suppresses *Shigella flexneri* IcsA autochaperone region mutations. *Microbiol-SGM* 158, 2835-2850.

Tran, E. N., Doyle, M. T. & Morona, R. (2013). LPS unmasking of *Shigella flexneri* reveals preferential localisation of tagged outer membrane protease IcsP to septa and new poles. *PloS One* 8, e70508.

Valencia-Gallardo, C. M., Carayol, N. & Nhieu, G. T. V. (2015). Cytoskeletal mechanics during *Shigella* invasion and dissemination in epithelial cells. *Cell Microbiol* 17, 174-182.

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

Vendruscolo, M., Paci, E., Karplus, M. & Dobson, C. M. (2003). Structures and relative free energies of partially folded states of proteins. *Proc Natl Acad Sci U S A* 100, 14817-14821.

von Seidlein, L., Kim, D. R., Ali, M., Lee, H., Wang, X., Thiem, V. D., Canh do, G., Chaicumpa, W., Agtini, M. D. & other authors (2006). A multicentre study of *Shigella* diarrhoea in six Asian countries: disease burden, clinical manifestations, and microbiology. *PLoS Med* 3, 1556-1569.

Winkler, J., Seybert, A., Konig, L., Pruggnaller, S., Haselmann, U., Sourjik, V., Weiss, M., Frangakis, A. S., Mogk, A. & other authors (2010). Quantitative and spatio-temporal features of protein aggregation in *Escherichia coli* and consequences on protein quality control and cellular ageing. *Embo J* 29, 910-923.

Xu, D., Zhang, J., Roy, A. & Zhang, Y. (2011). Automated protein structure modeling in CASP9 by I-TASSER pipeline combined with QUARK-based ab initio folding and FG-MD-based structure refinement. *Proteins* 79 Suppl 10, 147-160.

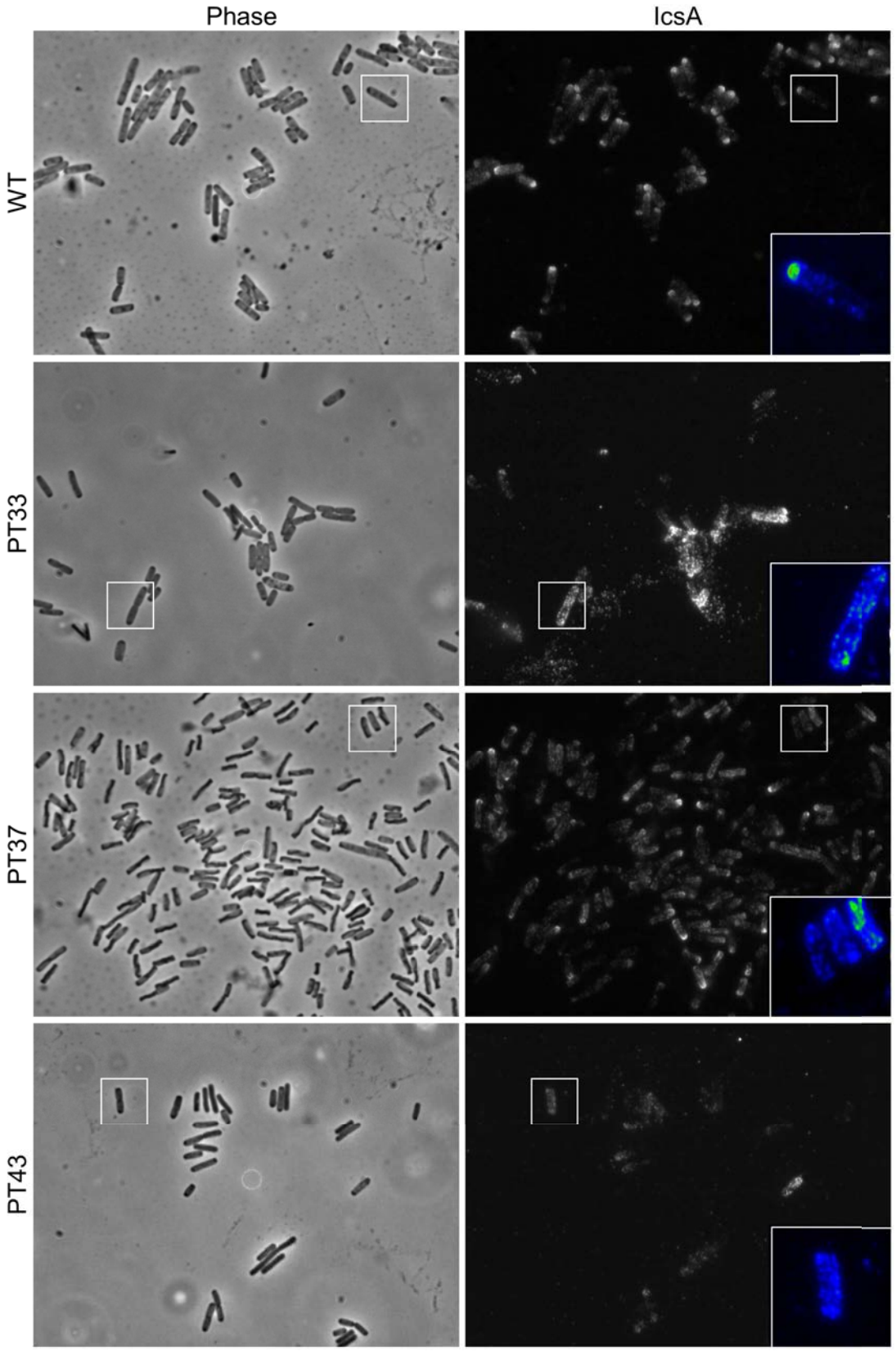


Figure 6.5[S] continued next page

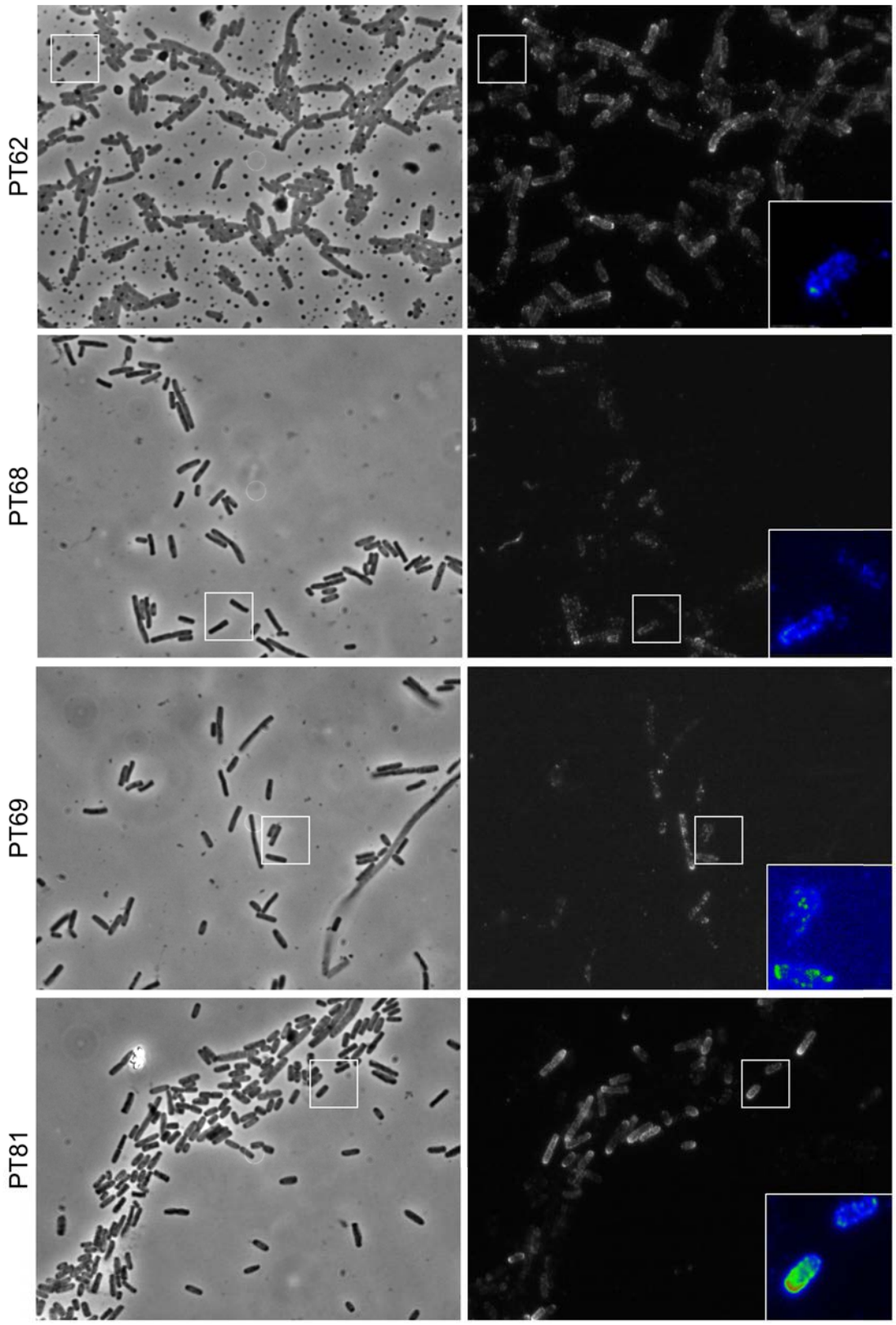


Figure 6.5[S] continued next page

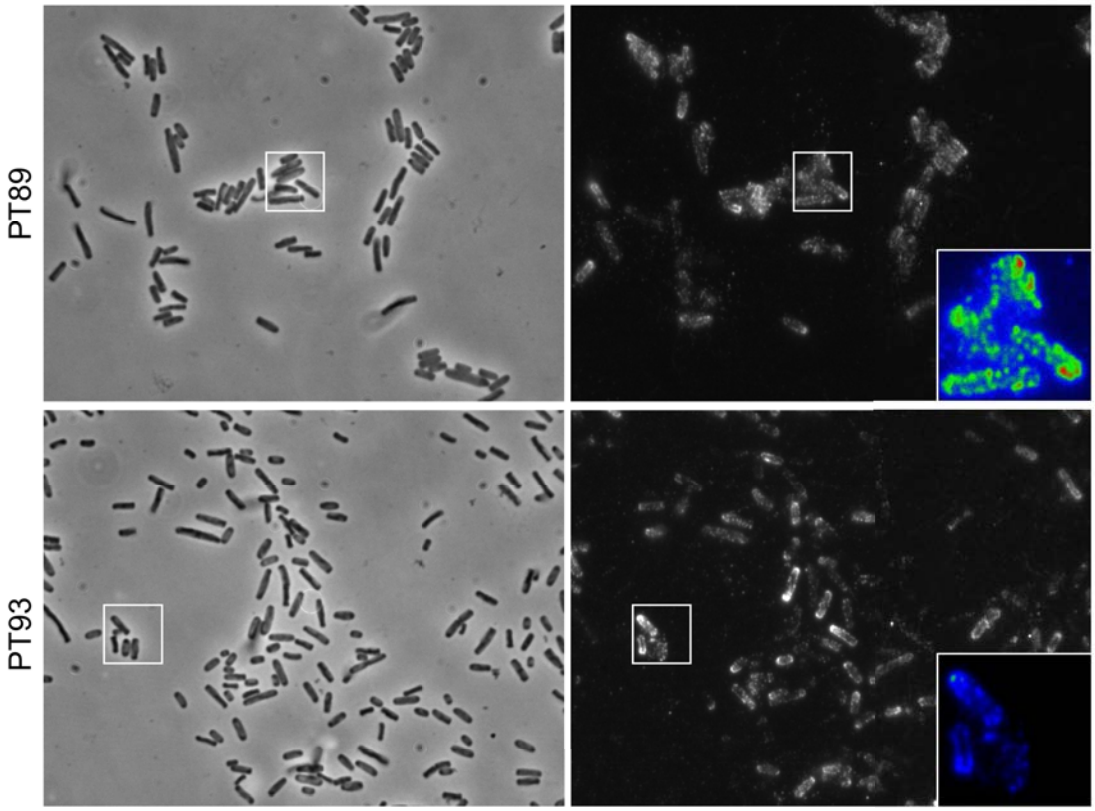


Figure 6.5[S]: Screening IcsA PT region random mutants for surface expression aberrations.

Plasmid pIcsA that was randomly mutagenised over the PT coding region (see *Methods*) was transformed into RMA2041 and 100 putative mutant transformants were screened for aberrant surface expression via microscopy after immunofluorescent labelling of IcsA. 9/100 isolates (PT#) displayed abnormal surface expression compared to wild type control RMA2041[pIcsA]. Sample phase contrast images (left), fluorescence images (right), *Insets* = intensity-recoloured image of representative boxed bacterium, images are 6 μm by 6 μm.

Chapter Seven

Conclusions

Chapter 7: Conclusions

7.1. IcsA polarity

During this work major factors that affect IcsA polar targeting were examined. Firstly, the role of proteolytic refinement of IcsA polarity by IcsP was investigated in Chapter 3 by assessing the surface distribution of IcsP relative IcsA. It was found that IcsP surface localisation is asymmetrically localised. Indeed, IcsP was found at septation sites of dividing bacteria, and also at the new pole. Further, the concentration of IcsP at septa was significantly higher than that of the new pole. This leads to the conclusion that the new pole IcsP is derived (inherited) from the septal material. IcsP surface distribution is life cycle dependent and is not symmetrically distributed which was the previous model (Steinhauer *et al*, 1999). Therefore, it is reasonable to assume that IcsA cleavage by IcsP occurs most frequently at the site of septation and at new poles. IcsP is important for the initial set-up of IcsA at old poles of young *S. flexneri* cells, and is further required for maintenance of IcsA polarity for elongating cells. However, this raises the further question – how is IcsP localised to the site of septation of dividing *Shigellae* and retained at new poles? Few outer membrane proteins have been reported to localise towards the septum. An exception to this is the outer membrane lipoprotein LpoB (and LpoB-like lipoproteins) which complexes with septal divisome proteins such as periplasmic penicillin binding protein 1B (Typas *et al*, 2010, Egan & Vollmer, 2013). Whether the localisation of IcsP (an integral outer membrane protein) interacts with divisome components, or other landmarks at the septum, remains to be investigated.

While the localisation of IcsP is a key observation for conceptualising IcsA polar biogenesis, the finding remained shrouded by the proposed modulation by LPS surface structure. There are reports in the literature that alterations in the LPS O-antigen structure changes membrane fluidity which subsequently alters IcsA stable localisation at the pole, and further, that the short and very long O-antigen modal lengths may allow asymmetric masking of IcsA exposure. However, findings presented in Chapter 4 revealed that polar material contained equivalent LPS O-antigen lengths distributions comparatively to whole cells. Further, when O-antigen was removed altogether, IcsA surface distribution was equivalent in comparison to wild-type *Shigella* that is O-antigen positive. Thus, IcsA polarity is not affected by LPS structure with regards to either asymmetric masking or due to drifting away from the pole towards the lateral edge of the bacterium. Thus, with respect to IcsP, it is unlikely that its

role is to cleave laterally diffusing IcsA since this does not appear to occur (at least measurably). Actually, it is most likely that IcsP only cleaves IcsA that has been misdirected to septal or new polar positions in the first instance.

Together, the conclusions presented in Chapters 3 and 4 further highlight that the initial targeting mechanism is very important for overall IcsA polar localisation. To further investigate early IcsA polar targeting, the IcsA polar targeting (PT) regions were examined in Chapter 6. It was revealed that the central PT (cPT) was most important in polar accumulation. Polar targeting functional residues were identified that clustered as two hot-spots. Further, these hot-spots cluster around a smaller element sufficient for polar accumulation which contained the conserved autotransport PATR motif. Currently, it remains elusive how this region exerts polar targeting properties. It is possible that the cPT may prescribe a transient aggregate nucleation event that causes IcsA coalescence to the pole. Moreover, it was observed that some cPT polarity disrupting substitutions also coincided with reduced total IcsA cellular concentration, and in Chapter 4, that increased cellular levels resulted in an increased tendency for bipolarity. These observations can be explained by a previously proposed aggregative model of polarity setup (Rokney *et al*, 2009) which predicts that decreases in total IcsA concentration limit the chances of nucleation and subsequent polar coalescence, and that increases in total IcsA concentration would increase the likelihood of off-target accumulation at both poles. Although further studies are required to completely understand this process, I predict that future experiments investigating potential binding partners to the cPT would be extremely informative. For this, site specific crosslinking experiments (i.e. using amber codons and UV crosslinking) during pulse-chase style assays, coupled with advanced super resolution microscopy (for single molecule tracking in live cells) would be useful.

7.2. Autotransporters and the PATR motif

Regarding autotransporter assembly, there appears to be a high level of mixing of conserved subdomains, motifs, and repeats in the construction of the passenger. For these reasons it is still poorly understood how these diverse passengers are efficiently translocated to the bacterial cell surface. IcsA contains a highly conserved 32-aa motif ($IcsA_{526-557}$) that was annotated as an 'autotransporter-associated β-repeat'. Despite its conservation, this motif remained completely uncharacterised in any protein. As such, Chapter 5 describes the first report in characterising the function of this repeat in autotransporter biogenesis. Of the conserved residues of this motif, the most conserved were four interspersed glycine residues. It was found that these residues are required for presentation of the passenger moiety on the bacterial cell surface which, consequently, resulted in a reduction of N-WASP recruitment to *S. flexneri* during epithelial cell infection. Further, the motif plays a role in the efficiency of passenger transport to the cell surface. Thus, the motif was renamed the 'passenger-associated transport repeat' (PATR) which is a more informative description of its role.

Therefore, the PATR is an important biogenesis element, at least in the case of IcsA autotransport. Chapter 5 also investigated the wider use of the repeat in other Gram negative autotransporter proteins. A wider view of the importance of the PATR in autotransporters was gained via database-mining and bioinformatic analyses. The usage of other domains and motifs by autotransporters were analysed showing that the PATR motif is present in passengers very commonly and often distally repeated within the same passenger. It was found that PATR-containing autotransporters have distinctly different (and significantly different) virulence functional features and architectures compared to autotransporters without the motif (i.e. subtilisin-type serine proteases and polymorphic repeats). Therefore, there seems to exist at least two subtypes of autotransporters: PATR-type and non-PATR autotransporters. PATR-type and non-PATR autotransporters are also used differentially between different Gram negative genera (section 5.12). Further, PATR-type autotransporters are significantly larger, and the largest of these cluster with low isoelectric point. The recently postulated notion that large autotransporters will require, as yet unidentified, additional secretion driving forces (Drobnak *et al*, 2015a), coupled with the observation that passenger charge can effect secretion (Kang'ethe & Bernstein, 2013b), leads to the intriguing hypothesis that the clustering is indicative of this theoretical driving force. Additionally, it could also be speculated that the differences in size and pI is a reflection of the differential usage of the PATR between different pathological niches or pathogenic functions. I predict that further bioinformatic database analyses (for instance sorting autotransporters on pathological niche

and function and assessing relative representation of physical characteristics such as PATR, pI, and size) will be useful in investigating this notion. However, this will be reliant on additional annotation of these databases with pathogenic information, which in itself is reliant on further experimentation of uncharacterised autotransporters.

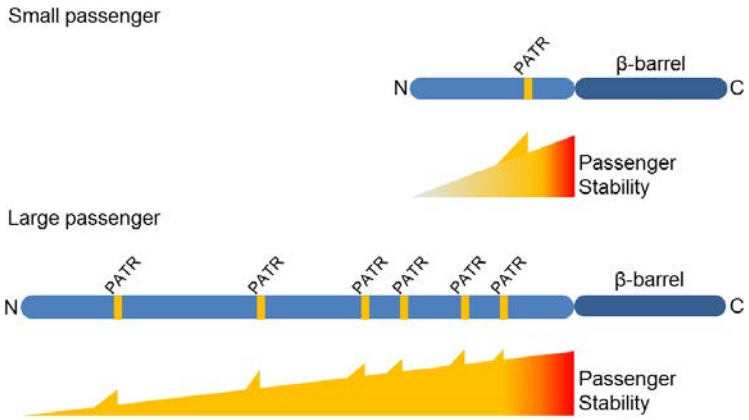


Figure 7.1: Concept of 'linear ratcheting' by distally repeated PATR motifs.

A diagrammatic hypothesis whereby the PATR motif prescribes a stable folding region providing a driving force for the secretion of the passenger to the bacterial surface. Autotransporters that are very long may have many distally repeated PATR motifs to provide a linear ratcheting effect during secretion.

Although an argument of the wide importance of the PATR motif in the construction of autotransporters has been made, the exact mechanism of transport augmentation by the PATR remains uncertain. *In silico* modelling suggested the conserved glycines are present in β-wedge apices indicating a requirement in folding flexibility. It is well recognised that the vectorial folding of the passenger plays a key role in the driving force in translocation (Renn & Clark, 2008, Peterson *et al*, 2010, Soprova *et al*, 2010, Braselmann & Clark, 2012, Kang'ethe & Bernstein, 2013b, Drobnak *et al*, 2015a). Therefore, the stepwise folding of the PATR during secretion might be a contributor to the driving force. The effect of the stable folding of the C-terminus of the passenger has been likened to a Brownian Ratchet in the past (Leyton *et al*, 2012). In extension of this, the distally repeated PATR motifs found in the largest PATR-type autotransporters may afford a 'linear ratcheting' effect providing efficient secretion (see Figure 7.1). However, it also remains alternatively possible that the PATR is required for in intricate and timely interactions with other factors required with other factors during outer membrane insertion and translocation such as chaperones and the BAM-complex. It is interesting to speculate that the PATR might contribute to the strong peripheral interaction observed between the orphan passenger (without β-barrel) and the outer membrane (section 5.13). Regardless, further experimentation is required to uncover the

PATR molecular mechanism. Specific pulse-chase, crosslinking experiments would again be useful in the endeavour to uncover PATR interactors. Mutagenesis of multiple different PATR motifs within a Group 1 PATR-type autotransporter such as ShdA (see sub-section 5.6.3) would also be useful in investigating the linear ratchet hypothesis. To elucidate the folding characteristics of the PATR, biophysical *in vitro* folding-unfolding experiments (such as circular dichroism) and protease susceptibility analyses should be conducted.

7.3. The PATR/cPT region functional interplay

Taking the work presented in Chapters 5 and 6 as a whole, there appears to be an unusual link between IcsA polarity and IcsA secretion efficiency due to the physical overlap between the cPT region and the IcsA PATR motif. Therefore, this region of IcsA represents a multifunctional IcsA biogenesis coordinating region. Yet, how can this region coordinate both polarity augmentation and passenger secretion efficiency which are quite distinct events? I speculate that although these regions are physically linked, they function differently within separate bacterial sub-cellular compartments. In the cytoplasm, the cPT region is exposed to factors that allow it to augment polar accumulation. Conversely, in the periplasm or during outer membrane transit, the PATR is required for some stage of passenger transport and assembly to the bacterial cell surface. Thus, it may be the specific sub-cellular environment that allows this region to conduct discrete functions. Unfortunately I have not been able to experimentally demonstrate these notions within the time of candidature. It would be ideal to further test whether there is any polarity disrupting effects of substituting the IcsA PATR residues. It would also be interesting to determine whether there is any role in the kinetics of passenger secretion for the IcsA cPT residues with a role in polarity augmentation. The methodologies designed during this work, including population based IcsA polarity quantification and *in vivo* passenger secretion kinetics analyses, will be extremely useful in further investigations of this multifunctional region.

There is also the question of whether the presence of the PATR motif in other autotransporters would confer polar targeting properties. Although I cannot exclude this possibility for all PATR-type autotransporters, it does not currently appear to be a general case. By way of example, one could argue that the PATR-type autotransporter EprS from *P. aeruginosa* would not require polar distributions since it functions as a secreted inflammation augmenting toxin (Kida *et al*, 2013). Further, the PATR-type autotransporter ShdA from *S. enterica* has been long studied (Kingsley *et al*, 2000, Kingsley *et al*, 2002, Kingsley *et al*, 2004b,

Kingsley *et al*, 2004a) yet there has been no reports of polar surface localisation. Finally, the PATR-type autotransporter NalP is from *N. meningitidis* (van Ulsen *et al*, 2003) which has a coccus cell shape, and hence lacks conventional poles.

7.4. Summary

The theme of this thesis was to explore the spatial distribution mechanisms of the *Shigella* classical autotransporter IcsA in the contexts of both; targeted placement around the macrostructure of the cell, and the correct export to the bacterial cell surface. This work has provided substantial contributions in the understanding of both the asymmetrical polar placement of IcsA and its autotransport mechanism. The discovery and characterisation of the PATR motif is predicted to have a wide reaching impact on understanding the secretion of many autotransporter virulence factors. Overall, this thesis highlights the critical importance of the correct spatial organisation of virulence factors within and around the bacterial cell.

Appendices

Oligo List

Strains Lists

Bibliography

Appendix A - Oligonucleotides Generated

Code	Sequence (5' to 3')	Notes
md1	GGCCATCACCATCACCATCACCATCAC	F, OA, His_8 Tag plus NotI overhangs
md2	GGCCGTGATGGTGATGGTGATGGTGAT	R, OA, His_8 Tag plus NotI overhangs
md3	GAAAACGCTGAACTTACT	F, anneals 5' to PT coding region
md4	ACTATTATTGAGAATAACATG	R, anneals 3' to PT coding region
md5	GAGTTCAATCTGAAAGAAAACGCTGAACTTACT	F, DTRM of PT coding region
md6	CCGACATTTGACGAACTATTATTGAGAATAACATG	R, DTRM of PT coding region
md7	CTACTGGTGCTATTGACTATACAAAAAGCTATATC	F, PT coding region sequencing
md8	TTTGCTAGCATGGTGAGCAAGGGCGAGGA	F, delete *murG* from pAM2, insert NheI
md9	AAAGCTAGCACCTCGTACACGAGCCTGC	R, delete *murG* from pAM2, insert NheI
md10	CCCGCTAGCATGAATCAAATTCACAAATTTTTTTGTAATATGACCC	F, amplify *icsA*, add flanking NheI
md11	GGGGCTAGCGAAGGTATATTTCACACCCAAAATACCTTG	R, amplify *icsA*,, add flanking NheI
md12	CGCCGACATCATAACGGTTCTGGC	F, sequencing primer for inserts into pMDRM19
md13	CAAATCACCAGATAGGCAACCGTAATGTG	F, *icsA* sequencing, nt $icsA_{3148\text{-}3177}$
md14	TCTCGAGCATGTTATTCTCAATAATAGTTCGTCAAATG	F, IP replacement of PT coding region with XhoI
md15	BLANK	
md16	CTCGCCCTCGCCCTCGATC	R, sequencing primer for inserts into pMDRM19
md17	GGTACCAGGAGGAAACGATGAATCAAATTCACAAATTTTTTTGTAATATGA	F, amplify *icsA*,, add 5' KpnI and RBS
md18	GTCGACTCAGAAGGTATATTTCACACCC	R, amplify *icsA*,, add 3' SalI
md19	AGGCTGTTACAGGAGACAATCTG	F, *icsA* sequencing, nt $icsA_{496\text{-}525}$
md20	CTGAGCTGACCTGTAGTAGGATC	R, *icsA* sequencing, nt $icsA_{2427\text{-}2451}$
md21	CTAGCCCGCGGCTCGAGA	F, OA, MCS for pMDRM19
md22	GTAGTCTCGAGCCGCGGG	R, OA, MCS for pMDRM19
md23	CCGCGGCTCGAGACTAGC	F, screening for MCS for pMDRM19
md24	CCGGTGGAGTGGCGGCCC	R, screening for MCS for pMDRM19
md25	CCCTAACAAAATCAGGGGCGGGAACTCTCATTTTGGC	F, QC, $IcsA_{E532A}$
md26	GCCAAAATGAGAGTTCCCGCCCCTGATTTTGTTAGGG	R, QC, $IcsA_{E532A}$
md27	CCCTAACAAAATCAGGGGAGGCAACTCTCATTTTGGC	F, QC $IcsA_{G533A}$
md28	GCCAAAATGAGAGTTGCCTCCCCTGATTTTGTTAGGG	R, QC $IcsA_{G533A}$
md29	GACAGTTGAAGCTATGGCACGTACCGCTGGTGTTATTG	F, QC $IcsA_{T563A}$
md30	CAATAACACCAGCGGTACGTGCCATAGCTTCAACTGTC	R, QC $IcsA_{T563A}$
md31	GGACAGTTGAAGCTATGACAGCTACCGCTGGTGTTATTG	F, QC $IcsA_{R564A}$
md32	CAATAACACCAGCGGTAGCTGTCATAGCTTCAACTGTCC	R, QC $IcsA_{R564A}$
md33	ACTCGAGAGTAAGTTCAGCGTTTTCTTTCAGATTG	R, IP replacement of PT coding region with XhoI
md34	CTACGGTTTGAGCTGGAATGATACAG	R, *icsA* sequencing, nt $icsA_{1441\text{-}1465}$
md35	GATCTACGGTTTGAGCTGGAATG	F, *icsA PT* region sequencing, nt $icsA_{1438\text{-}1459}$,
md36	GTACTATTGCTGTCAGATGCTC	F, screening for IcsA EcoRI-SalI locus
md37	CGAAAGGGCCTCGTGATACG	R, screening for IcsA EcoRI-SalI locus
md38	CGGAATCTTTTCAGGGGTTTATC	F, sequencing primer for pIcsA and derivatives
md39	CAGAGAAATGCAGGACATCAAC	R, sequencing primer for pIcsA and derivatives
md40	GTCGACGCGACTACTCATTTGAGTAGACTCTTG	R, amplify *icsAp*, add 3' SalI, nt $icsA_{2248\text{-}2274}$
md41	BLANK	
md42	CAAAATCAGGGGAGGGAACTAACATTTTGGCGG	F, QC, $IcsA_{L535N}$
md43	CCGCCAAAATGTTAGTTCCCTCCCCTGATTTTG	R, QC, $IcsA_{L535N}$

md44	GCGGCATGGACGAGCTG	F, sequence pMDRM19 backbone after mCherry
md45	GCTAACAGCGCGATTTGC	R, sequence pMDRM19 backbone before insert
md46	TTTGCTAGCGCGACTACTCATTTGAGTAGACTGTTG	R, amplify *icsAp*, add 3' NheI, nt $icsA_{2248\text{-}2274}$
md47	GCTCTCATGTTTTGGTTGAGGC	F, sequencing primer for pIcsA and derivatives
md48	TTGCTAGCATGAAGCTACGGATTAACTCTGATATTATGATTAG	F, amplify from $icsA_{313}$, add 5' SalI, nt $icsA_{313\text{-}344}$
md49	ATTAGTCTCGAGAGTAAGTTCAGCGTTTTCTTTCAGATTG	R, IP replacement of PT coding region with XhoI
md50	GAGAAACTCGAGCATGTTATTCTCAATAATAGTTCGTCAAATG	F, IP replacement of PT coding region with XhoI
md51	GTAGGACAAATCCGCCGG	F, sequence pMDRM19 backbone after mCherry
md52	GAAGGACCGAGCGCAGC	R, sequence pMDRM19 backbone
md53	ATAGGTACCAGGAGGAACGATGAATCAAATTCACAAATTTTTTTGTAATATGA	F, amplify *icsA*, add 5' KpnI and RBS
md54	TGAGTCGACTCAGCGACTACTCATTTGAGTAGAC	R, amplify *icsAp*, add 3' NheI, nt $icsA_{2253\text{-}2274}$
md55	CAAAAGTGTCTATAATCACGGCAG	F, pBAD inserts screening /sequencing
md56	GGCATGGGGTCAGGTGG	R, pBAD inserts screening / sequencing
md57	TTTGGTACCAGGAGGAAACGATGAATCAAATTCACAAATTTTTTTGTAATATGA	F, amplify *icsA,*, add 5' KpnI and RBS
md58	TTTGTCGACTCAGAAGGTATATTTCACACCC	R, amplify *icsA*, add 3' SalI STOP
md59	CCCAGTCACGACGTTGTAAAACG	F, pUC M13F
md60	AGCGGATAACAATTTCACACAGG	R, pUC M13R
md61	ATTTAGGTGACACTATAG	R, SP6
md62	ACAGTTGAAGCTATGACACG	F, IP blunt deletion of IcsA PATR without scar
md63	TTTTCCGTCCCAGCTATTTATA	R, IP blunt deletion of IcsA PATR without scar
md64	CATGTTATTCTCAATAATAGTTCGTC	F, IP blunt deletion of IcsA PT without scar
md65	AGTAAGTTCAGCGTTTTCTTTC	R, IP blunt deletion of IcsA PT without scar
md66	GGCCAGCGCCCATGCTACTCCTCTTTCGGGTACTC	F, IP deletion of $IcsA_{SS}$ adding 5' $SfmC_{SS}$ overhang
md67	CATCAATAACTTTATTTTAGTCATCATGCACTATATTATCAGTAAGTGGTTG	R, IP deletion of $IcsA_{SS}$ adding 3' $SfmC_{SS}$ overhang
md68	ATGATGACTAAAATAAAGTTATTGATGCTCATTATATTTTATTTAATCATTTCGGCCAGCGCCCATGCT	F, OA, GA, $SfmC_{SS}$
md69	CATGGACGAGCTGTACAAGACTCCTCTTTCGGGTACTC	F, GA, IP add mCherry ends between $IcsA_{SS}$ / IcsAα
md70	CCTCGCCCTTGCTCACCATAGCAAAAGCTATTGGCCCC	R, GA, IP add mCherry ends between $IcsA_{SS}$ / IcsAα
md71	ATGGTGAGCAAGGGCGAG	F, amplify *mCherry* ORF
md72	CTTGTACAGCTCGTCCATG	R, amplify *mCherry* ORF, no stop codon
md73	AGCATGGGCGCTGGCCGAAATGATTAAATAAAATATAATGAGCATCAATAACTTTATTTTAGTCATCAT	R, OA, GA, $SfmC_{SS}$, reverse complement
md74	CCCTAACAAAATCAGCGGAGGGAACTCTCATTTTGGC	F, QC, $IcsA_{G531A}$
md75	GCCAAAATGAGAGTTCCCTCCGCTGATTTTGTTAGGG	R, QC, $IcsA_{G531A}$
md76	CCTACTCTGCTTTCACCAACATCAATGCAGGC	F, QC, $IcsA_{G545A}$
md77	GCCTGCATTGATGTTGGTGAAAGCAGAGTAGG	R, QC, $IcsA_{G545A}$
md78	CTCTGGTTTCACCAACATCAATGCAGCCATTC	F, QC, $IcsA_{G552A}$
md79	GAATGGCTGCATTGATGTTGGTGAAACCAGAG	R, QC, $IcsA_{G552A}$

F = forward oligo, R = reverse oligo, OA = used for oligo annealing, PT = polar targeting $IcsA_{506\text{-}620}$, DTRM = domain-targeted random mutagenesis, IP = inverse PCR nt, nucleotide position, RBS = ribosome binding site, QC = QuikChange mutagenesis, p = passenger sequence, PATR = passenger-associated translocation repeat $IcsA_{526\text{-}557}$, SS = signal sequence, GA = piece for Gibson Assembly.

Appendix B - Bacterial Laboratory Strains

Strain	Notes
***E. coli* K-12**	
DH5α	F- Φ80*lac*ZΔM15 Δ(*lac*ZYA-*arg*F) U169 *rec*A1 *end*A1 *hsd*R17(r_k^-, m_k^+) *pho*A *sup*E44 *thi*-1 *gyr*A96 *rel*A1 λ-
UT5600	F- *ara-14 leuB6 secA6 lacY1 proC14 tsx-67 ⊿(ompT-fepC)266 entA403 trpE38 rfbD1 rpsL109 xyl-5 mtl-1 thi-1*
XL10Gold	Tet^RΔ(*mcrA*)*183* Δ(*mcrCB-hsdSMR-mrr*)*173 endA1 supE44 thi-1 recA1 gyrA96 relA1 lac* Hte [F´ *proAB lacI*^{q}Z*ΔM15* Tn*10*(Tet^R) Amy Cam^R].
Top10F'	F´[*lac*I^q Tn*10* (Tet^R)] *mcr*A Δ(*mrr-hsd*RMS-*mcr*BC) Φ80*lac*ZΔM15 Δ*lac*X74 *rec*A1 *ara*D139 Δ(*ara-leu*)7697 *gal*U *gal*K *rps*L *end*A1 *nup*G
Tuner	F– *ompT hsdS*B(r_B^- m_B^-) *gal dcm lacY1*
S. flexneri	
2457T	Wild-type S. flexneri 2a
RMA2041	2457T ⊿*icsA::Tc*^{R}
RMA2043	2457T ⊿*icsA::Tc*^{R} ⊿*rmlD::Km*^{R}
RMA2090	2457T ⊿*icsA::Tc*^{R}[pIcsA]
RMA2107	2457T ⊿*icsA::Tc*^{R} ⊿*rmlD::Km*^{R}[pIcsA]
RMA4353	2457T ⊿*icsA::Tc*^{R} *icsP::Km*^{R} (VP-)
RMA4378	2457T ⊿*icsA::Tc*^{R} *icsP::Km*^{R}
ETRM22	2457T *icsP::Km*^{R}
ETRM230	2457T ⊿*rmlD::Km*^{R}
ETRM233	2457T ⊿*rmlD::FRT*
ETRM240	2457T ⊿*rmlD::FRT icsP::Km*^{R}
MYRM522	RMA2041 *degP::Cm*^{R}

Tc^R = tetracycline resistance, Km^R = kanamycin resistance, Cm^R = chloramphenicol resistance, VP = virulence plasmid.

Appendix C - *E. coli* and *S. flexneri* Strains Generated

Strain	Parent[plasmid(s)]	Notes
MDRM1	Tuner[pMG421][pREP4]	P_{T5}, Strep-IcsA$_{506-620}$-TEV-GFP^{+}, LacI^{hi}, Km^{R}Ap^{R}
MDRM2	Tuner[pMG422][pREP4]	P_{T5}, Strep-IcsA$_{506-620(i532)}$-TEV-GFP^{+}, LacI^{hi}, Km^{R}Ap^{R}
MDRM3	Tuner[pMG426][pREP4]	P_{T5}, Strep-IcsA$_{506-620(i563)}$-TEV-GFP^{+}, LacI^{hi}, Km^{R}Ap^{R}
MDRM4	Tuner[pMG100::*gfp*$^{+}$][pREP4]	P_{T5}, GFP^{+}, LacI^{hi}, Km^{R}Ap^{R}
MDRM5	DH5α[pBAD30::pMG421]	P_{BAD}, Strep-IcsA$_{506-620}$-TEV-GFP^{+}, (EcoRI/NheI), Ap^{R}
MDRM6	DH5α[pBAD30::pMG422]	P_{BAD}, Strep-IcsA$_{506-620(i532)}$-TEV-GFP^{+}, (EcoRI/NheI), Ap^{R}
MDRM7	DH5α[pBAD30::pMG426]	P_{BAD}, Strep-IcsA$_{506-620(i563)}$-TEV-GFP^{+}, (EcoRI/NheI), Ap^{R}
MDRM8	DH5α[pKMRM1::His$_8$]	P_{IcsA}, pIcsA$_{i87::His8}$, Ap^{R}
MDRM9	Blank	
MDRM10	UT5600[pMDRM8]	OmpT^{-}, P_{IcsA}, pIcsA$_{i87::His8}$, Ap^{R}
MDRM11	RMA2041[pMDRM8]	2457T IcsA^{-}, VP^{+}, P_{IcsA}, pIcsA$_{i87::His8}$, Tc^{R}Ap^{R}
MDRM12	RMA2041[pBR322]	2457T IcsA-, VP^{-}, vector control, Tc^{R}Ap^{R}
MDRM13	Blank	
MDRM14	Blank	
MDRM15	Blank	
MDRM16	DH5α[pIcsA]	P_{IcsA}, pIcsA storage, Ap^{R}
MDRM17	UT5600[pIcsA]	OmpT^{-}, P_{IcsA}, pIcsA storage, Ap^{R}
MDRM18	XL10Gold[pIcsA$_{DTRM(506-620)}$]	P_{IcsA}, pIcsA, IcsA$_{506-620}$ random mutant library, Ap^{R}
MDRM19	DH5α[pAM2⊥*murG*::NheI]	P_{Trc}, in frame fusions to mCherry using NheI site, Ap^{R}
MDRM20	DH5α[pGEMT-E::*icsA*]	*icsA* ORF (except stop codon) flanked by NheI, Ap^{R}
MDRM21	DH5α[pMDRM19::*icsA*]	P_{Trc}, IcsA$_{1-1102}$-mCherry, Ap^{R} (1)
MDRM22	DH5α[pMDRM19::*icsA*]	P_{Trc}, IcsA$_{1-1102}$-mCherry, Ap^{R} (2)
MDRM23	Top10F'[pMDRM22]	P_{Trc}, IcsA$_{1-1102}$-mCherry, LacI^{q}, Ap^{R}
MDRM24	XL10Gold[pIcsA$_{E532A}$]	P_{IcsA}, pIcsA$_{E532A}$ storage, Ap^{R} (1)
MDRM25	XL10Gold[pIcsA$_{E532A}$]	P_{IcsA}, pIcsA$_{E532A}$ storage, Ap^{R} (2)
MDRM26	XL10Gold[pIcsA$_{G533A}$]	P_{IcsA}, pIcsA$_{G533A}$ storage, Ap^{R}
MDRM27	XL10Gold[pIcsA$_{T563A}$]	P_{IcsA}, pIcsA$_{T563A}$ storage, Ap^{R}
MDRM28	XL10Gold[pIcsA$_{R564A}$]	P_{IcsA}, pIcsA$_{R564A}$ storage, Ap^{R} (1)
MDRM29	XL10Gold[pIcsA$_{R564A}$]	P_{IcsA}, pIcsA$_{R564A}$ storage, Ap^{R} (2)
MDRM30	DH5α[pGEMT-E::pMDRM22ΔPT]	pMDRM22 in pGEMT-E, Xhol replace IcsA$_{506-620}$, Ap^{R}
MDRM31	Blank	
MDRM32	Blank	
MDRM33	DH5α[pGEMT-E::*icsA*$_{i87::His8}$]	IcsA$_{i87::His8}$ ORF flanked by KpnI-RBS/SalI, Ap^{R}
MDRM34	DH5α[pBAD33:: *icsA*$_{i87::His8}$]	P_{BAD}, IcsA$_{i87::His8}$ expression, pMDRM33 KpnI/SalI, Cml^{R}
MDRM35	UT5600[pMDRM34]	OmpT^{-}, P_{BAD}, IcsA$_{i87::His8}$ expression vector, Cml^{R}
MDRM36	UT5600[pBAD33]	OmpT^{-}, P_{BAD}, vector control, Cml^{R}
MDRM37	RMA2041[pMDRM24]	2457T IcsA^{-}, VP^{+}, P_{IcsA}, pIcsA$_{E532A}$, Tc^{R}Ap^{R}
MDRM38	RMA2041[pMDRM26]	2457T IcsA^{-}, VP^{+}, P_{IcsA}, pIcsA$_{G533A}$, Tc^{R}Ap^{R}
MDRM39	RMA2041[pMDRM27]	2457T IcsA^{-}, VP^{+}, P_{IcsA}, pIcsA$_{T563A}$, Tc^{R}Ap^{R}
MDRM40	RMA2041[pMDRM28]	2457T IcsA^{-}, VP^{+}, P_{IcsA}, pIcsA$_{R564A}$, Tc^{R}Ap^{R}
MDRM41	DH5α[pMDRM22ΔPT]	P_{Trc}, IcsA$_{1-1102(Δ506-620)}$-mCherry, Xhol replace PT, Ap^{R}

MDRM42	Top10F'[pMDRM41]	P_{Trc}, $IcsA_{1-1102(\Delta 506-620)}$-mCherry, $LacI^q$, Tc^RAp^R
MDRM43	RMA2041[$pIcsA_{M562R}$]	2457T IcsA$^-$, VP$^+$, P_{IcsA}, $pIcsA_{M562R}$, Tc^RAp^R
MDRM44	RMA2041[$pIcsA_{W522R}$]	2457T IcsA$^-$, VP$^+$, P_{IcsA}, $pIcsA_{W522R}$, Tc^RAp^R
MDRM45	RMA2041[$pIcsA_{G531W}$]	2457T IcsA$^-$, VP$^+$, P_{IcsA}, $pIcsA_{G531W}$, Tc^RAp^R
MDRM46	DH5α[pMDRM43]	P_{IcsA}, $pIcsA_{M562R}$ storage, Ap^R
MDRM47	DH5α[pMDRM44]	P_{IcsA}, $pIcsA_{W522R}$ storage, Ap^R
MDRM48	DH5α[pMDRM45]	P_{IcsA}, $pIcsA_{G531W}$ storage, Ap^R
MDRM49	DH5α[pKMRM16]	P_{IcsA}, $pIcsA_{i532}$ storage, Ap^R
MDRM50	DH5α[pKMRM11]	P_{IcsA}, $pIcsA_{i563}$ storage, Ap^R
MDRM51	DH5α[pCACTUS::$icsA_{W522R}$]	Suicide vector with $icsA_{W522R}$ allele, Cml^R, (1)
MDRM52	DH5α[pCACTUS::$icsA_{W522R}$]	Suicide vector with $icsA_{W522R}$ allele, Cml^R, (2)
MDRM53	MDRM44[pKJE7]	2457T IcsA$^-$, VP$^+$, P_{IcsA}, $pIcsA_{W522R}$, high DnaKJ/GrpE, $Tc^RAp^RCml^R$
MDRM54	RMA2043[pMDRM44]	2457T IcsA$^-$ RmlD$^-$, VP$^+$, P_{IcsA}, $pIcsA_{W522R}$, $Tc^RKm^RAp^R$
MDRM55	RMA2043[pMDRM28]	2457T IcsA$^-$ RmlD$^-$, VP$^+$, P_{IcsA}, $pIcsA_{R564A}$, $Tc^RKm^RAp^R$
MDRM56	RMA4353[pIcsA]	2457T IcsA$^-$ IcsP$^-$, VP$^-$, P_{IcsA}, pIcsA, $Tc^RKm^RAp^R$
MDRM57	RMA4353[$pIcsA^{NC}$]	2457T IcsA$^-$ IcsP$^-$, VP$^-$, P_{IcsA}, $pIcsA^{NC}$, $Tc^RKm^RAp^R$
MDRM58	RMA4353[pMDRM28]	2457T IcsA$^-$ IcsP$^-$, VP$^-$, P_{IcsA}, $pIcsA_{R564A}$, $Tc^RKm^RAp^R$
MDRM59	XL10Gold[$pIcsA^{NC}{}_{G533A}$]	P_{IcsA}, $pIcsA^{NC}{}_{G533A}$, Ap^R
MDRM60	XL10Gold[$pIcsA^{NC}{}_{R564A}$]	P_{IcsA}, $pIcsA^{NC}{}_{R564A}$, Ap^R
MDRM61	RMA2041[pMDRM59]	2457T IcsA$^-$, VP$^+$, P_{IcsA}, $pIcsA^{NC}{}_{G533A}$, Tc^RAp^R
MDRM62	RMA2043[pMDRM26]	2457T IcsA$^-$ RmlD$^-$, VP$^+$, P_{IcsA}, $pIcsA_{G533A}$, $Tc^RKm^RAp^R$
MDRM63	RMA2043[pMDRM27]	2457T IcsA$^-$ RmlD$^-$, VP$^+$, P_{IcsA}, $pIcsA_{T563A}$, $Tc^RKm^RAp^R$
MDRM64	RMA2043[pMDRM45]	2457T IcsA$^-$ RmlD$^-$, VP$^+$, P_{IcsA}, $pIcsA_{G531W}$, $Tc^RKm^RAp^R$
MDRM65	RMA2043[pMDRM43]	2457T IcsA$^-$ RmlD$^-$, VP$^+$, P_{IcsA}, $pIcsA_{M562R}$, $Tc^RKm^RAp^R$
MDRM66	RMA2041[pMDRM60]	2457T IcsA$^-$, VP$^+$, P_{IcsA}, $pIcsA^{NC}{}_{R564A}$, Tc^RAp^R
MDRM67	2457T[pKJE7]	WT, high DnaKJ/GrpE, Cml^R
MDRM68	RMA2041[pBR322]	2457T IcsA$^-$, VP$^+$, vector control, Tc^RAp^R
MDRM69	RMA2043[pBR322]	2457T IcsA$^-$ RmlD$^-$, VP$^+$, vector control, $Tc^RKm^RAp^R$
MDRM70	RMA4378[pBR322]	2457T IcsA$^-$ IcsP$^-$, VP$^+$, vector control, $Tc^RKm^RAp^R$
MDRM71	RMA4378[pIcsA]	2457T IcsA$^-$ IcsP$^-$, VP$^+$, P_{IcsA}, pIcsA, $Tc^RKm^RAp^R$
MDRM72	RMA4378[$pIcsA^{NC}$]	2457T IcsA$^-$ IcsP$^-$, VP$^+$, P_{IcsA}, $pIcsA^{NC}$, $Tc^RKm^RAp^R$
MDRM73	RMA4378[pMDRM26]	2457T IcsA$^-$ IcsP$^-$, VP$^+$, P_{IcsA}, $pIcsA_{G533A}$, $Tc^RKm^RAp^R$
MDRM74	RMA4378[pMDRM27]	2457T IcsA$^-$ IcsP$^-$, VP$^+$, P_{IcsA}, $pIcsA_{T563A}$, $Tc^RKm^RAp^R$
MDRM75	RMA4378[pMDRM28]	2457T IcsA$^-$ IcsP$^-$, VP$^+$, P_{IcsA}, $pIcsA_{R564A}$, $Tc^RKm^RAp^R$
MDRM76	RMA4378[pMDRM46]	2457T IcsA$^-$ IcsP$^-$, VP$^+$, P_{IcsA}, $pIcsA_{M562R}$, $Tc^RKm^RAp^R$
MDRM77	RMA4378[pMDRM47]	2457T IcsA$^-$ IcsP$^-$, VP$^+$, P_{IcsA}, $pIcsA_{W522R}$, $Tc^RKm^RAp^R$
MDRM78	RMA4378[pMDRM48]	2457T IcsA$^-$ IcsP$^-$, VP$^+$, P_{IcsA}, $pIcsA_{G531W}$, $Tc^RKm^RAp^R$
MDRM79	Top10F'[pMDRM19]	for mCherry fusion, $LacI^q$, Tc^RAp^R
MDRM80	DH5α[pGEMT-E::$icsAp$]	$IcsA_{1-758}$, no stop codon, flanked by NheI, Ap^R
MDRM81	DH5α[pGEMT-E::$icsA_{104-758}$]	$IcsA_{104-758}$, no stop codon, flanked by NheI, Ap^R
MDRM82	Top10F'[pMDRM19::$icsA_{104-758}$]	P_{Trc}, $IcsA_{104-758}$-mCherry, $LacI^q$, Tc^RAp^R
MDRM83	Top10F'[pMDRM19::$icsAp$]	P_{Trc}, $IcsA_{1-758}$-mCherry, $LacI^q$, Tc^RAp^R
MDRM84	MYRM522[pMDRM26]	2457T IcsA$^-$ DegP$^-$, VP$^+$, P_{IcsA}, $pIcsA_{G533A}$, $Tc^RKm^RAp^R$
MDRM85	MYRM522[pMDRM27]	2457T IcsA$^-$ DegP$^-$, VP$^+$, P_{IcsA}, $pIcsA_{T563A}$, $Tc^RKm^RAp^R$
MDRM86	MYRM522[pMDRM28]	2457T IcsA$^-$ DegP$^-$, VP$^+$, P_{IcsA}, $pIcsA_{R564A}$, $Tc^RKm^RAp^R$
MDRM87	MYRM522[pMDRM46]	2457T IcsA$^-$ DegP$^-$, VP$^+$, P_{IcsA}, $pIcsA_{M562R}$, $Tc^RKm^RAp^R$

MDRM88	MYRM522[pMDRM47]	2457T IcsA$^-$ DegP$^-$, VP$^+$, P_{IcsA}, pIcsA$_{W522R}$, Tc^{R}Km^{R}Ap^R
MDRM89	MYRM522[pMDRM48]	2457T IcsA$^-$ DegP$^-$, VP$^+$, P_{IcsA}, pIcsA$_{G531W}$, Tc^{R}Km^{R}Ap^R
MDRM90	MYRM522[pKMRM11]	2457T IcsA$^-$ DegP$^-$, VP$^+$, P_{IcsA}, pIcsA$_{i563}$, Tc^{R}Km^{R}Ap^R
MDRM91	MYRM522[pKMRM16]	2457T IcsA$^-$ DegP$^-$, VP$^+$, P_{IcsA}, pIcsA$_{i532}$, Tc^{R}Km^{R}Ap^R
MDRM92	Top10F'[pMDRM82ΔPT]	P_{Trc}, IcsA$_{104\text{-}758(\Delta 506\text{-}620}$-mCherry, LacI^q, Tc^{R}Ap^R
MDRM93	Top10F'[pMDRM83ΔPT]	P_{Trc}, IcsA$_{1\text{-}758(\Delta 506\text{-}620}$-mCherry, LacI^q, Tc^{R}Ap^R
MDRM94	RMA2041[pMDRM19]	2457T IcsA$^-$, VP$^+$, P_{Trc}, mCherry, Tc^{R}Ap^R
MDRM95	RMA2041[pMDRM82]	2457T IcsA$^-$, VP$^+$, P_{Trc}, IcsA$_{104\text{-}758}$-mCherry, Tc^{R}Ap^R
MDRM96	RMA2041[pMDRM83]	2457T IcsA$^-$, VP$^+$, P_{Trc}, IcsA$_{1\text{-}758}$-mCherry, Tc^{R}Ap^R
MDRM97	RMA2041[pMDRM92]	2457T IcsA$^-$, VP$^+$, P_{Trc}, IcsA$_{104\text{-}758(\Delta 506\text{-}620}$-mCherry, Tc^{R}Ap^R
MDRM98	RMA2043[pMDRM19]	2457T IcsA$^-$ RmlD$^-$, VP$^+$, P_{Trc}, mCherry, Tc^{R}Ap^R
MDRM99	RMA2043[pMDRM82]	2457T IcsA$^-$ RmlD$^-$, VP$^+$, P_{Trc}, IcsA$_{104\text{-}758}$-mCherry, Tc^{R}Ap^R
MDRM100	RMA2043[pMDRM83]	2457T IcsA$^-$ RmlD$^-$, VP$^+$, P_{Trc}, IcsA$_{1\text{-}758}$-mCherry, Tc^{R}Ap^R
MDRM101	RMA2043[pMDRM92]	2457T IcsA$^-$ RmlD$^-$, VP$^+$, P_{Trc}, IcsA$_{104\text{-}758(\Delta 506\text{-}620}$-mCherry, Tc^{R}Ap^R
MDRM102	DH5α[pBAD30::*icsA*$_{i87::His8}$]	P_{BAD}, IcsA$_{i87::His8}$ expression, pMDRM33 kpnI/SalI, Ap^R (1)
MDRM103	DH5α[pBAD30::*icsA*$_{i87::His8}$]	P_{BAD}, IcsA$_{i87::His8}$ expression, pMDRM33 kpnI/SalI, Ap^R (2)
MDRM104	RMA2041[pMDRM34]	2457T IcsA$^-$, VP$^+$, P_{BAD}, IcsA$_{i87::His8}$ express, Tc^{R}Cml^R
MDRM105	RMA2041[pMDRM102]	2457T IcsA$^-$, VP$^+$, P_{BAD}, IcsA$_{i87::His8}$ express, Tc^{R}Ap^R
MDRM106	RMA2043[pMDRM34]	2457T IcsA$^-$ RmlD$^-$, VP$^+$, P_{BAD}, IcsA$_{i87::His8}$ express, Tc^{R}Cml^R
MDRM107	RMA2043[pMDRM102]	2457T IcsA$^-$ RmlD$^-$, VP$^+$, P_{BAD}, IcsA$_{i87::His8}$ express, Tc^{R}Ap^R
MDRM108	RMA2041[pBAD30]	2457T IcsA$^-$, VP$^+$, vector control, Tc^{R}Ap^R
MDRM109	RMA2041[pBAD33]	2457T IcsA$^-$, VP$^+$, vector control, Tc^{R}Cml^R
MDRM110	RMA2043[pBAD30]	2457T IcsA$^-$ RmlD$^-$, VP$^+$, vector control, Tc^{R}Ap^R
MDRM111	RMA2043[pBAD33]	2457T IcsA$^-$ RmlD$^-$, VP$^+$, vector control, Tc^{R}Cml^R
MDRM112	DH5α[pGEMT-E::*icsA*]	*icsA* ORF flanked by KpnI-RBS/Stop-SalI, Ap^R
MDRM113	DH5α[pGEMT-E::*icsAp*]	*icsA*$_{1\text{-}758}$ flanked by KpnI-RBS/Stop-SalI, Ap^R (1)
MDRM114	DH5α[pGEMT-E::*icsAp*]	*icsA*$_{1\text{-}758}$ flanked by KpnI-RBS/Stop-SalI, Ap^R (2)
MDRM115	DH5α[pBAD30::*icsA*]	P_{BAD}, IcsA expression, pMDRM112 KpnI/SalI, Ap^R
MDRM116	DH5α[pBAD30::*icsAp*]	P_{BAD}, IcsA$_{1\text{-}758}$ expression, pMDRM113 KpnI/SalI, Ap^R
MDRM117	UT5600[pBAD30]	OmpT$^-$, vector control, Ap^R
MDRM118	UT5600[pMDRM116]	OmpT$^-$, P_{BAD}, IcsA$_{1\text{-}758}$ express, Ap^R
MDRM119	RMA2041[pMDRM116]	2457T IcsA$^-$, VP$^+$, P_{BAD}, IcsA$_{1\text{-}758}$ express, Ap^R
MDRM120	RMA2043[pMDRM116]	2457T IcsA$^-$ RmlD$^-$, VP$^+$, P_{BAD}, IcsA$_{1\text{-}758}$ express, Ap^R
MDRM121	2457T[pBAD30]	WT, vector control, Ap^R
MDRM122	2457T[pMDRM116]	WT, P_{BAD}, IcsA$_{1\text{-}758}$ express, Ap^R
MDRM123	DH5α[pBAD33::*icsAp*]	P_{BAD}, IcsA$_{1\text{-}758}$ expression, pMDRM113 KpnI/SalI, Cml^R
MDRM124	RMA2090[pMDRM123]	RMA2041[pIcsA], VP$^+$, P_{BAD}, IcsA$_{1\text{-}758}$, Tc^{R}Ap^{R}Cml^R
MDRM125	RMA2090[pBAD33]	RMA2041[pIcsA], VP$^+$, vector control, Tc^{R}Ap^{R}Cml^R
MDRM126	MDRM68[pMDRM123]	RMA2041[pBR322], VP$^+$, P_{BAD}, IcsA$_{1\text{-}758}$, Tc^{R}Ap^{R}Cml^R
MDRM127	MDRM68[pBAD33]	RMA2041[pBR322], VP$^+$, 2 x vect. control Tc^{R}Ap^{R}Cml^R(1)
MDRM128	MDRM68[pBAD33]	RMA2041[pBR322], VP$^+$, 2 x vect. control Tc^{R}Ap^{R}Cml^R(2)
MDRM129	RMA2041[pMDRM115]	2457T IcsA$^-$, VP$^+$, P_{BAD}, IcsA express, Tc^{R}Ap^R
MDRM130	DH5α[pGEMT-E::*icsAp*$_{i87::His8}$]	*icsA*$_{1\text{-}758(i87::his8)}$ flanked by KpnI-RBS/Stop-SalI, Ap^R
MDRM131	UT5600[pMDRM115]	OmpT$^-$, P_{BAD}, IcsA expression, Ap^R (1)
MDRM132	UT5600[pMDRM115]	OmpT$^-$, P_{BAD}, IcsA expression, Ap^R (2)
MDRM133	XL10Gold[pMDRM115$_{G533A}$]	P_{BAD}, IcsA$_{G533A}$ expression, Ap^R (1)

MDRM134	XL10Gold[pMDRM115$_{G533A}$]	P_{BAD}, IcsA$_{G533A}$ expression, Ap^R (2)
MDRM135	RMA2043[pMDRM115]	2457T IcsA$^-$ RmlD$^-$, VP$^+$, P_{BAD}, IcsA express, Tc^{R}Km^{R}Ap^R
MDRM136	RMA2043[pMDRM115]	2457T IcsA$^-$ RmlD$^-$, VP$^+$, P_{BAD}, IcsA express, Tc^{R}Km^{R}Ap^R
MDRM137	UT5600[pMDRM133]	OmpT$^-$, P_{BAD}, IcsA$_{G533A}$ expression, Ap^R
MDRM138	DH5α[pIcsAΔPT]	P_{IcsA}, IcsA$_{\Delta 506-620}$, no scar, Ap^R
MDRM139	DH5α[pBAD30::*icsA*$p_{i87::His8}$]	P_{BAD}, IcsA$_{1-758(i87::His8)}$ express, Ap^R (1)
MDRM140	DH5α[pBAD30::*icsA*$p_{i87::His8}$]	P_{BAD}, IcsA$_{1-758(i87::His8)}$ express, Ap^R (2)
MDRM141	DH5α[pMDRM115ΔPATR]	P_{BAD}, IcsA$_{\Delta 526-557}$ express, no scar, Ap^R (1)
MDRM142	DH5α[pMDRM115ΔPATR]	P_{BAD}, IcsA$_{\Delta 526-557}$ express, no scar, Ap^R (2)
MDRM143	RMA2041[pMDRM138]	2457T IcsA$^-$, VP$^+$, P_{IcsA}, IcsA$_{\Delta 506-620}$, Tc^{R}Ap^R (1)
MDRM144	RMA2041[pMDRM138]	2457T IcsA$^-$, VP$^+$, P_{IcsA}, IcsA$_{\Delta 506-620}$, Tc^{R}Ap^R (1)
MDRM145	XL10Gold[pIcsA$_{G531A}$]	P_{IcsA}, pIcsA$_{G531A}$, Ap^R
MDRM146	DH5α[pIcsA$_{G545A}$]	P_{IcsA}, pIcsA$_{G545A}$, Ap^R
MDRM147	XL10Gold[pIcsA$_{G552A}$]	P_{IcsA}, pIcsA$_{G552A}$, Ap^R
MDRM148	XL10Gold[pMDRM115$_{G531A}$]	P_{BAD}, IcsA$_{G531A}$ expression, Ap^R
MDRM149	DH5α[pMDRM115$_{G545A}$]	P_{BAD}, IcsA$_{G545A}$ expression, Ap^R
MDRM150	XL10Gold[pMDRM115$_{G552A}$]	P_{BAD}, IcsA$_{G552A}$ expression, Ap^R
MDRM151	RMA2043[pMDRM138]	2457T IcsA$^-$ RmlD$^-$, VP$^+$, P_{IcsA}, IcsA$_{\Delta 506-620}$, Tc^{R}Ap^R (1)
MDRM152	RMA2043[pMDRM138]	2457T IcsA$^-$ RmlD$^-$, VP$^+$, P_{IcsA}, IcsA$_{\Delta 506-620}$, Tc^{R}Ap^R (2)
MDRM153	UT5600[pMDRM141]	OmpT$^-$, P_{BAD}, IcsA$_{\Delta 526-557}$ express, no scar, Ap^R (1)
MDRM154	UT5600[pMDRM141]	OmpT$^-$, P_{BAD}, IcsA$_{\Delta 526-557}$ express, no scar, Ap^R (2)
MDRM155	DH5α[pIcsAΔPATR]	P_{IcsA}, IcsA$_{\Delta 526-557}$, no scar, Ap^R (1)
MDRM156	RMA4378[pMDRM115]	2457T IcsA$^-$ IcsP$^-$, VP$^+$, P_{BAD}, IcsA expression, Tc^{R}Km^{R}Ap^R
MDRM157	RMA4378[pMDRM115]	2457T IcsA$^-$ IcsP$^-$, VP$^+$, P_{BAD}, IcsA expression, Tc^{R}Km^{R}Ap^R
MDRM158	RMA4378[pBAD30]	2457T IcsA$^-$ IcsP$^-$, VP$^+$, P_{BAD}, vector control, Tc^{R}Km^{R}Ap^R
MDRM159	RMA4378[pBAD30]	2457T IcsA$^-$ IcsP$^-$, VP$^+$, P_{BAD}, vector control, Tc^{R}Km^{R}Ap^R
MDRM160	DH5α[pIcsAΔPATR]	P_{IcsA}, IcsA$_{\Delta 526-557}$, no scar, Ap^R (2)
MDRM161	DH5α[pGEMT-E::SfmC$_{SS}$]	SfmC$_{SS}$ storage
MDRM162	RMA4378[pMDRM141]	2457T IcsA$^-$ IcsP$^-$, VP$^+$, P_{BAD}, IcsA$_{\Delta PATR}$ express, Tc^{R}Km^{R}Ap^R
MDRM163	RMA4378[pMDRM141]	2457T IcsA$^-$ IcsP$^-$, VP$^+$, P_{BAD}, IcsA$_{\Delta PATR}$ express, Tc^{R}Km^{R}Ap^R
MDRM164	RMA4378[pMDRM148]	2457T IcsA$^-$ IcsP$^-$, VP$^+$, P_{BAD}, IcsA$_{G531A}$ express, Tc^{R}Km^{R}Ap^R
MDRM165	RMA4378[pMDRM148]	2457T IcsA$^-$ IcsP$^-$, VP$^+$, P_{BAD}, IcsA$_{G531A}$ express, Tc^{R}Km^{R}Ap^R
MDRM166	RMA4378[pMDRM133]	2457T IcsA$^-$ IcsP$^-$, VP$^+$, P_{BAD}, IcsA$_{G533A}$ express, Tc^{R}Km^{R}Ap^R
MDRM167	RMA4378[pMDRM133]	2457T IcsA$^-$ IcsP$^-$, VP$^+$, P_{BAD}, IcsA$_{G533A}$ express, Tc^{R}Km^{R}Ap^R
MDRM168	RMA4378[pMDRM138]	2457T IcsA$^-$ IcsP$^-$, VP$^+$, P_{IcsA}, IcsA$_{\Delta 506-620}$, Tc^{R}Km^{R}Ap^R
MDRM169	RMA4378[pMDRM138]	2457T IcsA$^-$ IcsP$^-$, VP$^+$, P_{IcsA}, IcsA$_{\Delta 506-620}$, Tc^{R}Km^{R}Ap^R
MDRM170	RMA4378[pMDRM149]	2457T IcsA$^-$ IcsP$^-$, VP$^+$, P_{BAD}, IcsA$_{G545A}$ express, Tc^{R}Km^{R}Ap^R
MDRM171	RMA4378[pMDRM149]	2457T IcsA$^-$ IcsP$^-$, VP$^+$, P_{BAD}, IcsA$_{G545A}$ express, Tc^{R}Km^{R}Ap^R
MDRM172	RMA4378[pMDRM150]	2457T IcsA$^-$ IcsP$^-$, VP$^+$, P_{BAD}, IcsA$_{G552A}$ express, Tc^{R}Km^{R}Ap^R
MDRM173	RMA4378[pMDRM150]	2457T IcsA$^-$ IcsP$^-$, VP$^+$, P_{BAD}, IcsA$_{G552A}$ express, Tc^{R}Km^{R}Ap^R
MDRM174	DH5α[pMDRM115ΔPT]	P_{BAD}, IcsA$_{\Delta 506-620}$ express, no scar, Ap^R
MDRM175	DH5α[pMDRM115::*mCherry*]	P_{BAD}, IcsA$_{SS}$-mCherry-$_{105-1102}$ express, no scar, Ap^R
MDRM176	RMA2041[pMDRM175]	2457T IcsA$^-$, VP$^+$, P_{BAD}, IcsA$_{SS}$-mCherry-IcsA$_{105-1102}$, Tc^{R}Ap^R
MDRM177	RMA2041[pMDRM145]	2457T IcsA$^-$, VP$^+$, P_{IcsA}, pIcsA$_{G531A}$, Tc^{R}Ap^R
MDRM178	RMA2041[pMDRM146]	2457T IcsA$^-$, VP$^+$, P_{IcsA}, pIcsA$_{G545A}$, Tc^{R}Ap^R
MDRM179	RMA2041[pMDRM147]	2457T IcsA$^-$, VP$^+$, P_{IcsA}, pIcsA$_{G552A}$, Tc^{R}Ap^R
MDRM180	RMA2041[pMDRM155]	2457T IcsA$^-$, VP$^+$, P_{IcsA}, pIcsA$_{\Delta PATR}$, Tc^{R}Ap^R

MDRM181	RMA2043[pMDRM145]	2457T IcsA$^-$ RmlD$^-$, VP$^+$, P_{IcsA}, pIcsA$_{G531A}$, Tc^{R}Ap^R
MDRM182	RMA2043[pMDRM146]	2457T IcsA$^-$ RmlD$^-$, VP$^+$, P_{IcsA}, pIcsA$_{G545A}$, Tc^{R}Ap^R
MDRM183	RMA2043[pMDRM147]	2457T IcsA$^-$ RmlD$^-$, VP$^+$, P_{IcsA}, pIcsA$_{G552A}$, Tc^{R}Ap^R
MDRM184	RMA2043[pMDRM155]	2457T IcsA$^-$ RmlD$^-$, VP$^+$, P_{IcsA}, pIcsA$_{\Delta PATR}$, Tc^{R}Ap^R

P = promoter type, Km^R = kanamycin resistance Ap^R = ampicillin resistance, Tc^R = tetracycline resistance, Cml^R = chloramphenicol resistance, VP = virulence plasmid, RBS = ribosome binding site, DTRM = domain-targeted random mutagenesis, PT = polar targeting IcsA$_{506-620}$, NC = non-cleavable, WT = wild type *S. flexneri* 2a 2457T, p = passenger IcsA$_{1-758}$, PATR = passenger-associated translocation repeat IcsA$_{526-557}$.

Appendix D - Thesis Bibliography

Achtman M, Manning PA, Edelbluth C & Herrlich P (1979) Export without proteolytic processing of inner and outer membrane proteins encoded by F sex factor tra cistrons in *Escherichia* coli minicells. *Proc Natl Acad Sci U S A* 76(10): 4837-4841.

Adler B, Sasakawa C, Tobe T, Makino S, Komatsu K & Yoshikawa M (1989) A dual transcriptional activation system for the 230 kb plasmid genes coding for virulence-associated antigens of *Shigella flexneri*. *Mol Microbiol* 3(5): 627-635.

Ambrosi C, Pompili M, Scribano D, Zagaglia C, Ripa S & Nicoletti M (2012) Outer membrane protein A (OmpA): a new player in *shigella flexneri* protrusion formation and inter-cellular spreading. *PLoS One* 7(11): e49625.

Ames GF, Prody C & Kustu S (1984) Simple, rapid, and quantitative release of periplasmic proteins by chloroform. *J Bacteriol* 160(3): 1181-1183.

Anton IM, Jones GE, Wandosell F, Geha R & Ramesh N (2007) WASP-interacting protein (WIP): working in polymerisation and much more. *Trends Cell Biol* 17(11): 555-562.

Ashida H, Mimuro H & Sasakawa C (2015) *Shigella* manipulates host immune responses by delivering effector proteins with specific roles. *Front Immunol* 6219.

Baker KS, Dallman TJ, Ashton PM*, et al* (2015) Intercontinental dissemination of azithromycin-resistant shigellosis through sexual transmission: a cross-sectional study. *Lancet Infect Dis* 10.1016/S1473-3099(15)00002-X.

Bardhan P, Faruque AS, Naheed A & Sack DA (2010) Decrease in shigellosis-related deaths without *Shigella spp.*-specific interventions, Asia. *Emerg Infect Dis* 16(11): 1718-1723.

Barnard TJ, Dautin N, Lukacik P, Bernstein HD & Buchanan SK (2007) Autotransporter structure reveals intra-barrel cleavage followed by conformational changes. *Nat Struct Mol Biol* 14(12): 1214-1220.

Barzu S, Benjelloun-Touimi Z, Phalipon A, Sansonetti P & Parsot C (1997) Functional analysis of the *Shigella flexneri* IpaC invasin by insertional mutagenesis. *Infect Immun* 65(5): 1599-1605.

Bennion D, Charlson ES, Coon E & Misra R (2010) Dissection of beta-barrel outer membrane protein assembly pathways through characterizing BamA POTRA 1 mutants of *Escherichia coli*. *Mol Microbiol* 77(5): 1153-1171.

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(10): 3867-3871.

Bernstein HD (2015) Looks can be deceiving: recent insights into the mechanism of protein secretion by the autotransporter pathway. *Mol Microbiol* 10.1111/mmi.13031.

Besingi RN, Chaney JL & Clark PL (2013) An alternative outer membrane secretion mechanism for an autotransporter protein lacking a C-terminal stable core. *Mol Microbiol* 90(5): 1028-1045.

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(3): 683-693.

Bolivar F, Rodriguez RL, Greene PJ, Betlach MC, Heyneker HL, Boyer HW, Crosa JH & Falkow S (1977) Construction and characterization of new cloning vehicles. II. A multipurpose cloning system. *Gene* 2(2): 95-113.

Brandon LD & Goldberg MB (2001) Periplasmic transit and disulfide bond formation of the autotransported *Shigella* protein IcsA. *J Bacteriol* 183(3): 951-958.

Brandon LD, Goehring N, Janakiraman A, Yan AW, Wu T, Beckwith J & Goldberg MB (2003) IcsA, a polarly localized autotransporter with an atypical signal peptide, uses the Sec apparatus for secretion, although the Sec apparatus is circumferentially distributed. *Mol Microbiol* 50(1): 45-60.

Braselmann E & Clark PL (2012) Autotransporters: The cellular environment reshapes a folding mechanism to promote protein transport. *J Phys Chem Lett* 3(8): 1063-1071.

Brosius J (1989) Superpolylinkers in cloning and expression vectors. *DNA* 8(10): 759-777.

Brotcke Zumsteg A, Goosmann C, Brinkmann V, Morona R & Zychlinsky A (2014) IcsA is a *Shigella flexneri* adhesion regulated by the type III secretion system and required for pathogenesis. *Cell Host Microbe* 15(4): 435-445.

Calloni G, Chen T, Schermann SM, Chang HC, Genevaux P, Agostini F, Tartaglia GG, Hayer-Hartl M & Hartl FU (2012) DnaK functions as a central hub in the *E. coli* chaperone network. *Cell Rep* 1(3): 251-264.

Campbell-Valois FX, Sachse M, Sansonetti PJ & Parsot C (2015) Escape of Actively Secreting *Shigella flexneri* from ATG8/LC3-Positive Vacuoles Formed during Cell-To-Cell Spread Is Facilitated by IcsB and VirA. *MBio* 6(3): e02567-02514.

Carrio MM & Villaverde A (2005) Localization of chaperones DnaK and GroEL in bacterial inclusion bodies. *J Bacteriol* 187(10): 3599-3601.

Castanie-Cornet MP, Bruel N & Genevaux P (2013) Chaperone networking facilitates protein targeting to the bacterial cytoplasmic membrane. *Biochim Biophys Acta* 10.1016/j.bbamcr.2013.11.007.

Celik N, Webb CT, Leyton DL*, et al* (2012) A bioinformatic strategy for the detection, classification and analysis of bacterial autotransporters. *PLoS One* 7(8): e43245.

Chamaillard M, Hashimoto M, Horie Y*, et al* (2003) An essential role for NOD1 in host recognition of bacterial peptidoglycan containing diaminopimelic acid. *Nat Immunol* 4(7): 702-707.

Charles M, Perez M, Kobil JH & Goldberg MB (2001) Polar targeting of *Shigella* virulence factor IcsA in *Enterobacteriacae* and *Vibrio*. *Proc Natl Acad Sci U S A* 98(17): 9871-9876.

Chatterjee S & Rothenberg E (2012) Interaction of bacteriophage l with its E. coli receptor, LamB. *Viruses* 4(11): 3162-3178.

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

Collins KD, Lacal J & Ottemann KM (2014) Internal sense of direction: sensing and signaling from cytoplasmic chemoreceptors. *Microbiol Mol Biol Rev* 78(4): 672-684.

d'Hauteville H, Dufourcq Lagelouse R, Nato F & Sansonetti PJ (1996) Lack of cleavage of IcsA in *Shigella flexneri* causes aberrant movement and allows demonstration of a cross-reactive eukaryotic protein. *Infect Immun* 64(2): 511-517.

Dai J, Wang S, Guerlebeck D, Laturnus C, Guenther S, Shi Z, Lu C & Ewers C (2010) Suppression subtractive hybridization identifies an autotransporter adhesin gene of *E. coli* IMT5155 specifically associated with avian pathogenic *Escherichia coli* (APEC). *BMC Microbiol* 10236.

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(1): 181-194.

Datsenko KA & Wanner BL (2000) One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc Natl Acad Sci U S A* 97(12): 6640-6645.

Dautin N & Bernstein HD (2007) Protein secretion in gram-negative bacteria via the autotransporter pathway. *Annu Rev Microbiol* 6189-112.

de Boer PA, Crossley RE & Rothfield LI (1989) A division inhibitor and a topological specificity factor coded for by the minicell locus determine proper placement of the division septum in *E. coli*. *Cell* 56(4): 641-649.

de Pedro MA, Quintela JC, Holtje JV & Schwarz H (1997) Murein segregation in *Escherichia coli*. *J Bacteriol* 179(9): 2823-2834.

de Pedro MA, Young KD, Holtje JV & Schwarz H (2003) Branching of *Escherichia coli* cells arises from multiple sites of inert peptidoglycan. *J Bacteriol* 185(4): 1147-1152.

Demali KA, Jue AL & Burridge K (2006) IpaA targets beta1 integrins and rho to promote actin cytoskeleton rearrangements necessary for *Shigella* entry. *J Biol Chem* 281(51): 39534-39541.

Desvaux M, Cooper LM, Filenko NA, Scott-Tucker A, Turner SM, Cole JA & Henderson IR (2006) The unusual extended signal peptide region of the type V secretion system is phylogenetically restricted. *FEMS Microbiol Lett* 264(1): 22-30.

Desvaux M, Scott-Tucker A, Turner SM, Cooper LM, Huber D, Nataro JP & Henderson IR (2007) A conserved extended signal peptide region directs posttranslational protein translocation via a novel mechanism. *Microbiol-SGM* 153(Pt 1): 59-70.

Domingo Meza-Aguilar J, Fromme P, Torres-Larios A, Mendoza-Hernandez G, Hernandez-Chinas U, Arreguin-Espinosa de Los Monteros RA, Eslava Campos CA & Fromme R (2014) X-ray crystal structure of the passenger domain of plasmid encoded toxin(Pet), an autotransporter enterotoxin from enteroaggregative *Escherichia coli* (EAEC). *Biochem Biophys Res Commun* 445(2): 439-444.

Doyle MT, Tran EN & Morona R (2015a) The passenger-associated transport repeat promotes virulence factor secretion efficiency and delineates a distinct autotransporter subtype. *Mol Microbiol* 97(2): 315-329.

Doyle MT, Grabowicz M, May KL & Morona R (2015b) Lipopolysaccharide surface structure does not influence IcsA polarity. *FEMS Microbiol Lett* 362(8): fnv042.

Drobnak I, Braselmann E & Clark PL (2015a) Multiple driving forces required for efficient secretion of autotransporter virulence proteins. *J Biol Chem* 290(16): 10104-10116.

Drobnak I, Braselmann E, Chaney JL, Leyton DL, Bernstein HD, Lithgow T, Luirink J, Nataro JP & Clark PL (2015b) Of linkers and autochaperones: an unambiguous nomenclature to identify common and uncommon themes for autotransporter secretion. *Mol Microbiol* 95(1): 1-16.

DuPont HL, Levine MM, Hornick RB & Formal SB (1989) Inoculum size in shigellosis and implications for expected mode of transmission. *J Infect Dis* 159(6): 1126-1128.

Dworkin J (2009) Cellular polarity in prokaryotic organisms. *Cold Spring Harb Perspect Biol* 1(6): a003368.

Egan AJ & Vollmer W (2013) The physiology of bacterial cell division. *Ann N Y Acad Sci* 12778-28.

Egile C, d'Hauteville H, Parsot C & Sansonetti PJ (1997) SopA, the outer membrane protease responsible for polar localization of IcsA in *Shigella flexneri*. *Mol Microbiol* 23(5): 1063-1073.

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(6): 1319-1332.

Emsley P, Charles IG, Fairweather NF & Isaacs NW (1996) Structure of *Bordetella pertussis* virulence factor P.69 pertactin. *Nature* 381(6577): 90-92.

Fasano A, Noriega FR, Liao FM, Wang W & Levine MM (1997) Effect of *Shigella* enterotoxin 1 (ShET1) on rabbit intestine *in vitro* and *in vivo*. *Gut* 40(4): 505-511.

Fasano A, Noriega FR, Maneval DR, Jr., Chanasongcram S, Russell R, Guandalini S & Levine MM (1995) *Shigella* enterotoxin 1: an enterotoxin of *Shigella flexneri* 2a active in rabbit small intestine *in vivo* and *in vitro*. *J Clin Invest* 95(6): 2853-2861.

Fehniger TE, Radolf JD, Walfield AM, Cunningham TM, Miller JN & Lovett MA (1986) Native surface association of a recombinant 38-kilodalton Treponema pallidum antigen isolated from the Escherichia coli outer membrane. *Infect Immun* 52(2): 586-593.

Finn RD, Bateman A, Clements J*, et al* (2014a) Pfam: the protein families database. *Nucleic Acids Res* 42(Database issue): D222-230.

Finn RD, Bateman A, Clements J*, et al* (2014b) Pfam: the protein families database. *Nucleic Acids Research* 42(D1): D222-D230.

Fixen KR, Janakiraman A, Garrity S, Slade DJ, Gray AN, Karahan N, Hochschild A & Goldberg MB (2012) Genetic reporter system for positioning of proteins at the bacterial pole. *MBio* 3(2).

Fukuda I, Suzuki T, Munakata H, Hayashi N, Katayama E, Yoshikawa M & Sasakawa C (1995) Cleavage of *Shigella* surface protein VirG occurs at a specific site, but the secretion is not essential for intracellular spreading. *J Bacteriol* 177(7): 1719-1726.

Fukumatsu M, Ogawa M, Arakawa S, Suzuki M, Nakayama K, Shimizu S, Kim M, Mimuro H & Sasakawa C (2012) *Shigella* targets epithelial tricellular junctions and uses a noncanonical clathrin-dependent endocytic pathway to spread between cells. *Cell Host Microbe* 11(4): 325-336.

Gangwer KA, Mushrush DJ, Stauff DL, Spiller B, McClain MS, Cover TL & Lacy DB (2007) Crystal structure of the *Helicobacter pylori* vacuolating toxin p55 domain. *Proc Natl Acad Sci U S A* 104(41): 16293-16298.

Gawarzewski I, DiMaio F, Winterer E, Tschapek B, Smits SH, Jose J & Schmitt L (2014) Crystal structure of the transport unit of the autotransporter adhesin involved in diffuse adherence from *Escherichia coli*. *J Struct Biol* 187(1): 20-29.

Ghosh AS & Young KD (2005) Helical disposition of proteins and lipopolysaccharide in the outer membrane of Escherichia coli. *J Bacteriol* 187(6): 1913-1922.

Giangrossi M, Prosseda G, Tran CN, Brandi A, Colonna B & Falconi M (2010) A novel antisense RNA regulates at transcriptional level the virulence gene *icsA* of *Shigella flexneri*. *Nucleic Acids Res* 38(10): 3362-3375.

Gibbs KA, Isaac DD, Xu J, Hendrix RW, Silhavy TJ & Theriot JA (2004) Complex spatial distribution and dynamics of an abundant Escherichia coli outer membrane protein, LamB. *Mol Microbiol* 53(6): 1771-1783.

Girardin SE, Boneca IG, Carneiro LA*, et al* (2003) Nod1 detects a unique muropeptide from gram-negative bacterial peptidoglycan. *Science* 300(5625): 1584-1587.

Goldberg MB (2001) Actin-based motility of intracellular microbial pathogens. *Microbiol Mol Biol Rev* 65(4): 595-626.

Goldberg MB & Theriot JA (1995) *Shigella flexneri* surface protein IcsA is sufficient to direct actin-based motility. *Proc Natl Acad Sci U S A* 92(14): 6572-6576.

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(8): 2189-2196.

Gorden J & Small PL (1993) Acid resistance in enteric bacteria. *Infect Immun* 61(1): 364-367.

Grijpstra J, Arenas J, Rutten L & Tommassen J (2013) Autotransporter secretion: varying on a theme. *Res Microbiol* 164(6): 562-582.

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(14): 4121-4130.

Haft DH, Selengut JD & White O (2003) The TIGRFAMs database of protein families. *Nucleic Acids Research* 31(1): 371-373.

Hale TL (1991) Genetic basis of virulence in *Shigella* species. *Microbiol Rev* 55(2): 206-224.

Henderson IR & Lam AC (2001) Polymorphic proteins of *Chlamydia spp.*-autotransporters beyond the Proteobacteria. *Trends Microbiol* 9(12): 573-578.

Henderson IR, Navarro-Garcia F, Desvaux M, Fernandez RC & Ala'Aldeen D (2004) Type V protein secretion pathway: the autotransporter story. *Microbiol Mol Biol Rev* 68(4): 692-744.

Heras B, Totsika M, Peters KM, Paxman JJ, Gee CL, Jarrott RJ, Perugini MA, Whitten AE & Schembri MA (2014) The antigen 43 structure reveals a molecular Velcro-like mechanism of autotransporter-mediated bacterial clumping. *Proc Natl Acad Sci U S A* 111(1): 457-462.

High N, Mounier J, Prevost MC & Sansonetti PJ (1992) IpaB of Shigella flexneri causes entry into epithelial cells and escape from the phagocytic vacuole. *Embo J* 11(5): 1991-1999.

Hilbi H, Moss JE, Hersh D, Chen Y, Arondel J, Banerjee S, Flavell RA, Yuan J, Sansonetti PJ & Zychlinsky A (1998) *Shigella*-induced apoptosis is dependent on caspase-1 which binds to IpaB. *J Biol Chem* 273(49): 32895-32900.

Hobb RI, Fields JA, Burns CM & Thompson SA (2009) Evaluation of procedures for outer membrane isolation from *Campylobacter jejuni*. *Microbiol-SGM* 155(Pt 3): 979-988.

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(4): 779-791.

Horton RM (1993) In Vitro Recombination and Mutagenesis of DNA (Chapter 25). *Methods in Molecular Biology, Vol 15: PCR Protocols: Current Methods and Applications,* Vol. 15 (White BA, ed.) p.^pp. 251-261. Humana Press Inc., Totowa, NJ.

Horton RM, Cai ZL, Ho SN & Pease LR (1990) Gene splicing by overlap extension: tailor-made genes using the polymerase chain reaction. *Biotechniques* 8(5): 528-535.

Hromockyj AE, Tucker SC & Maurelli AT (1992) Temperature regulation of Shigella virulence: identification of the repressor gene virR, an analogue of hns, and partial complementation by tyrosyl transfer RNA (tRNA1(Tyr)). *Mol Microbiol* 6(15): 2113-2124.

Huang KC, Mukhopadhyay R & Wingreen NS (2006) A curvature-mediated mechanism for localization of lipids to bacterial poles. *PLoS Comput Biol* 2(11): e151.

Hunter S, Jones P, Mitchell A*, et al* (2012) InterPro in 2011: new developments in the family and domain prediction database. *Nucleic Acids Res* 40(Database issue): D306-312.

Ieva R & Bernstein HD (2009) Interaction of an autotransporter passenger domain with BamA during its translocation across the bacterial outer membrane. *Proc Natl Acad Sci U S A* 106(45): 19120-19125.

Ieva R, Skillman KM & Bernstein HD (2008) Incorporation of a polypeptide segment into the beta-domain pore during the assembly of a bacterial autotransporter. *Mol Microbiol* 67(1): 188-201.

Ieva R, Tian P, Peterson JH & Bernstein HD (2011) Sequential and spatially restricted interactions of assembly factors with an autotransporter beta domain. *Proc Natl Acad Sci U S A* 108(31): E383-391.

Jain E, Bairoch A, Duvaud S, Phan I, Redaschi N, Suzek BE, Martin MJ, McGarvey P & Gasteiger E (2009) Infrastructure for the life sciences: design and implementation of the UniProt website. *BMC Bioinform* 10136.

Jain S & Goldberg MB (2007) Requirement for YaeT in the outer membrane assembly of autotransporter proteins. *J Bacteriol* 189(14): 5393-5398.

Jain S, van Ulsen P, Benz I, Schmidt MA, Fernandez R, Tommassen J & Goldberg MB (2006) Polar localization of the autotransporter family of large bacterial virulence proteins. *J Bacteriol* 188(13): 4841-4850.

Janakiraman A, Fixen KR, Gray AN, Niki H & Goldberg MB (2009) A genome-scale proteomic screen identifies a role for DnaK in chaperoning of polar autotransporters in *Shigella*. *J Bacteriol* 191(20): 6300-6311.

Jensen VB, Harty JT & Jones BD (1998) Interactions of the invasive pathogens *Salmonella typhimurium*, *Listeria monocytogenes*, and *Shigella flexneri* with M cells and murine Peyer's patches. *Infect Immun* 66(8): 3758-3766.

Jin Q, Yuan Z, Xu J*, et al* (2002) Genome sequence of *Shigella flexneri* 2a: insights into pathogenicity through comparison with genomes of *Escherichia coli* K12 and O157. *Nucleic Acids Res* 30(20): 4432-4441.

Johnson TA, Qiu J, Plaut AG & Holyoak T (2009) Active-site gating regulates substrate selectivity in a chymotrypsin-like serine protease the structure of *haemophilus influenzae* immunoglobulin A1 protease. *J Mol Biol* 389(3): 559-574.

Jones P, Binns D, Chang HY*, et al* (2014) InterProScan 5: genome-scale protein function classification. *Bioinformatics* 30(9): 1236-1240.

Jong WS, ten Hagen-Jongman CM, den Blaauwen T, Slotboom DJ, Tame JR, Wickstrom D, de Gier JW, Otto BR & Luirink J (2007) Limited tolerance towards folded elements during secretion of the autotransporter Hbp. *Mol Microbiol* 63(5): 1524-1536.

Junker M, Besingi RN & Clark PL (2009) Vectorial transport and folding of an autotransporter virulence protein during outer membrane secretion. *Mol Microbiol* 71(5): 1323-1332.

Junker M, Schuster CC, McDonnell AV, Sorg KA, Finn MC, Berger B & Clark PL (2006) Pertactin beta-helix folding mechanism suggests common themes for the secretion and folding of autotransporter proteins. *Proc Natl Acad Sci U S A* 103(13): 4918-4923.

Kabanov DS & Prokhorenko IR (2010) Structural analysis of lipopolysaccharides from Gram-negative bacteria. *Biochemistry (Mosc)* 75(4): 383-404.

Kajava AV & Steven AC (2006) The turn of the screw: variations of the abundant beta-solenoid motif in passenger domains of Type V secretory proteins. *J Struct Biol* 155(2): 306-315.

Kang'ethe W & Bernstein HD (2013a) Stepwise folding of an autotransporter passenger domain is not essential for its secretion. *J Biol Chem* 288(49): 35028-35038.

Kang'ethe W & Bernstein HD (2013b) Charge-dependent secretion of an intrinsically disordered protein via the autotransporter pathway. *Proc Natl Acad Sci U S A* 110(45): E4246-4255.

Kaufmann A, Stierhof YD & Henning U (1994) New outer membrane-associated protease of *Escherichia coli* K-12. *J Bacteriol* 176(2): 359-367.

Kayath CA, Hussey S, El hajjami N, Nagra K, Philpott D & Allaoui A (2010) Escape of intracellular *Shigella* from autophagy requires binding to cholesterol through the type III effector, IcsB. *Microbes Infect* 12(12-13): 956-966.

Kazmierczak BI & Hendrixson DR (2013) Spatial and numerical regulation of flagellar biosynthesis in polarly flagellated bacteria. *Mol Microbiol* 88(4): 655-663.

Keddy KH, Sooka A, Crowther-Gibson P*, et al* (2012) Systemic Shigellosis in South Africa. *Clin Infect Dis*, 22474223.

Khan S, Mian HS, Sandercock LE, Chirgadze NY & Pai EF (2011) Crystal structure of the passenger domain of the *Escherichia coli* autotransporter EspP. *J Mol Biol* 413(5): 985-1000.

Kida Y, Taira J, Yamamoto T, Higashimoto Y & Kuwano K (2013) EprS, an autotransporter protein of *Pseudomonas aeruginosa,* possessing serine protease activity induces inflammatory responses through protease-activated receptors. *Cell Microbiol* 15(7): 1168-1181.

Kim AS, Kakalis LT, Abdul-Manan N, Liu GA & Rosen MK (2000) Autoinhibition and activation mechanisms of the Wiskott-Aldrich syndrome protein. *Nature* 404(6774): 151-158.

Kingsley RA, van Amsterdam K, Kramer N & Baumler AJ (2000) The *shdA* gene is restricted to serotypes of *Salmonella enterica* subspecies I and contributes to efficient and prolonged fecal shedding. *Infect Immun* 68(5): 2720-2727.

Kingsley RA, Keestra AM, de Zoete MR & Baumler AJ (2004a) The ShdA adhesin binds to the cationic cradle of the fibronectin 13FnIII repeat module: evidence for molecular mimicry of heparin binding. *Mol Microbiol* 52(2): 345-355.

Kingsley RA, Santos RL, Keestra AM, Adams LG & Baumler AJ (2002) *Salmonella enterica* serotype Typhimurium ShdA is an outer membrane fibronectin-binding protein that is expressed in the intestine. *Mol Microbiol* 43(4): 895-905.

Kingsley RA, Abi Ghanem D, Puebla-Osorio N, Keestra AM, Berghman L & Baumler AJ (2004b) Fibronectin binding to the *Salmonella enterica* serotype Typhimurium ShdA autotransporter protein is inhibited by a monoclonal antibody recognizing the A3 repeat. *J Bacteriol* 186(15): 4931-4939.

Kocks C, Marchand JB, Gouin E, d'Hauteville H, Sansonetti PJ, Carlier MF & Cossart P (1995) The unrelated surface proteins ActA of *Listeria monocytogenes* and IcsA of *Shigella flexneri* are sufficient to confer actin-based motility on *Listeria innocua* and *Escherichia coli* respectively. *Mol Microbiol* 18(3): 413-423.

Koppelman CM, Den Blaauwen T, Duursma MC, Heeren RM & Nanninga N (2001) *Escherichia coli* minicell membranes are enriched in cardiolipin. *J Bacteriol* 183(20): 6144-6147.

Kotloff KL, Winickoff JP, Ivanoff B, Clemens JD, Swerdlow DL, Sansonetti PJ, Adak GK & Levine MM (1999) Global burden of infections: implications *Shigella* for vaccine development and implementation of control strategies. *Bull World Health Organ* 77(8): 651-666.

Kovach ME, Elzer PH, Hill DS, Robertson GT, Farris MA, Roop RM, 2nd & Peterson KM (1995) Four new derivatives of the broad-host-range cloning vector pBBR1MCS, carrying different antibiotic-resistance cassettes. *Gene* 166(1): 175-176.

Kownhar H, Shankar EM, Rajan R, Vengatesan A & Rao UA (2007) Prevalence of *Campylobacter jejuni* and enteric bacterial pathogens among hospitalized HIV infected versus non-HIV infected patients with diarrhoea in southern India. *Scand J Infect Dis* 39(10): 862-866.

Kuhnel K & Diezmann D (2011) Crystal structure of the autochaperone region from the *Shigella flexneri* autotransporter IcsA. *J Bacteriol* 193(8): 2042-2045.

LaBrec E, Schneider H, Magnani T & SB F (1964) Epithelial cell penetration as an essential step in the pathogenesis of bacillary dysentery. *J Bacteriol* 881503-1518.

Lafont F, Tran Van Nhieu G, Hanada K, Sansonetti P & van der Goot FG (2002) Initial steps of *Shigella* infection depend on the cholesterol/sphingolipid raft-mediated CD44-IpaB interaction. *Embo J* 21(17): 4449-4457.

Lai EM, Nair U, Phadke ND & Maddock JR (2004) Proteomic screening and identification of differentially distributed membrane proteins in *Escherichia coli*. *Mol Microbiol* 52(4): 1029-1044.

Laloux G & Jacobs-Wagner C (2014) How do bacteria localize proteins to the cell pole? *J Cell Sci* 127(Pt 1): 11-19.

Lan R, Alles MC, Donohoe K, Martinez MB & Reeves PR (2004) Molecular evolutionary relationships of enteroinvasive *Escherichia coli* and *Shigella* spp. *Infect Immun* 72(9): 5080-5088.

Landick R, Wade JT & Grainger DC (2015) H-NS and RNA polymerase: a love-hate relationship? *Curr Opin Microbiol* 2453-59.

Leo JC, Oberhettinger P, Schutz M & Linke D (2015) The inverse autotransporter family: intimin, invasin and related proteins. *Int J Med Microbiol* 305(2): 276-282.

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(1): 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(7): 540-553.

Levine OS & Levine MM (1991) Houseflies (*Musca domestica*) as mechanical vectors of shigellosis. *Rev Infect Dis* 13(4): 688-696.

Leyton DL, Rossiter AE & Henderson IR (2012) From self sufficiency to dependence: mechanisms and factors important for autotransporter biogenesis. *Nat Rev Microbiol* 10(3): 213-225.

Leyton DL, de Luna MG, Sevastsyanovich YR, Tveen Jensen K, Browning DF, Scott-Tucker A & Henderson IR (2010) The unusual extended signal peptide region is not required for secretion and function of an *Escherichia coli* autotransporter. *FEMS Microbiol Lett* 311(2): 133-139.

Leyton DL, Sevastsyanovich YR, Browning DF, Rossiter AE, Wells TJ, Fitzpatrick RE, Overduin M, Cunningham AF & Henderson IR (2011) Size and conformation limits to secretion of disulfide-bonded loops in autotransporter proteins. *J Biol Chem* 286(49): 42283-42291.

Leyton DL, Johnson MD, Thapa R*, et al* (2014) A mortise-tenon joint in the transmembrane domain modulates autotransporter assembly into bacterial outer membranes. *Nat Commun* 54239.

Lima IF, Havt A & Lima AA (2015) Update on molecular epidemiology of Shigella infection. *Curr Opin Gastroenterol* 31(1): 30-37.

Lindner AB, Madden R, Demarez A, Stewart EJ & Taddei F (2008) Asymmetric segregation of protein aggregates is associated with cellular aging and rejuvenation. *Proc Natl Acad Sci U S A* 105(8): 3076-3081.

Livio S, Strockbine NA, Panchalingam S*, et al* (2014) *Shigella* isolates from the global enteric multicenter study inform vaccine development. *Clin Infect Dis* 59(7): 933-941.

Lommel S, Benesch S, Rottner K, Franz T, Wehland J & Kuhn R (2001) Actin pedestal formation by enteropathogenic *Escherichia coli* and intracellular motility of *Shigella flexneri* are abolished in N-WASP-defective cells. *EMBO Rep* 2(9): 850-857.

Lu Q, Yao Q, Xu Y*, et al* (2014) An iron-containing dodecameric heptosyltransferase family modifies bacterial autotransporters in pathogenesis. *Cell Host Microbe* 16(3): 351-363.

Luck SN, Turner SA, Rajakumar K, Sakellaris H & Adler B (2001) Ferric dicitrate transport system (Fec) of *Shigella flexneri* 2a YSH6000 is encoded on a novel pathogenicity island carrying multiple antibiotic resistance genes. *Infect Immun* 69(10): 6012-6021.

Lugtenberg B, Meijers J, Peters R, van der Hoek P & van Alphen L (1975) Electrophoretic resolution of the "major outer membrane protein" of *Escherichia coli* K12 into four bands. *FEBS Lett* 58(1): 254-258.

Lutkenhaus J & Addinall SG (1997) Bacterial cell division and the Z ring. *Annu Rev Biochem* 6693-116.

Lynne AM, Skyberg JA, Logue CM & Nolan LK (2007) Detection of Iss and Bor on the surface of Escherichia coli. *J Appl Microbiol* 102(3): 660-666.

Magdalena J & Goldberg MB (2002) Quantification of *Shigella* IcsA required for bacterial actin polymerization. *Cell Motil Cytoskeleton* 51(4): 187-196.

Magrane M & Consortium U (2011) UniProt Knowledgebase: a hub of integrated protein data. *Database* 2011bar009.

Makino S, Sasakawa C, Kamata K, Kurata T & Yoshikawa M (1986) A genetic determinant required for continuous reinfection of adjacent cells on large plasmid in *S. flexneri* 2a. *Cell* 46(4): 551-555.

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(13): 4666-4676.

May KL, Grabowicz M, Polyak SW & Morona R (2012) Self-association of the *Shigella flexneri* IcsA autotransporter protein. *Microbiol-SGM* 158(Pt 7): 1874-1883.

Meng G, Spahich N, Kenjale R, Waksman G & St Geme JW, 3rd (2011) Crystal structure of the *Haemophilus influenzae* Hap adhesin reveals an intercellular oligomerization mechanism for bacterial aggregation. *Embo J* 30(18): 3864-3874.

Michaelis AM & Gitai Z (2010) Dynamic polar sequestration of excess MurG may regulate enzymatic function. *J Bacteriol* 192(18): 4597-4605.

Miki H & Takenawa T (2003) Regulation of actin dynamics by WASP family proteins. *J Biochem* 134(3): 309-313.

Monack DM & Theriot JA (2001) Actin-based motility is sufficient for bacterial membrane protrusion formation and host cell uptake. *Cell Microbiol* 3(9): 633-647.

Moraes TF, Spreter T & Strynadka NC (2008) Piecing together the type III injectisome of bacterial pathogens. *Curr Opin Struct Biol* 18(2): 258-266.

Moreau V, Frischknecht F, Reckmann I, Vincentelli R, Rabut G, Stewart D & Way M (2000) A complex of N-WASP and WIP integrates signalling cascades that lead to actin polymerization. *Nat Cell Biol* 2(7): 441-448.

Morona R & Van Den Bosch L (2003a) Multicopy *icsA* is able to suppress the virulence defect caused by the *wzz(SF)* mutation in *Shigella flexneri*. *FEMS Microbiol Lett* 221(2): 213-219.

Morona R & Van Den Bosch L (2003b) Lipopolysaccharide O antigen chains mask IcsA (VirG) in Shigella flexneri. *FEMS Microbiol Lett* 221(2): 173-180.

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(4): 1059-1068.

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. *Microbiol-SGM* 149(Pt 4): 925-939.

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(1): 237-248.

Murayama SY, Sakai T, Makino S, Kurata T, Sasakawa C & Yoshikawa M (1986) The use of mice in the Sereny test as a virulence assay of *shigellae* and enteroinvasive *Escherichia coli*. *Infect Immun* 51(2): 696-698.

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(5): 1395-1406.

Musiime V, Kalyesubula I, Kaddu-Mulindwa D & Byarugaba J (2009) Enteric bacterial pathogens in HIV-infected children with acute diarrhea in Mulago referral and teaching hospital, Kampala, Uganda. *J Int Assoc Physicians AIDS Care (Chic)* 8(3): 185-190.

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(19): 5069-5078.

Niyogi SK (2005) Shigellosis. *J Microbiol* 43(2): 133-143.

Noinaj N, Rollauer SE & Buchanan SK (2015) The beta-barrel membrane protein insertase machinery from Gram-negative bacteria. *Curr Opin Struct Biol* 3135-42.

Noinaj N, Kuszak AJ, Balusek C, Gumbart JC & Buchanan SK (2014) Lateral opening and exit pore formation are required for BamA function. *Structure* 22(7): 1055-1062.

Noinaj N, Kuszak AJ, Gumbart JC, Lukacik P, Chang H, Easley NC, Lithgow T & Buchanan SK (2013) Structural insight into the biogenesis of beta-barrel membrane proteins. *Nature* 501(7467): 385-390.

Oaks EV, Wingfield ME & Formal SB (1985) Plaque formation by virulent *Shigella flexneri*. *Infect Immun* 48(1): 124-129.

Ogawa M, Suzuki T, Tatsuno I, Abe H & Sasakawa C (2003) IcsB, secreted via the type III secretion system, is chaperoned by IpgA and required at the post-invasion stage of *Shigella* pathogenicity. *Mol Microbiol* 48(4): 913-931.

Ogawa M, Yoshimori T, Suzuki T, Sagara H, Mizushima N & Sasakawa C (2005) Escape of intracellular *Shigella* from autophagy. *Science* 307(5710): 727-731.

Oliver DC, Huang G, Nodel E, Pleasance S & Fernandez RC (2003) A conserved region within the *Bordetella pertussis* autotransporter BrkA is necessary for folding of its passenger domain. *Mol Microbiol* 47(5): 1367-1383.

Oomen CJ, van Ulsen P, van Gelder P, Feijen M, Tommassen J & Gros P (2004) Structure of the translocator domain of a bacterial autotransporter. *Embo J* 23(6): 1257-1266.

Osborn MJ & Munson R (1974) Separation of the inner (cytoplasmic) and outer membranes of Gram-negative bacteria. *Methods Enzymol* 31(Pt A): 642-653.

Otto BR, Sijbrandi R, Luirink J, Oudega B, Heddle JG, Mizutani K, Park SY & Tame JR (2005) Crystal structure of hemoglobin protease, a heme binding autotransporter protein from pathogenic *Escherichia coli*. *J Biol Chem* 280(17): 17339-17345.

Page AL, Ohayon H, Sansonetti PJ & Parsot C (1999) The secreted IpaB and IpaC invasins and their cytoplasmic chaperone IpgC are required for intercellular dissemination of *Shigella flexneri*. *Cell Microbiol* 1(2): 183-193.

Papadopoulos M & 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(13): 3385-3393.

Pavlova O, Peterson JH, Ieva R & Bernstein HD (2013) Mechanistic link between beta barrel assembly and the initiation of autotransporter secretion. *Proc Natl Acad Sci U S A* 110(10): E938-947.

Perdomo JJ, Gounon P & Sansonetti PJ (1994) Polymorphonuclear leukocyte transmigration promotes invasion of colonic epithelial monolayer by *Shigella flexneri*. *J Clin Invest* 93(2): 633-643.

Peterson JH, Tian P, Ieva R, Dautin N & Bernstein HD (2010) Secretion of a bacterial virulence factor is driven by the folding of a C-terminal segment. *Proc Natl Acad Sci U S A* 107(41): 17739-17744.

Philpott DJ, Yamaoka S, Israel A & Sansonetti PJ (2000) Invasive *Shigella flexneri* activates NF-kappa B through a lipopolysaccharide-dependent innate intracellular response and leads to IL-8 expression in epithelial cells. *J Immunol* 165(2): 903-914.

Pina-Pedrero S, Olvera A, Perez-Simo M & Bensaid A (2012) Genomic and antigenic characterization of monomeric autotransporters of *Haemophilus parasuis*: an ongoing process of reductive evolution. *Microbiol-SGM* 158(Pt 2): 436-447.

Prevost MC, Lesourd M, Arpin M, Vernel F, Mounier J, Hellio R & Sansonetti PJ (1992) Unipolar reorganization of F-actin layer at bacterial division and bundling of actin filaments by plastin correlate with movement of *Shigella flexneri* within HeLa cells. *Infect Immun* 60(10): 4088-4099.

Prosseda G, Falconi M, Nicoletti M, Casalino M, Micheli G & Colonna B (2002) Histone-like proteins and the *Shigella* invasivity regulon. *Res Microbiol* 153(7): 461-468.

Prosseda G, Fradiani PA, Di Lorenzo M, Falconi M, Micheli G, Casalino M, Nicoletti M & Colonna B (1998) A role for H-NS in the regulation of the virF gene of *Shigella* and enteroinvasive *Escherichia coli*. *Res Microbiol* 149(1): 15-25.

Pugsley AP & Buddelmeijer N (2004) Traffic spotting: poles apart. *Mol Microbiol* 53(6): 1559-1562.

Purdy GE, Hong M & Payne SM (2002) *Shigella flexneri* DegP facilitates IcsA surface expression and is required for efficient intercellular spread. *Infect Immun* 70(11): 6355-6364.

Purdy GE, Fisher CR & Payne SM (2007) IcsA surface presentation in *Shigella flexneri* requires the periplasmic chaperones DegP, Skp, and SurA. *J Bacteriol* 189(15): 5566-5573.

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 Wzz_{pHS2}. *Microbiol-SGM* 154(Pt 4): 1104-1116.

Rawlings ND, Waller M, Barrett AJ & Bateman A (2014) MEROPS: the database of proteolytic enzymes, their substrates and inhibitors. *Nucleic Acids Res* 42(Database issue): D503-509.

Reeves PR, Hobbs M, Valvano MA*, et al* (1996) Bacterial polysaccharide synthesis and gene nomenclature. *Trends Microbiol* 4(12): 495-503.

Renn JP & Clark PL (2008) A conserved stable core structure in the passenger domain beta-helix of autotransporter virulence proteins. *Biopolymers* 89(5): 420-427.

Renn JP, Junker M, Besingi RN, Braselmann E & Clark PL (2012) ATP-independent control of autotransporter virulence protein transport via the folding properties of the secreted protein. *Chem Biol* 19(2): 287-296.

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(4): 861-872.

Robert V, Volokhina EB, Senf F, Bos MP, Van Gelder P & Tommassen J (2006) Assembly factor Omp85 recognizes its outer membrane protein substrates by a species-specific C-terminal motif. *PLoS Biol* 4(11): e377.

Rokney A, Shagan M, Kessel M, Smith Y, Rosenshine I & Oppenheim AB (2009) *E. coli* transports aggregated proteins to the poles by a specific and energy-dependent process. *J Mol Biol* 392(3): 589-601.

Roman-Hernandez G, Peterson JH & Bernstein HD (2014) Reconstitution of bacterial autotransporter assembly using purified components. *Elife* 3e04234.

Romantsov T, Helbig S, Culham DE, Gill C, Stalker L & Wood JM (2007) Cardiolipin promotes polar localization of osmosensory transporter ProP in *Escherichia coli*. *Mol Microbiol* 64(6): 1455-1465.

Rose RE (1988) The nucleotide sequence of pACYC184. *Nucleic Acids Res* 16(1): 355.

Rossiter AE, Leyton DL, Tveen-Jensen K*, et al* (2011) The essential beta-barrel assembly machinery complex components BamD and BamA are required for autotransporter biogenesis. *J Bacteriol* 193(16): 4250-4253.

Roy A, Kucukural A & Zhang Y (2010) I-TASSER: a unified platform for automated protein structure and function prediction. *Nat Protoc* 5(4): 725-738.

Ruiz-Perez F & Nataro JP (2014) Bacterial serine proteases secreted by the autotransporter pathway: classification, specificity, and role in virulence. *Cell Mol Life Sci* 71(5): 745-770.

Ruiz-Perez F, Henderson IR & Nataro JP (2010) Interaction of FkpA, a peptidyl-prolyl cis/trans isomerase with EspP autotransporter protein. *Gut Microbes* 1(5): 339-344.

Ruiz-Perez F, Henderson IR, Leyton DL, Rossiter AE, Zhang Y & Nataro JP (2009) Roles of periplasmic chaperone proteins in the biogenesis of serine protease autotransporters of *Enterobacteriaceae*. *J Bacteriol* 191(21): 6571-6583.

Sabate R, de Groot NS & Ventura S (2010) Protein folding and aggregation in bacteria. *Cell Mol Life Sci* 67(16): 2695-2715.

Saibil HR (2013) Machinery to reverse irreversible aggregates. *Science* 339(6123): 1040-1041.

Sakai T, Sasakawa C & Yoshikawa M (1988) Expression of four virulence antigens of *Shigella flexneri* is positively regulated at the transcriptional level by the 30 kiloDalton VirF protein. *Mol Microbiol* 2(5): 589-597.

Sandlin RC, Goldberg MB & Maurelli AT (1996) Effect of O side-chain length and composition on the virulence of Shigella flexneri 2a. *Mol Microbiol* 22(1): 63-73.

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(1): 229-237.

Sansonetti PJ (2001) Rupture, invasion and inflammatory destruction of the intestinal barrier by *Shigella*, making sense of prokaryote-eukaryote cross-talks. *FEMS Microbiol Rev* 25(1): 3-14.

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(3): 193-203.

Sansonetti PJ, Arondel J, Cavaillon JM & Huerre M (1995) Role of interleukin-1 in the pathogenesis of experimental shigellosis. *J Clin Invest* 96(2): 884-892.

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(3): 1471-1480.

Sansonetti PJ, Phalipon A, Arondel J, Thirumalai K, Banerjee S, Akira S, Takeda K & Zychlinsky A (2000) Caspase-1 activation of IL-1beta and IL-18 are essential for *Shigella flexneri*-induced inflammation. *Immunity* 12(5): 581-590.

Santapaola D, Del Chierico F, Petrucca A, Uzzau S, Casalino M, Colonna B, Sessa R, Berlutti F & Nicoletti M (2006) Apyrase, the product of the virulence plasmid-encoded *phoN2* (*apy*) gene of *Shigella flexneri*, is necessary for proper unipolar IcsA localization and for efficient intercellular spread. *J Bacteriol* 188(4): 1620-1627.

Sauri A, Ten Hagen-Jongman CM, van Ulsen P & Luirink J (2012) Estimating the size of the active translocation pore of an autotransporter. *J Mol Biol* 416(3): 335-345.

Sauri A, Soprova Z, Wickstrom D, de Gier JW, Van der Schors RC, Smit AB, Jong WS & Luirink J (2009) The Bam (Omp85) complex is involved in secretion of the autotransporter haemoglobin protease. *Microbiol-SGM* 155(Pt 12): 3982-3991.

Scholz O, Thiel A, Hillen W & Niederweis M (2000) Quantitative analysis of gene expression with an improved green fluorescent protein. *Eur J Biochem* 267(6): 1565-1570.

Schroeder GN & Hilbi H (2008) Molecular pathogenesis of *Shigella spp.*: controlling host cell signaling, invasion, and death by type III secretion. *Clin Microbiol Rev* 21(1): 134-156.

Schuch R, Sandlin RC & Maurelli AT (1999) A system for identifying post-invasion functions of invasion genes: requirements for the Mxi-Spa type III secretion pathway of *Shigella flexneri* in intercellular dissemination. *Mol Microbiol* 34(4): 675-689.

Schuster-Bockler B, Schultz J & Rahmann S (2004) HMM Logos for visualization of protein families. *BMC Bioinformatics* 57.

Scribano D, Petrucca A, Pompili M*, et al* (2014) Polar localization of PhoN2, a periplasmic virulence-associated factor of *Shigella flexneri*, is required for proper IcsA exposition at the old bacterial pole. *PLoS One* 9(2): e90230.

Senerovic L, Tsunoda SP, Goosmann C, Brinkmann V, Zychlinsky A, Meissner F & Kolbe M (2012) Spontaneous formation of IpaB ion channels in host cell membranes reveals how *Shigella* induces pyroptosis in macrophages. *Cell Death Dis* 3e384.

Shere KD, Sallustio S, Manessis A, D'Aversa TG & Goldberg MB (1997) Disruption of IcsP, the major *Shigella* protease that cleaves IcsA, accelerates actin-based motility. *Mol Microbiol* 25(3): 451-462.

Shikata S, Shimada K, Kataoka H, Horinouchi S & Beppu T (1992) Detection of large COOH-terminal domains processed from the precursor of *Serratia marcescens* serine protease in the outer membrane of Escherichia coli. *J Biochem* 111(5): 627-632.

Simpore J, Ouermi D, Ilboudo D*, et al* (2009) Aetiology of acute gastro-enteritis in children at Saint Camille Medical Centre, Ouagadougou, Burkina Faso. *Pak J Biol Sci* 12(3): 258-263.

Skillman KM, Barnard TJ, Peterson JH, Ghirlando R & Bernstein HD (2005) Efficient secretion of a folded protein domain by a monomeric bacterial autotransporter. *Mol Microbiol* 58(4): 945-958.

Snapper SB, Takeshima F, Anton I*, et al* (2001) N-WASP deficiency reveals distinct pathways for cell surface projections and microbial actin-based motility. *Nat Cell Biol* 3(10): 897-904.

Soprova Z, Sauri A, van Ulsen P, Tame JR, den Blaauwen T, Jong WS & Luirink J (2010) A conserved aromatic residue in the autochaperone domain of the autotransporter Hbp is critical for initiation of outer membrane translocation. *J Biol Chem* 285(49): 38224-38233.

Speelman P, Kabir I & Islam M (1984) Distribution and spread of colonic lesions in shigellosis: a colonoscopic study. *J Infect Dis* 150(6): 899-903.

Steinhauer J, Agha R, Pham T, Varga AW & Goldberg MB (1999) The unipolar *Shigella* surface protein IcsA is targeted directly to the bacterial old pole: IcsP cleavage of IcsA occurs over the entire bacterial surface. *Mol Microbiol* 32(2): 367-377.

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(1): 23-30.

Suzek BE, Huang H, McGarvey P, Mazumder R & Wu CH (2007) UniRef: comprehensive and non-redundant UniProt reference clusters. *Bioinformatics* 23(10): 1282-1288.

Suzuki T & Sasakawa C (2001) Molecular basis of the intracellular spreading of *Shigella*. *Infect Immun* 69(10): 5959-5966.

Suzuki T, Lett MC & Sasakawa C (1995) Extracellular transport of VirG protein in *Shigella*. *J Biol Chem* 270(52): 30874-30880.

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(36): 21878-21885.

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(10): 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(4): 223-233.

Suzuki T, Nakanishi K, Tsutsui H, Iwai H, Akira S, Inohara N, Chamaillard M, Nunez G & Sasakawa C (2005) A novel caspase-1/toll-like receptor 4-independent pathway of cell death induced by cytosolic *Shigella* in infected macrophages. *J Biol Chem* 280(14): 14042-14050.

Szabady RL, Peterson JH, Skillman KM & Bernstein HD (2005) An unusual signal peptide facilitates late steps in the biogenesis of a bacterial autotransporter. *Proc Natl Acad Sci U S A* 102(1): 221-226.

Tajima N, Kawai F, Park SY & Tame JR (2010) A novel intein-like autoproteolytic mechanism in autotransporter proteins. *J Mol Biol* 402(4): 645-656.

Tamaki S, Matsuzawa H & Matsuhashi M (1980) Cluster of mrdA and mrdB genes responsible for the rod shape and mecillinam sensitivity of Escherichia coli. *J Bacteriol* 141(1): 52-57.

Taneja N, Mewara A, Kumar A, Verma G & Sharma M (2012) Cephalosporin-resistant *Shigella flexneri* over 9 years (2001-09) in India. *J Antimicrob Chemother* 67(6): 1347-1353.

Tapia-Pastrana G, Chavez-Duenas L, Lanz-Mendoza H, Teter K & Navarro-Garcia F (2012) VirK is a periplasmic protein required for efficient secretion of plasmid-encoded toxin from enteroaggregative *Escherichia coli*. *Infect Immun* 80(7): 2276-2285.

Teh MY (2012) Analysis of *Shigella flexneri* cell surface virulence factors. Mixed Thesis, The University of Adelaide.

Teh MY & Morona R (2013) Identification of *Shigella flexneri* IcsA residues affecting interaction with N-WASP, and evidence for IcsA-IcsA co-operative interaction. *PLoS One* 8(2): e55152.

Teh MY, Tran EN & Morona R (2012a) Absence of O antigen suppresses *Shigella flexneri* IcsA autochaperone region mutations. *Microbiol-SGM* 158(Pt 11): 2835-2850.

Teh MY, Tran ENH & Morona R (2012b) Absence of O antigen suppresses Shigella flexneri IcsA autochaperone region mutations. *Microbiology-Sgm* 1582835-2850.

Tiruneh M (2009) Serodiversity and antimicrobial resistance pattern of *Shigella* isolates at Gondar University teaching hospital, Northwest Ethiopia. *Jpn J Infect Dis* 62(2): 93-97.

Tran CN, Giangrossi M, Prosseda G, Brandi A, Di Martino ML, Colonna B & Falconi M (2011) A multifactor regulatory circuit involving H-NS, VirF and an antisense RNA modulates transcription of the virulence gene *icsA* of *Shigella flexneri*. *Nucleic Acids Res* 39(18): 8122-8134.

Tran EN, Doyle MT & Morona R (2013) LPS unmasking of *Shigella flexneri* reveals preferential localisation of tagged outer membrane protease IcsP to septa and new poles. *PLoS One* 8(7): e70508.

Tran EN, Papadopoulos M & Morona R (2014) Relationship between O-antigen chain length and resistance to colicin E2 in *Shigella flexneri*. *Microbiol-SGM* 160(Pt 3): 589-601.

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(12): 3249-3262.

Treuner-Lange A & Sogaard-Andersen L (2014) Regulation of cell polarity in bacteria. *J Cell Biol* 206(1): 7-17.

Turner SA, Luck SN, Sakellaris H, Rajakumar K & Adler B (2003) Molecular epidemiology of the SRL pathogenicity island. *Antimicrob Agents Chemother* 47(2): 727-734.

Typas A, Banzhaf M, van den Berg van Saparoea B*, et al* (2010) Regulation of peptidoglycan synthesis by outer-membrane proteins. *Cell* 143(7): 1097-1109.

Ursell TS, Trepagnier EH, Huang KC & Theriot JA (2012) Analysis of surface protein expression reveals the growth pattern of the gram-negative outer membrane. *PLoS Comput Biol* 8(9): e1002680.

Valencia-Gallardo CM, Carayol N & Tran Van Nhieu G (2015) Cytoskeletal mechanics during *Shigella* invasion and dissemination in epithelial cells. *Cell Microbiol* 17(2): 174-182.

Van den Berg B (2010) Crystal structure of a full-length autotransporter. *J Mol Biol* 396(3): 627-633.

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(1): 11-18.

Van Den Bosch L, Manning PA & Morona R (1997) Regulation of O-antigen chain length is required for *Shigella flexneri* virulence. *Mol Microbiol* 23(4): 765-775.

Van der Ley P, De Graaff P & Tommassen J (1986a) Shielding of Escherichia coli outer membrane proteins as receptors for bacteriophages and colicins by O-antigenic chains of lipopolysaccharide. *J Bacteriol* 168(1): 449-451.

Van der Ley P, Kuipers O, Tommassen J & Lugtenberg B (1986b) O-antigenic chains of lipopolysaccharide prevent binding of antibody molecules to an outer membrane pore protein in Enterobacteriaceae. *Microb Pathog* 1(1): 43-49.

van Ulsen P (2011) Protein folding in bacterial adhesion: secretion and folding of classical monomeric autotransporters. *Adv Exp Med Biol* 715125-142.

van Ulsen P, van Alphen L, ten Hove J, Fransen F, van der Ley P & Tommassen J (2003) A Neisserial autotransporter NalP modulating the processing of other autotransporters. *Mol Microbiol* 50(3): 1017-1030.

Vandeputte-Rutten L, Kramer RA, Kroon J, Dekker N, Egmond MR & Gros P (2001) Crystal structure of the outer membrane protease OmpT from Escherichia coli suggests a novel catalytic site. *Embo J* 20(18): 5033-5039.

Vendruscolo M, Paci E, Karplus M & Dobson CM (2003) Structures and relative free energies of partially folded states of proteins. *Proc Natl Acad Sci U S A* 100(25): 14817-14821.

von Seidlein L, Kim DR, Ali M*, et al* (2006) A multicentre study of *Shigella* diarrhoea in six Asian countries: disease burden, clinical manifestations, and microbiology. *PLoS Med* 3(9): 1556-1569.

Voorhis DL, Dillon S, Formal SB & Isberg RR (1991) An O antigen can interfere with the function of the Yersinia pseudotuberculosis invasin protein. *Mol Microbiol* 5(2): 317-325.

Voulhoux R, Bos MP, Geurtsen J, Mols M & Tommassen J (2003) Role of a highly conserved bacterial protein in outer membrane protein assembly. *Science* 299(5604): 262-265.

Wagner JK, Heindl JE, Gray AN, Jain S & Goldberg MB (2009) Contribution of the periplasmic chaperone Skp to efficient presentation of the autotransporter IcsA on the surface of *Shigella flexneri*. *J Bacteriol* 191(3): 815-821.

Walton TA & Sousa MC (2004) Crystal structure of Skp, a prefoldin-like chaperone that protects soluble and membrane proteins from aggregation. *Mol Cell* 15(3): 367-374.

Wang XY, Tao F, Xiao D*, et al* (2006) Trend and disease burden of bacillary dysentery in China (1991-2000). *Bull World Health Organ* 84(7): 561-568.

Wassef JS, Keren DF & Mailloux JL (1989) Role of M cells in initial antigen uptake and in ulcer formation in the rabbit intestinal loop model of shigellosis. *Infect Immun* 57(3): 858-863.

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(3): 991-999.

Wei J, Goldberg MB, Burland V*, et al* (2003) Complete genome sequence and comparative genomics of *Shigella flexneri* serotype 2a strain 2457T. *Infect Immun* 71(5): 2775-2786.

Weibezahn J, Schlieker C, Tessarz P, Mogk A & Bukau B (2005) Novel insights into the mechanism of chaperone-assisted protein disaggregation. *Biol Chem* 386(8): 739-744.

WHO (2005) Guidelines for the control of shigellosis, including epidemics due to *Shigella*

dysenteriae type 1. http://whqlibdoc.who.int/publications/2005/9241592330.pdf p.^pp. WHO Document Production Services.

Wing HJ, Yan AW, Goldman SR & Goldberg MB (2004) Regulation of IcsP, the outer membrane protease of the *Shigella* actin tail assembly protein IcsA, by virulence plasmid regulators VirF and VirB. *J Bacteriol* 186(3): 699-705.

Winkler J, Seybert A, Konig L, Pruggnaller S, Haselmann U, Sourjik V, Weiss M, Frangakis AS, Mogk A & Bukau B (2010) Quantitative and spatio-temporal features of protein aggregation in *Escherichia coli* and consequences on protein quality control and cellular ageing. *Embo J* 29(5): 910-923.

Xu D, Zhang J, Roy A & Zhang Y (2011) Automated protein structure modeling in CASP9 by I-TASSER pipeline combined with QUARK-based ab initio folding and FG-MD-based structure refinement. *Proteins* 79 Suppl 10147-160.

Yang J, Nie H, Chen L, Zhang X, Yang F, Xu X, Zhu Y, Yu J & Jin Q (2007) Revisiting the molecular evolutionary history of *Shigella* spp. *J Mol Evol* 64(1): 71-79.

Yen YT, Tsang C, Cameron TA, Ankrah DO, Rodou A & Stathopoulos C (2010) Importance of conserved residues of the serine protease autotransporter beta-domain in passenger domain processing and beta-barrel assembly. *Infect Immun* 78(8): 3516-3528.

Zhai Y, Zhang K, Huo Y*, et al* (2011) Autotransporter passenger domain secretion requires a hydrophobic cavity at the extracellular entrance of the beta-domain pore. *Biochem J* 435(3): 577-587.

Zhang Y & Skolnick J (2005) TM-align: a protein structure alignment algorithm based on the TM-score. *Nucleic Acids Res* 33(7): 2302-2309.

Zychlinsky A, Prevost MC & Sansonetti PJ (1992) *Shigella flexneri* induces apoptosis in infected macrophages. *Nature* 358(6382): 167-169.

Zychlinsky A, Kenny B, Menard R, Prevost MC, Holland IB & Sansonetti PJ (1994) IpaB mediates macrophage apoptosis induced by *Shigella flexneri*. *Mol Microbiol* 11(4): 619-627.