



## **Terms and Conditions of Use of Digitised Theses from Trinity College Library Dublin**

### **Copyright statement**

All material supplied by Trinity College Library is protected by copyright (under the Copyright and Related Rights Act, 2000 as amended) and other relevant Intellectual Property Rights. By accessing and using a Digitised Thesis from Trinity College Library you acknowledge that all Intellectual Property Rights in any Works supplied are the sole and exclusive property of the copyright and/or other IPR holder. Specific copyright holders may not be explicitly identified. Use of materials from other sources within a thesis should not be construed as a claim over them.

A non-exclusive, non-transferable licence is hereby granted to those using or reproducing, in whole or in part, the material for valid purposes, providing the copyright owners are acknowledged using the normal conventions. Where specific permission to use material is required, this is identified and such permission must be sought from the copyright holder or agency cited.

### **Liability statement**

By using a Digitised Thesis, I accept that Trinity College Dublin bears no legal responsibility for the accuracy, legality or comprehensiveness of materials contained within the thesis, and that Trinity College Dublin accepts no liability for indirect, consequential, or incidental, damages or losses arising from use of the thesis for whatever reason. Information located in a thesis may be subject to specific use constraints, details of which may not be explicitly described. It is the responsibility of potential and actual users to be aware of such constraints and to abide by them. By making use of material from a digitised thesis, you accept these copyright and disclaimer provisions. Where it is brought to the attention of Trinity College Library that there may be a breach of copyright or other restraint, it is the policy to withdraw or take down access to a thesis while the issue is being resolved.

### **Access Agreement**

By using a Digitised Thesis from Trinity College Library you are bound by the following Terms & Conditions. Please read them carefully.

I have read and I understand the following statement: All material supplied via a Digitised Thesis from Trinity College Library is protected by copyright and other intellectual property rights, and duplication or sale of all or part of any of a thesis is not permitted, except that material may be duplicated by you for your research use or for educational purposes in electronic or print form providing the copyright owners are acknowledged using the normal conventions. You must obtain permission for any other use. Electronic or print copies may not be offered, whether for sale or otherwise to anyone. This copy has been supplied on the understanding that it is copyright material and that no quotation from the thesis may be published without proper acknowledgement.

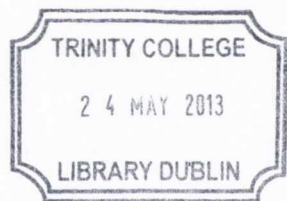
**Rational design of artificial genetic switches: Co-option of  
H-NS-repressed operons by the VirB virulence master  
regulator**

By Kelly Kane

A dissertation presented for the degree of Doctor of Philosophy, in the  
Faculty of Science, University of Dublin, Trinity College

Moyne Institute of Preventative Medicine  
School of Genetics and Microbiology  
Trinity College Dublin

July 2012



*Thesis 9890*

## Declarations

I, Kelly Kane, am the sole author of this thesis. This work presented with represents my own work except where duly acknowledged in the text and has not been previously presented for a higher degree at this or any other University. Some parts of this work have been published in the following papers;

**Kane, K. A., Dorman, C. J.** (2011) Rational design of an artificial genetic switch; Co-option of the H-NS-repressed *proU* operon by the VirB virulence master regulator. *J Bacteriol* 193(21) 5950-60

**Kane, K. A., Dorman, C. J.** (2012) VirB-mediated positive feedback control of the virulence gene regulatory cascade of *Shigella flexneri*. *J Bacteriol* (Epub ahead of print 20 July)

I agree that the library of Trinity College Dublin may lend or copy this thesis upon request.

## Summary

The H-NS protein represses the transcription of hundreds of genes in Gram-negative bacteria. De-repression is achieved by a multitude of mechanisms, many of which involve binding of a protein to DNA at the repressed promoter in a manner that compromises the maintenance of the H-NS-DNA nucleoprotein repression complex. The principal virulence gene promoters in *Shigella flexneri*, the cause of bacillary dysentery, are repressed by H-NS. Expression occurs via a regulatory cascade whereby the AraC-like protein VirF activates an intermediate regulator VirB, which in turn derepresses the operons that encode the main structural components and the effector proteins of the *S. flexneri* type III secretion system. Bioinformatic analysis suggests that VirB, a protein that closely resembles members of the ParB family of plasmid partitioning proteins, has been co-opted into its current role as an H-NS antagonist in *S. flexneri* through genetic rearrangements that positioned a binding site for VirB at a location where it could act to disrupt H-NS repression.

To test this hypothesis, the potential for VirB to act as a positive regulator of *proU*, an operon that is repressed by H-NS, was assessed. Although VirB has no known relationship with the osmo-regulated *proU* operon, it could relieve H-NS-mediated repression when the *parS*-like VirB binding site was placed appropriately upstream of the RpoD-dependent *proU* promoter. Experiments also revealed auto-regulation of the *virB* gene and positive feedback to its activator *virF*, necessary for maximal induction. These results reveal the remarkable facility with which novel regulatory circuits can evolve, at least among those promoters that are repressed by H-NS.

## Publications

### Full Papers

**Dorman, C. J., Kane, K. A.** (2009) DNA bridging and antibridging; a role for bacterial nucleoid-associated proteins in regulating the expression of laterally acquired genes.

*FEMS Microbiol Rev* **33**(3):587-92

**Kane, K. A., Dorman, C. J.** (2011) Rational design of an artificial genetic switch; Co-option of the H-NS-repressed *proU* operon by the VirB virulence master regulator. *J Bacteriol* **193**(21)5950-60

**Kane, K. A., Dorman, C. J.** (2012) VirB-mediated positive feedback control of the virulence gene regulatory cascade of *Shigella flexneri*. *J Bacteriol* (in press)

### Published Conference Abstracts

**Dorman, C. J., Kane, K. A.** (2009) Antagonism of H-NS mediated transcriptional repression by the introduction of a *cis*-acting activator binding site. Society for General Microbiology, Irish Division Meeting, Cork, Ireland.

**Dorman, C. J., Kane, K. A.** (2011) Rational design of new genetic switches; Antagonism of H-NS mediated transcriptional repression. 5<sup>th</sup> IECA Conference, Cancun, Mexico.

**Dorman, C. J., Kane, K. A.** (2011) H-NS and the global control of bacterial gene expression. Abstract Receipt Number: 30320. Proceedings of the International Union of Microbiological Societies, Sapporo, Japan.

## Acknowledgments

Firstly, I would like to thank Charlie for the wonderful opportunity to conduct my Ph.D in his lab and for all his helpful encouragement and advice throughout the last four years. I would also like to thank both past and present members of my thesis committee, Cyril, Tim, Stephen and Jay for their much appreciated help and guidance.

A special thanks to all of the prep room staff Gerry, Dave, Margaret, Miriam, Ronan and Stephen, thanks to all for your hard work, I was never short of necessary supplies. Furthermore, thanks to Stephan and Ronan for the daily laughs and banter during their morning rounds.

To Connie, thanks for all your help over the years (and ordering the armchair my first week!) To everyone in the Moyne, in particular the CJD lab for all their helpful discussions/troubleshooting/advice. Thanks to the “old” Moyne, Gar, Dan, Shane, Seratna, Enda, Alex, Tadhg, Suzy, Jimmy, Finian for Friday evenings/nights in Kennedys. Colin and Fi (KFC!), it’s never too early for champagne, especially in Galway! Joanne – our next publication is “Ph.D rants over gmail”. Andrew and Heather for an unbeliezzzable Belieze/Cancun trip! Fitzzy Fitzgerald, I couldn’t possibly put anything in writing for fear of being sued/arrested. Thanks to the “new” Moyne, Michael, Will, Conor and Aoife “the family”! I’ll never buy unnecessary tonic again!

My non-Moyne science buddies; Tiernan, Yvonne and especially Cats (sorry for abandoning you in Madrid to embark on a scientific career!)

Jacqueline, Laura, Sharon, Sarah and Louise, thanks for making my Xmas/Easters break a great escape from Dublin (albeit a little hazy!)

To the doll house, Niamh, Eadoin, Claire, Katy and Kate, for enduring years of “science is everywhere” and knowing to ignore my mumblings on Monday nights before Tuesday meetings. Somehow family dinner and a glass of wine never ended how it should!

To Isobel, for...reasons.

A huge thankyou to my parents for their continued support (both mentally and financially!) and for not asking too many questions when my “food money” went continually missing on Grafton street!

**For my parents,  
with thanks**



## Table of Contents

Title page	I	
Declaration	II	
Summary	III	
Publications	IV	
Acknowledgements	V	
Table of Contents	VI	
List of Figures	XI	
List of Tables	XII	
<b>Chapter 1</b>	<b>General Introduction</b>	
1.1	Overview	2
1.2	<i>Shigella</i>	3
1.2.1	Cellular pathogenesis of Shigellosis	5
1.3	Molecular determinants of <i>Shigella</i> pathogenicity	10
1.3.1	The large virulence plasmid	10
1.3.2	The Type III secretion system and its effector proteins	12
1.3.3	Regulatory proteins	17
1.3.3.1	VirF	17
1.3.3.2	VirB	19
1.3.3.3	MxiE	23
1.4	Chromosomal proteins involved in regulating virulence gene expression	24
1.4.1	H-NS	24
1.4.2	IHF	27
1.4.3	FIS	28
1.5	Two-component regulatory systems	29
1.6	Lux quorum sensing	31
1.7	Transcriptional regulation in prokaryotes	32

1.8	Regulation of virulence gene expression in <i>Shigella flexneri</i>	38
1.8.1	Regulatory Cascade	38
1.8.2	Regulation of <i>virF</i>	40
1.8.3	Regulation of <i>virB</i>	42
1.8.4	Regulation of <i>icsB</i> , <i>ipgD</i> , <i>spa</i> and <i>virA</i>	44
1.9	General regulation of type III secretion systems in bacteria	46
1.10	The ParB family of plasmid partitioning proteins	49
1.11	VirB; plasmid partitioning protein co-opted into a regulatory role	52
1.12	Summary	55
<b>Chapter 2</b>	<b>Materials and Methods</b>	
2.1	Chemicals and growth media	58
2.1.2	Chemicals and reagents	58
2.1.3	Growth media	58
2.1.4	Antibiotics	59
2.2	Bacterial strains and culture conditions	59
2.2.1	Bacterial strains	59
2.2.2	Bacterial culture conditions	59
2.3	Plasmids, bacteriophage and oligonucleotides	61
2.3.1	Plasmids	61
2.3.2	Bacteriophage	61
2.3.3	Oligonucleotides	61
2.4	Genetic techniques	66
2.4.1	Generalized transduction using P1 phage	61
2.4.2	Preparation of high titre P1 <i>vir</i> lysate	61
2.4.3	Transduction with P1 <i>vir</i>	66
2.5	Purification of Nucleic Acids	68
2.5.1	Isolation of plasmid DNA	68
2.5.2	Isolation of chromosomal DNA	59

2.5.3	Purification of linear DNA	70
2.5.4	Isolation of RNA	70
2.5.5	Purification of RNA by DNase treatment	72
2.5.6	Synthesis of cDNA by reverse transcription	73
2.6	Manipulation of DNA <i>in vitro</i>	73
2.6.1	Restriction endonuclease cleavage of DNA	73
2.6.2	Phosphate treatment of restriction endonuclease cleaved DNA	74
2.6.3	Ligation of DNA molecules	75
2.6.4	Amplification of DNA by PCR	75
2.6.4.1	Amplification of DNA using Phusion Polymerase	76
2.6.5	Agarose gel electrophoresis	77
2.6.6	Site-directed mutagenesis	78
2.6.7	Quantitative PCR	78
2.6.8	Allelic replacement using lambda Red	79
2.7	Transformation of bacterial strains with plasmid DNA	80
2.7.1	Preparation of calcium chloride competent cells	80
2.7.2	Transformation of calcium chloride competent cells	80
2.7.3	Preparation of electro-competent cells	81
2.7.4	Transformation of electro-competent cells	82
2.8	Assays based on spectrophotometry	82
2.8.1	Monitoring bacterial growth	82
2.8.2	Assay of $\beta$ -galactosidase activity	83
2.8.3	Determination of nucleic acid concentrations	84
2.8.4	Determination of protein concentration by the Bradford assay	84
2.9	Flow cytometry	85
2.10	Proteomic analysis	85
2.10.1	Preparation of total cellular extracts	85
2.10.2	Separation of proteins by SDS-PAGE	86
2.10.3	Visualisation of proteins in SDS-PAGE gels	87

2.10.4	Transfer of proteins to nitrocellulose membrane	88
2.10.5	Western Immunoblot analysis	88
2.10.6	Over-expression of VirB protein	89
2.10.7	Protein Purification	90
2.11	EMSA	91
2.12	DNase I footprinting using FAM labelled primers	91
<b>Chapter 3</b>	<b>The rational design of a new genetic switch</b>	
3.1	Introduction	94
3.2	Results	101
3.2.1	Creation of <i>proU-gfp</i> expression vectors	101
3.2.2	Creation of VirB dependence at the <i>proU</i> promoter	103
3.2.2.1	Insertion of the VirB binding site from the <i>icsB</i> promoter	103
3.2.2.2	Purification of VirB protein and EMSA of hybrid promoters	105
3.2.2.3	VirB protein levels and VirB function at low osmolarity	107
3.2.2.4	Insertion of a binding site for VirB confers VirB-dependent regulation at modified <i>proU</i> promoters	109
3.2.3	Chromosome integration of modified <i>proU</i> promoters	111
3.2.3.1	Expression of modified <i>proU</i> promoter on the chromosome mimics plasmid based experiments	111
3.2.3.2	VirB dependent de-repression of <i>proU</i> involves promoter P2 not P1	114
3.2.4	VirB-dependent derepression is conditional on binding site location	114
3.2.5	Removal of the URE abolishes VirB-dependent derepression of <i>proU</i>	117
3.2.6	Modified <i>proU</i> promoter constructs retain osmo-sensitivity	120
3.2.7	URE is a functional transcription regulatory element	120

3.3	Discussion	123
<b>Chapter 4</b>	<b>Auto-regulation and positive feedback are features of the <i>Shigella</i> virulence cascade</b>	
4.1	Introduction	133
4.2	Results	137
4.2.1	Artificial expression of VirB protein at 30°C can alleviate the repressive effects of H-NS at the <i>virB</i> and <i>virF</i> promoters	137
4.2.2	Auto-regulation of <i>virB</i> is dependent on a <i>cis</i> -acting VirB binding site	140
4.2.3	VirB de-repression of <i>virF</i> is facilitated by a <i>cis</i> -acting VirB binding site	143
4.2.4	VirB interaction with the regulatory region of the <i>virB</i> and <i>virF</i> promoters	145
4.2.5	VirB is required for the thermal induction of the <i>virB</i> and <i>virF</i> promoters	148
4.3	Discussion	151
Chapter 5	General Discussion	158
	Bibliography	169

## List of Figures

- Figure 1.1 Infection process of *Shigella flexneri*.
- Figure 1.2 Genetic map of the large virulence plasmid of *S. flexneri*.
- Figure 1.3 VirB is a member of the ParB family of plasmid partitioning protein:
- Figure 1.4 Comparison of the nucleotide sequence of the ParB binding site (*parS*) with the VirB binding site at *icsB*.
- Figure 1.5 Virulence gene cascade in *S. flexneri*
- Figure 1.6 Schematic overview of selected transcriptional regulatory networks in several type II secretion systems
- Figure 1.7 The alignment of the regulatory regions of the *parAB* operon and the *virB* gene and its binding site at the *icsB* promoter.
- Figure 3.1 (A) The alignment of the regulatory regions of *parAB* operon and the *icsB-ippD* intergenic region, (B) An illustration of the regulatory region of *proU*
- Figure 3.2 Creation of a *proU-gfp* fusion. (A) Strategy used for the construction of *proU-gfp*, (B) Expression data confirming *proU-gfp* osmoregulation
- Figure 3.3 Insertion of VirB binding site
- Figure 3.4 Purification and binding affinity of the VirB protein
- Figure 3.5 VirB protein levels are reduced in low salt but are still sufficient to de-repress promoters
- Figure 3.6 De-repression of the *proU* promoter by VirB
- Figure 3.7 (A) Strategy for construction of a single-copy *proU-gfp* fusion on the chromosome. (B) Transcription from derepressed *proU* promoters is mediated by promoter P2, not P1
- Figure 3.8 Partial de-repression of the *proU* promoter by VirB
- Figure 3.9 The URE and VirB-dependent *proU* de-repression
- Figure 3.10 Osmoregulation of modified *proU* promoters
- Figure 3.11 The URE imposes a barrier to transcription read-through
- Figure 3.12.1 VirB-dependent de-repression of the *proU* promoter - model
- Figure 3.12.2 Repression of the *proU* promoter – model
- Figure 4.1 Schematic overview of the interplay of transcription factors in the virulence gene regulatory cascade
- Figure 4.2 Artificial expression of VirB protein at 30°C can alleviate the repressive effects of H-NS at the *virB* and *virF* promoters
- Figure 4.3 Autoregulation of *virB* is dependent on a *cis*-acting VirB binding site
- Figure 4.4 VirB-mediated de-repression of the *virF* promoter if facilitated by a *cis*-acting VirB binding site
- Figure 4.5 VirB requires a *cis*-acting sequence for efficient binding at *virB*
- Figure 4.6 VirB requires a *cis*-acting sequence for efficient binding at *virF*
- Figure 4.7 Thermal induction of the *virB* promoter
- Figure 4.8 Thermal induction of the *virF* promoter
- Figure 4.9 The revised schematic overview of the interplay of transcription factors in the virulence gene regulatory cascade

## List of Tables

Table 2.1	List of strains
Table 2.2	List of plasmids
Table 2.3	List of primers

# **Chapter 1**

## **Introduction**



## 1.1 Overview

*Shigella flexneri* is a Gram negative facultative intracellular pathogen of humans and primates, and is the causative agent of bacillary dysentery. Shigellosis, a disease characterized by the destruction of the colonic epithelium, is responsible for over a million deaths annually (Peng *et al.*, 2010). *Shigella flexneri* contains a large virulence plasmid of approximately 230 kb that encodes most proteins directly involved in the entry and dissemination of bacteria into the epithelial cells, including a type III secretion system (Buchriesser *et al.*, 2001). Several environmental signals are integrated into the regulation of expression of this system, via changes in DNA architecture and the binding of certain nucleoid associated proteins, which act to regulate an extensive repertoire of structural and effector proteins that constitute this elaborate secretion system. The expression of these genes is tightly controlled, with the chromosomally encoded H-NS protein targeting and repressing the majority of the operons, under conditions that are inappropriate for type III secretion assembly. De-repression of these virulence genes, involves a regulatory cascade in which the AraC-like transcriptional activator VirF initiates transcription of *virB*, which encodes a Par-B like dedicated H-NS antagonist VirB. The VirB protein then serves to displace H-NS from a series of structural and effector gene promoters, allowing for full activation of the type III secretion system (Turner and Dorman, 2007).

The additional VirB intermediate step in the regulatory cascade differs from the typical regulatory networks of type III secretion systems in Gram-negative bacteria, whereby the AraC-like transcriptional regulator directly activates the structural genes. The VirB protein is also not a conventional transcription factor, it does not recruit RNA polymerase or initiate open complex formation, it instead acts solely as an H-NS antagonist (Turner and Dorman, 2007). It does not belong to any known family of transcription factors, but is grouped with the ParB family of plasmid partitioning proteins (McKenna *et al*, 2003). Similarities between VirB and the ParB family are evident at the amino acid level with the proteins showing extensive homology. The essential nucleation site to which VirB binds, denoted box1/box2, displays an almost identical nucleotide sequence to the *parS* sites bound by ParB.

The virulence plasmid contains two fully functioning plasmid partitioning systems, and it is apparent that VirB is the vestige of a third. Redundant in terms of the modern virulence plasmid, we hypothesize that genetic rearrangements during evolution allowed for the co-option of the VirB protein into a regulatory function as a H-NS antagonist.

## **1.2 *Shigella***

Pathogenic Gram-negative enteric bacilli such as *Escherichia*, *Salmonella* and *Shigella* species continue to pose a threat to public health (Kotloff *et al.*, 1999, Sansonetti, 1999, Newwell *et al.*, 2010). In contrast to the *Escherichia* and *Salmonella* species, which are estimated to have diverged from a common ancestor around 102 million years

ago, *Shigella* is considered to be a clonal lineage of *Escherichia coli* that arose approximately 80 million years ago (Pupo *et al.*, 1997; Battistuzzi *et al.*, 2004). The DNA sequence divergence between *S. flexneri* and *E. coli* K-12 is about 1.5%. This is marginal compared to a difference of 15% in the case of pathogenic *Salmonella enterica* and *E. coli* (Lan *et al.*, 2002). While *Shigella* and *E. coli* indeed share a high degree of homology, *Shigella* has diverged to occupy specialized niches (Falkow, 1996, Cooper and Lenski, 2000, Lan and Reeves, 2002). The *Shigella* species, of which there are four subgroups *Shigella boydii*, *Shigella dysenteriae*, *Shigella flexneri* and *Shigella sonnei*, are the aetiological agents of bacillary dysentery, an invasive disease of the colon in humans. Bacillary dysentery or Shigellosis, is endemic in developing countries due to poor sanitation often found in these regions. While all four subgroups are pathogenic, the epidemiology of each differs slightly. *S. flexneri* and *S. sonnei* account for the endemic forms of the disease with *S. flexneri* responsible for about 60% of all cases of dysentery in the developing world, while *S. sonnei* causes 77% of cases in industrialized nations (Hale, 1991; Katloff *et al.*, 1999). *S. dysenteriae* is usually the cause of brisk but deadly epidemics of dysentery, particularly in confined populations such as refugee camps (Edwards, 1999).

Bacillary dysentery is spread either directly by the oral-fecal route or by contaminated food, flies or water and is extremely contagious; as few as 10 microorganisms can cause clinical disease, while the LD<sub>50</sub> in healthy adults is about 500 microorganisms (DuPont *et al.*, 1989). The clinical symptoms range from mild diarrhea to severe dysentery, in which an acute inflammatory response in the large intestine gives rise to ulceration of

the bowel and mucoid, bloody diarrhea. Of the several million cases reported each year, approximately 1.1 million are fatal (Galan and Sansonetti, 1996). While usually self-limiting in healthy adults, the disease can be life-threatening in infants, children and the immunocompromised, often resulting in severe sepsis, intestinal occlusion and/or perforation, in addition to dehydration and subsequent chronic malnutrition. A combination of oral rehydration and antibiotics leads to resolution of infection and recovery.

Shigellosis causes 1.1 million deaths with over 164 million annual cases. More than 50 years of research has yielded numerous *Shigella* vaccine candidates that have exemplified the promise of vaccine-induced prevention of shigellosis (Levine *et al.*, 2007). However, although several strategies have been used to develop vaccines, none has been licensed for use outside China (Jennison and Verma, 2004). Owing to the wide range of *Shigella* serotypes and subtypes, there is a need for a multivalent vaccine representing prevalent species and serotypes, a goal that is yet to be obtained (Kweon, 2008). With the ever-increasing frequency of anti-microbial resistant strains it seems likely *Shigella* will continue to be a serious global health problem.

### **1.2.1 Cellular pathogenesis of Shigellosis**

*Shigella* infects the intestinal epithelium leading to the symptoms of dysentery via a multi step process: (i) entry into the host via M-cells, (ii) macrophage apoptosis, (iii) uptake into the epithelial, (iv) intra/intercellular spreading. (Figure 1.1) To gain access

to the intestinal mucosa *S. flexneri* must cross the intestinal epithelium, which evolved as a physical and functional barrier against invasion of commensal and pathogenic bacteria (Sansone, 1994). Unlike other enteric bacteria, such as *E. coli*, *Shigella* is unable to invade the polarized epithelial cell monolayer from the apical pole but instead utilizes the antigen sampling M-cells in the gut to gain entry to the basolateral membrane (Wassef *et al.*, 1989; Menard *et al.*, 1996). M cells are specialized epithelial cells, which continuously sample particles from the gut lumen and deliver them to the underlying mucosal lymphoid tissue, where immune responses can be initiated (Man *et al.*, 2004). *Shigella* triggers its uptake and is passively translocated through the M-cells in a phagosome, and engulfed by macrophage stationed in the basolateral layer below. In contrast to *Salmonella*, which remains and multiplies in the macrophage phagosome, *Shigella* lyses the phagocytic vacuole within minutes of internalization and once free in the cytoplasm, induces apoptosis of the macrophage (Zychlinsky *et al.*, 1992; 1996). The ability to lyse the endocytic vacuole and induce apoptosis depends on a type III secretion system (High *et al.*, 1992, Menard *et al.*, 1993). Macrophage cell death is accompanied by the release of the pro-inflammatory cytokines interleukin-1 $\beta$  (IL-1 $\beta$ ) and IL-18 (Zychlinsky *et al.*, 1994; Sansone *et al.*, 2000). Both cytokines are critical mediators of the acute and massive inflammatory response elicited by *S. flexneri*. Cytokine release results in the recruitment of polymorphonuclear (PMN) cells to the site of infection. Infiltrating PMN destroy the integrity of the epithelial lining, thus enabling more luminal bacteria to reach the submucosa without the need of M cells. This permits more bacteria to gain access to the basolateral membrane and leads to the characteristic symptoms of the disease due to the destruction of the tissue (Perdomo *et*

al, 1994a, b). *Shigella flexneri* can now invade colonic epithelial cells from the basolateral side in a manner similar to that of other enteric pathogens.

Enteric bacteria have evolved several different strategies for facilitating penetration of host cells. *Yersinia pseudotuberculosis* and *Listeria monocytogenes* enter cells by a zipper-like mechanism, in which bacteria establish contact with the cell surface via an interaction between a single bacterial surface ligand and a cellular receptor (Isberg, 1991; Mengaud *et al.*, 1996). In contrast, *Shigella* and *Salmonella* gain entry into the non-phagocytic epithelial cells via major cyto-skeletal rearrangements at the site of bacterial interaction at the cell surface (Finlay *et al.*, 1991; Adam *et al.*, 1995). *Shigella* injects effector proteins that subvert signaling pathways in the host cells leading to the formation of localized membrane protrusions reaching several microns in length to surround the bacterium. These so-call membrane ruffles coalesce around the bacterium leading to its phagocytosis in a membrane-bound vacuole. Inside the epithelial cell, *Shigella* rapidly lyses the phagocytic vacuole, thereby gaining access to the cytoplasm of the cell, where conditions are optimal for growth and multiplication (High *et al.*, 1992; Zychlinsky, 1996). Again, the induction of membrane ruffles on epithelial cells is dependent on the type III secretion machinery and its effector proteins.

*Shigella* is a non-motile organism having lost the ability to synthesize functional flagella (Al Mamun *et al.*, 1997). Comparisons of different *Shigella* strains revealed that the genetic basis of this motility defect varies, indicating that the loss of motility is a result of convergent evolution (Al Manum *et al.*, 1996; Tominaga *et al.*, 2005). Instead, *Shigella* infects neighbouring bacteria in the intestinal mucosa using actin-based

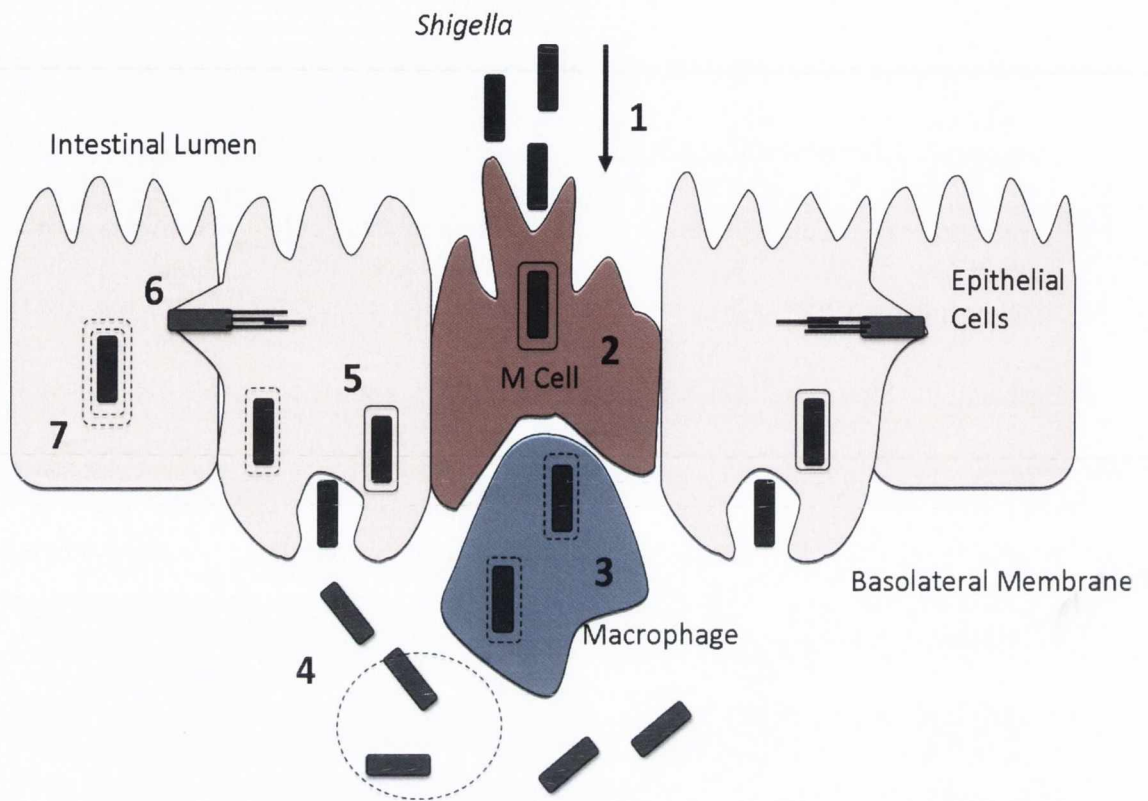


Figure 1.1. Infection process of *Shigella flexneri*

Schematic diagram showing a summary of the process by which *S. flexneri* invades colonic epithelial cells. Bacterial cells move across the intestinal lumen and are taken up by M-cells (1) allowing them to transverse the epithelial layer (2), where they are engulfed by macrophage (3). The *Shigellae* induce apoptosis of the macrophage (4) and can now gain access to the basolateral pole of the intestinal cells. Soon after internalisation the bacteria escape from phagocytic vesicles (5), where in the cytoplasm they can employ actin based motility to facilitate cell to cell spread (6). Within the neighboring cell, the bacteria can then escape from the double membrane (7).

motility. This movement was first observed through microcinematography (Ogawa *et al.*, 1968) when it was observed that intracellular *Shigellae* moved rapidly throughout the cytoplasm of cells. They first express Olm (organelle like movement), a phenotype reflecting intracellular movement along actin stress cables (Vasselon *et al.*, 1991). They subsequently express Ics (intracellular spread), a phenotype by which intracellular bacteria induce nucleation and polymerization of host cell actin filaments. This is followed by accumulation of this material at one end of the bacillus during cell division. This process generates rapid random movement and resulting in the formation of an actin-containing protusions several micrometers in length behind the bacterium which allow passage into adjacent cells (Bernardini *et al.*, 1989; Sansonetti *et al.*, 1991). A combination of these two movements is responsible for the rapid intra/inter cellular spreading of the bacteria.

While Shigellosis is self-limiting, a healthy immune system is a prerequisite for the efficient clearance of *Shigella* infection. It appears that the cytokine, interferon gamma IFN- $\gamma$ , plays an essential role in controlling early stages of infection and susceptibility to the organism (Raqib *et al.*, 1997; Way *et al.*, 1998; Philpott *et al.*, 2000). Although recruitment of polymorphonuclear leucocytes (PMNs) results in the destruction of the mucosal lining, it is ultimately these immune cells, which are immune to apoptosis by the bacteria and contain potential bactericidal properties, that entrap and kill the microbes, therefore resolving the infection. (Weiss *et al.*, 1985; Weinrauch *et al.*, 2002; Brinkmann *et al.*, 2004). Finally, the pattern recognition receptor CD14, appears to



play a role in resolving *Shigella* infections by activating signaling pathways in response to LPS (Aliprantis *et al.*, 2001; Kohler *et al.*, 2002).

### **1.3 Molecular determinants of *Shigella* pathogenicity**

#### **1.3.1 The large virulence plasmid**

In an attempt to study the genetics of invasiveness, early experiments were designed to confer *Shigella* virulence properties on *E. coli* K12 strains. Conjugal transfer of chromosomal genes from *Shigella* Hfr donors to non-virulent *E. coli* strains created avirulent hybrids (Formal and Hornick, 1978). The reason these attempts were unsuccessful was subsequently shown to be due to plasmid-encoded genes playing a key role in *Shigella* virulence. All invasive *S. flexneri* strains, irrespective of serotype, were found to harbour a large plasmid (Sansone *et al.*, 1982). Transfer of the entire *S. flexneri* virulence plasmid to *E. coli* successfully demonstrated its major role in invasion; transconjugants of the normally non-invasive *E. coli* strain harboring the virulence plasmid were able to invade HeLa cells. Although its size and restriction endonuclease digestion pattern vary considerably from one serotype to another, the high molecular mass plasmid is functionally interchangeable between *Shigella* subgroups indicating a high degree of homology (Sansone *et al.*, 1983). Since these original investigations were published, research into the virulence genes, their products and the regulatory mechanisms that control expression have provided detailed insight into *Shigella* virulence. Much information has been acquired from recent virulence

plasmid sequencing projects, which have contributed to our understanding of *Shigella* evolution (Buchrieser *et al.*, 2000; Jiang *et al.*, 2005; Wei *et al.*, 2003). The analysis of DNA sequences showed that the plasmids of serotypes 2a 2457T, serotype 2a 301 and serotype 5 M90T are largely identical, varying slightly in insertion (IS) elements, which make up approximately 25% of the plasmids (Wei *et al.*, 2003). The number and variety of insertion elements seen in *Shigella* virulence plasmids are among the highest that have been described in any bacterial plasmids. The high frequency of IS elements is thought to have contributed to the overall evolution of the plasmid creating a modern plasmid with a mosaic structure.

It appears that this plasmid was originally acquired by horizontal genetic transfer during evolution, as it possesses a very high A+T content (65 – 75%) when compared to the chromosome of the genome which has an average A+T content of 48% (Adler, *et al.*, 1989; Hale, 1991). Phylogenetic studies of *Shigella* and EIEC have shown they are a single pathovar of *E. coli* (Lan *et al.*, 2004). Some studies have shown that the virulence plasmids and the chromosomes share a similar evolutionary history (Escobar-Paramo *et al.* 2003), which suggests that an ancestral plasmid entered an *E. coli* strain only once, which gave rise to the *Shigella*/EIEC lineage. However, most evidence obtained of late from comparative genomic studies clearly indicates that *Shigella* and EIEC evolved from multiple *E. coli* strains by convergent evolution, due to the transfer of a variety of diverse ancestral plasmids to *E. coli* on several occasions, producing the modern *Shigella* species (Yang *et al.* 2007).

### 1.3.2 The Type III secretion system and its effector proteins

At the core of the *S. flexneri* virulence machinery is the Mxi-Spa type III secretion system and its associated effector proteins. Type III secretion systems are found in more than 25 bacterial species interacting with animal or plant hosts (Abe *et al.*, 2005). This Sec-independent secretion system enables Gram-negative bacteria to secrete and inject pathogenicity proteins into the cytosol of eukaryotic host cells. Similar systems are found in other enteric pathogens including *Salmonella*, which possess two systems located on two pathogenicity islands, and is involved in host-cell entry, and *Yersinia* where it is involved in the secretion of anti-host factors (Schlumberger *et al.*, 2006; Matsumoto and Young, 2009). On the *Shigella* virulence plasmid, the genes encoding the Mxi-Spa system and its effector proteins are grouped within a 31-kbp portion called the 'entry region' (Figure 1.2). Here, there are 37 open reading frames organized into two clusters, transcribed in opposite directions which facilitate coordinated expression. Based on their function they can be classified into four groups, type III secretion system apparatus, effectors/translocators, chaperones and transcriptional regulators.

The Mxi-Spa system is encoded by the *mxiGHIJKLMADCA* and the *spa15, 47, 13, 32, 33, 24, 9, 29, 40* genes. This locus encodes the components needed for assembly and function of the type III secretion system. Together with the Ipa invasion proteins, the apparatus is responsible for the secretion of 25 effector proteins, found at different

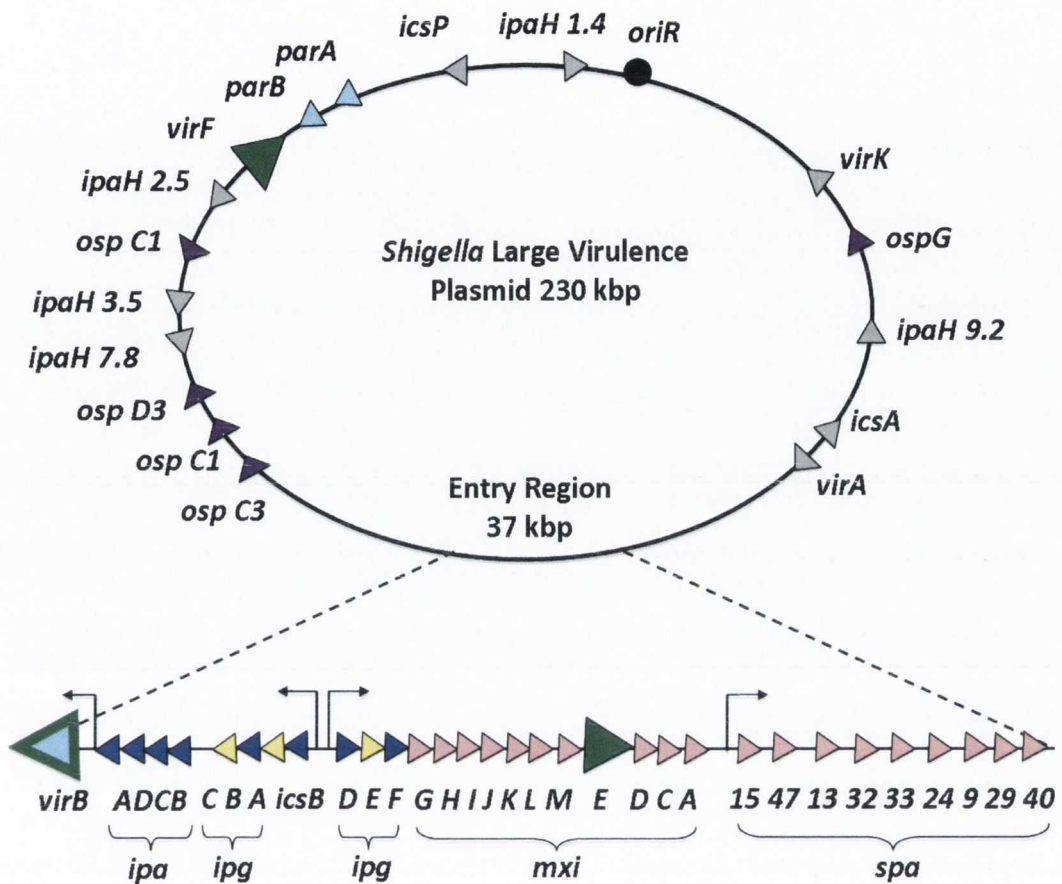


Figure 1.2. Genetic map of the large virulence plasmid of *S. flexneri*, showing the positions of some of the regulatory and structural genes (not all genes shown). Individual genes are represented by coloured arrow-heads and the origin of replication by a black filled circle at *oriR*. The 37 kbp entry region is enlarged to show position and organization of the virulence gene operons. Angled arrows represent promoters driving individual operons. Arrow-heads in red show the localization of the genes encoding proteins of the type III secretion system machinery, dark blue indicate effector/translocator proteins, yellow represent chaperone proteins, pale blue show protein involved in plasmid partitioning and green and enlarged indicate regulatory proteins. The *virB* gene is shown here as a blue and green arrow-head as it is a ParB-like protein with regulatory functions.

location on the virulence plasmid, from the bacterial cytoplasm into the host cell (Figure 1.2).

The four Ipa (invasion plasmid antigen) proteins, IpaB (580 aa), IpaC (382 aa), IpaD (332 aa) and IpaA (633 aa) are encoded in the *ipgC-ipaBCDA* operon located at the left end of the 31-kb invasion gene cluster (Figure 1.2). The most prominent proteins secreted by *S. typhimurium* are homologous to the *Shigella* IpaA through IpaD proteins and accordingly have been named *Salmonella* invasion proteins SipA through SipD (Collazo and Galan, 1997; Osiecki *et al.*, 2001). The Ipa polypeptides are required for host cell entry in the infection process of *S. flexneri*, and are the dominant antigens producing a humoral response during *Shigella* infection (Hale *et al.*, 1985). These proteins are needed for controlled secretion and translocation of effector proteins, and are expressed alongside the T3SS components and stored in the cytoplasm. The IpgC cytoplasmic chaperone serves to stabilize the IpaB and IpaC and the IpgE chaperone stabilizes IpgD. This helps to avoid premature association of the proteins. Once outside the bacterium, IpaB and IpaC form a soluble complex as shown by coimmunoprecipitation of the two proteins with antibodies against either one of them (Menard *et al.*, 1994). This complex interacts with the epithelial cell membrane to form a pore, through which other Ipa proteins are thought to enter the cytoplasm. The rate of protein release is believed to be regulated by a complex formed by IpaB and IpaD (Menard *et al.*, 1994). The role of these proteins is not confined to invasion: The cytoskeleton rearrangements which cause the bacterium to be engulfed are induced by IpaA and IpaC (Tran Van Nhieu *et al.*, 1997, 1999). IpaB and IpaD are also required

for lysis of the phagocytic membrane and IpaB induces apoptosis in infected macrophages (De Gayter *et al.*, 2000; Zychlinsky *et al.*, 1993).

There are many genes outside the entry region that contribute to the virulence phenotype. After liberation from the vacuole the bacterium acquires motility by recruiting and polymerizing host actin through a process that depends on IcsA, an outer membrane protein encoded by the virulence plasmid. Initially IcsA is distributed evenly around the outside of the bacterial cell, however the protein is cleaved by IcsP, an outer membrane protease, leading to the localization of IcsA at the pole of the cell (Egile, 1997). IcsA then binds N-WASP, a host cytoskeletal protein, which catalyzes the polymerization of the actin leading to forward momentum and the formation of an actin tail (Suzuki and Sasakawa, 2001). Located upstream of *icsA* and transcribed from the opposite DNA strand is the *virA* gene. It codes for the VirA protein, which is also required for spreading and promotes efficient internalization of *Shigella* in host cells through a mechanism that involves destabilisation of microtubules with associated membrane ruffling. The *virK* gene has been described as contributing to the correct positioning of IcsA (Wing *et al.*, 2005).

Scattered around the plasmid are five alleles of the *ipaH* locus (Figure 1.2) which code for proteins that are secreted when the bacterium is in an intracellular niche (Jin *et al.*, 2002). Also dispersed on the plasmid are the genes encoding the proteins OspC1, OspB, OspD1, OspG and OspE1 (Buchrieser *et al.*, 2000). The Mxi-Spa-secreted effector IpaH 9.8 functions as an E3 ubiquitin ligase and inhibits MAPK kinase-

dependent signaling by targeting components of this pathway for degradation by the proteasome (Rhode *et al.*, 2007). Interestingly, the chromosome-encoded IpaH family protein members reduce *Shigella*-induced inflammation in a mouse pulmonary infection model. Several IpaH proteins might be functionally redundant, as only a *S. flexneri* strain deficient in all chromosomal *ipaH* genes displayed a detectable phenotype (Ashida *et al.*, 2007) OspG is another Mxi-Spa-secreted effector protein, which blocks NF- $\kappa$ B-dependent immune signalling, thus dampening intestinal inflammation (Kim *et al.*, 2005). Several members of the outer Shigella protein (Osp) family shape the transcriptional response of epithelial cells invaded by *S. flexneri*, thus modulating the immune response in favour of the bacteria (Buchrieser *et al.*, 2000; Kim *et al.*, 2005). A role in the modulation of the immune response was also established for one member of the multigene IpaH family (Okuda *et al.*, 2005).

In addition to the fundamental bacterial weaponry, the *Shigella* virulence plasmid contains transcriptional activators that coordinate the expression of the type III secretion system and its effectors. These include the proteins VirF, VirB and MxiE. The expression of this elaborate system is likely to impart a heavy metabolic burden on the bacterial cell which could compromise its survival if expressed in an inappropriate environment. In addition these proteins are highly antigenic and would be likely to promote an immune response in the host if expressed constitutively. Therefore, *S. flexneri* has evolved a complex regulatory system which involves both the activator proteins VirF, VirB and MxiE along with chromosomally encoded proteins to ensure expression of the structural and effector genes is tightly controlled.

### 1.3.3 Regulatory proteins

#### 1.3.3.1 VirF

The VirF protein is a member of the AraC family of transcriptional regulator proteins (Porter and Dorman, 2002). The AraC family contains over 100 members, identified by virtue of a conserved C terminal. Most of these proteins are transcriptional activators, with the exception of CelD, and can be divided into 3 subgroups based on the signals to which they respond (Gallegos *et al.*, 1997, Martin and Rosner, 2001). The first group of proteins regulates transcription upon binding a chemical signal, usually a carbohydrate, and is typified by the AraC protein, the prototype for the family. AraC regulates the *ara* operon in several Gram-negative bacteria in response to arabinose (Schleif, 1996). The protein itself is present in dimeric form in solution and comprises an amino terminal domain to which arabinose binds controlling dimerization of the protein, and a carboxy terminal domain that contains two potential helix turn helix motifs involved in sequence specific DNA recognition and binding (Soisson *et al.*, 1997a,b, Porter and Dorman, 2002). The two, functionally distinct domains within the protein are connected by a linker region (Eustance and Schleif, 1996). The second subgroup of proteins in the family, are involved in the stress response. These proteins, including MarA, Rob and SoxS are monomers and most notably lack the N-terminal domain of the previous subgroup. The third subgroup regulates transcription in response to physical stimuli. This group includes many regulators of virulence gene



expression including Rns, a regulator of CS1 pili expression in *E. coli*, PerA a regulator of *esp* invasion expression in enteropathogenic EPEC and VirF protein of *S. flexneri* (Caron *et al.*, 1989; Gomez-Duarte and Kaper, 1995; Porter *et al.*, 1998; Porter and Dorman, 2002). VirF shows a particularly close relationship with Rns (36% amino acid homology), and Rns can functionally replace VirF in activation of structural gene expression (Porter *et al.*, 1998). In general a high degree of homology is observed between groups of the AraC family at the carboxy terminal DNA binding domain while the N-terminal effector domains tend to be more divergent (Munson and Scott, 1999). The VirF protein like other members of the family contains two HTH DNA binding motifs within its C-terminal domain, which have been shown to be essential for DNA binding. Mutants that were dominant negative when co-expressed with the wild-type VirF protein were isolated, indicating a role for protein-protein oligomerization in normal VirF function (Porter and Dorman, 2002).

VirF resides at the top of a regulatory cascade, where it acts as a transcriptional activator of *virB*, an H-NS antagonist required for expression of all other structural genes (discussed in detail in 1.8.3). It also acts at the *icsA* promoter, a structural gene required for intercellular spreading of bacteria in the host, along with the VirB protein (Adler *et al.*, 1989; Le Gall *et al.*, 2005).

### 1.3.3.2 VirB

The VirB (called InvE in the *S. sonnei* literature) protein was first identified through transposon mutagenesis of the virulence plasmid (Adler *et al.*, 1989) when it was found to be absolutely required for activation of almost all the structural virulence genes. It is a small (35.4 kDa), basic protein that shows no homology to any known family of conventional transcriptional activators (Watanabe *et al.*, 1990; Dorman and Porter, 1998). Instead, the VirB protein resembles the ParB/SopB proteins involved in the partitioning of plasmids and the maintenance of stable copy number on the P1/P7 and F plasmids respectively (Abeles *et al.*, 1985; Watanabe *et al.*, 1990; Radnedge *et al.*, 1996) (Figure 1.3). The homology is most pronounced in the amino-terminal two-thirds of the proteins that include a HTH DNA-binding motif that displays 80% identity. The ability of VirB to bind DNA depends on the integrity of this HTH region, located between residues 148 and 171, and binding is essential for gene activation (Beloin *et al.*, 2002; McKenna *et al.*, 2003). The VirB and ParB protein sequences become more divergent at the C-terminal ends that facilitate oligomerisation. VirB can form dimer, trimers and higher-order oligomers *in vivo* and *in vitro* (Watanabe *et al.*, 1990; McKenna *et al.*, 2003). The ability to form oligomers is critical for VirB function (McKenna *et al.*, 2003). It occurs through two domains. A leucine zipper located between residues 193 and 228 promotes the initial protein-protein interaction that is required for dimer formation; deletion of the leucine zipper or amino acid substitutions within it result in loss of oligomerisation and gene activation. A region at the C terminus that is predicted to form a triple coil structure, contributes to formation of higher order oligomers

**A**

```

VB -----MVLKNDLISIKE-----GQKKE--FTIHSNKVSTKA 32
PB -----MSKKNRPTIGRTINPSILSGFDSSA---SDRVQVVKLSTGRQATLEE 48
SB MKRAPVIPKHTLNTQPVEDTSLSTPAAPMVDLSLIARVGMARGNAITLFPVGRDVRITLE 60

VB KIPHKRIQDLTEVFNKTIIVRDOESLTEESADIKIKKLCOFFEVIKREIDGRTEILD 90
PB VIFPNQVESDLEVDQHNCRDQASLTFKSLKSRSLKSGQFYPAIGVRRATKTEILD 107
SB VLRGDSVEKTSRVWSG--NERDOELLTEDALDLPSEFLTGOQTPAFC--SRVSVTEIAD 118

VB GTRRRASATYAGADDEVVYSKEYIS--TLDARKLANDIOTAKEHSIRSLSIGENFIRVSGM 149
PB GSRRRASATLENVGRVIVTDQELIS-VQEAQNLAKVQVOTALQHSIREISLRMRMINDGM 166
SB GSRRRKKAALTESDYRVIIVGELDDEQMAALSRLGNVYRPTSAYERGORYASRLQNEFAGN 178

VB SYKDIAKKKNLDRAKVTRALQAAASVPCFIISLFFIASEINENDYKILFNYYKGLKAKNES 209
PB SQKDIAAKGLSCAKVTRALQAAASPEFLVALFVQSELTESDYKTLCAVGDEMGNKMLE 226
SB -ISALADANISRKIITNCINTAKLPKSVVALFSPGELIARSQDALQKAFTDK---L 234

VB *SSTLPIIKKEETKOLD--TNLPPFIYKKEITNIIKSKNRKQNPSLK---VDSIFISK 263
PB FDQLIQNISPEINQILSIEEMAEDEVNKKIIRLITKEASLLTDKGSQKQSVTEWKFEB 286
SB IKQQASNHEQKAGVIFEAEVITLLTSVLEKTSASRTSLSSRRHQFAPG--ATVLYKG 292

VB KIRTYIKRKENKTNRTLFTLSKINKTVQREIDEAR---DIIIRHLSSE----- 309
PB KIRFARKKVKGRAFSYEENRLSKRVGTGRTRQDCCGSLERASIKRSLKLPFKFHY 333
SB K--MVLNLDERSVPTECEKTEAELKELEKPA----- 323

```

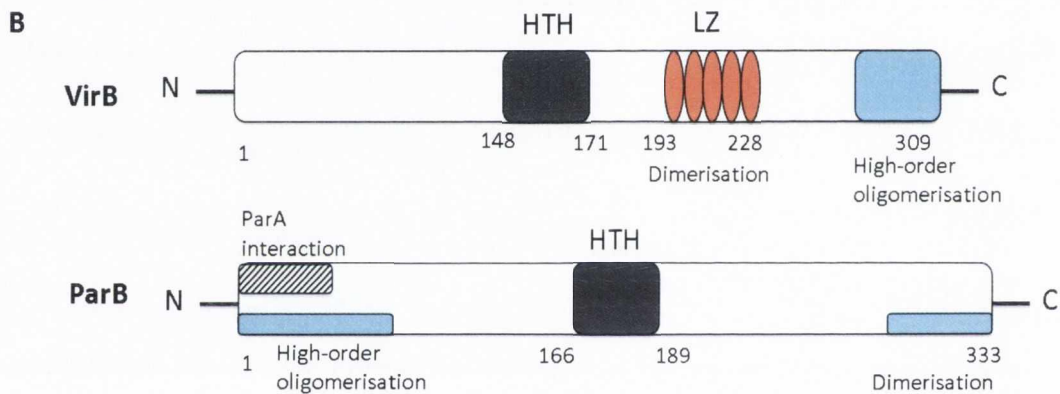


Figure 1.3. VirB is a member of the ParB family of plasmid partitioning proteins. (A) The amino acid sequences of the VirB protein (VB), the ParB protein (PB) and the SopB protein (SB) aligned to show homology (blackened letters) present between the three proteins (Beloin *et al.*, 2002).

(B) Schematic diagram comparing the functional domains of ParB and VirB proteins. Amino acid similarities in the helix turn helix motif (HTH) are represented by grey boxes and the oligomerisation domains are shown here with blue boxes. VirB protein also contains a leucine zipper (LZ) needed for dimerisation (orange ovals).

(Beloin *et al.*, 2002). Oligomerisation of the protein also contributes to DNA binding since oligomerisation mutants bind to DNA with a lower affinity *in vivo* (Belion *et al.*, 2002) (Figure 1.3).

The DNA sequence to which VirB binds at the *ipa* operon is an inverted repeat denoted box 1 and box 2, 5' GTTTCATcATGAAT 3' (Taniya *et al.*, 2003; Turner and Dorman., 2007) Site-directed mutagenesis revealed the box 2 heptameric motif to be essential for DNA binding (Taniya *et al.*, 2003; Turner and Dorman. 2007; Castellanos *et al.*, 2009) Matches of this sequence can be found in the regulatory regions of other proteins regulated by VirB such as *icsB*, *virA* and *spa15* (Taniya *et al.*, 2003).

Interestingly, this inverted repeat bound by VirB exhibits a highly similar nucleotide sequence to the box A box B site bound by ParB on the P1/P7 plasmid. ParB binds to a region known as *parS*, which contains 4 heptameric and two hexameric imperfect repeats, each heptamer is an individual binding site for ParB (Hayes and Austin, 1993). The intergenic regulatory region between *icsB-ipgD* shows DNA sequence homology, not only to the Box A/Box B repeats, but the entire *parS* region. This extended DNA sequence homology is not found at other VirB dependent promoters.

VirB activates the *virA*, *icsB*, *ipgD* and *spa* promoters, all of which are genes repressed by the H-NS nucleoid-associated protein (Beloin and Dorman, 2003). A detailed molecular dissection of the interplay between VirB and H-NS at the *icsB* promoter identified an antagonistic mechanism of transcription activation previously uncharacterized for VirB (discussed in 1.8.4).

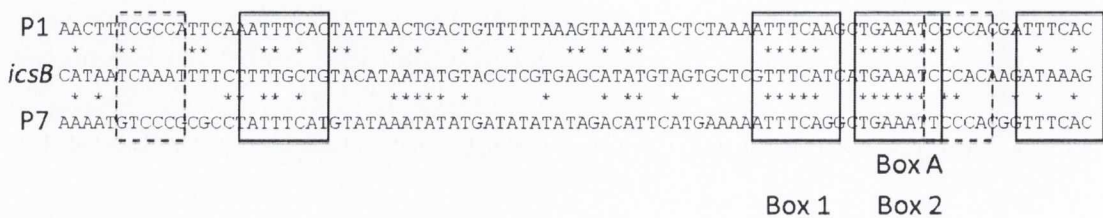


Figure 1.4. Comparison of the nucleotide sequence of the ParB binding site (*parS*) with the VirB binding site at *icsB* shows a highly similar nucleation site sequence (box1/box2) that extends throughout the *parS/icsB* region. The nucleotide sequences of *parS* from the P1 plasmid, the P7 plasmid and the region to which VirB binds upstream of the *icsB* promoter are aligned. Asterisks between DNA sequences denote matching base pairs. The essential box1/box2 VirB binding site is shown, along with the boxA (rectangle boxes) and box B (dashed rectangle boxes) sites recognized by ParB are also shown.

### 1.3.3.2. MxiE

The *mxiE* gene is located within the *mxi* operon in the entry region of the virulence plasmid (Figure 1.2). The MxiE protein, like VirF, is a member of the AraC-like family of transcriptional regulators (Kane *et al.*, 2002). The virulence phenotype of a nonpolar *mxiE* mutant was characterized and found to retain the ability to invade mammalian cells in tissue culture and to secrete Ipa (type III effectors required for host cell invasion), although it was less efficient than wild-type *Shigella* at cell-to-cell spread (Kane *et al.*, 2002). Despite its invasive properties in culture, the *mxiE* mutant was completely avirulent in an animal model.

The MxiE protein plays an important role in controlling the expression of a subset of genes in response to protein secretion. Demers *et al* (1998), identified six virulence plasmid genes that are activated by MxiE in the intracellular environment: 5 *ipaH* genes and *virA*. Subsequently, Kane *et al* (2002) revealed 4 *osp* genes under the control of MxiE. The presence of eight MxiE boxes on the virulence plasmid suggests that 11 genes encoding secreted proteins may be regulated by the activity of secretion (Mavris *et al.*, 2002b). Microarray analysis in 2005 comparing the transcriptional profiles of wild type and *mxiE* mutant strains indicated the expression of 13 genes (*ipaH*, *ipaH4*, *ipaH7*, *ipaH9*, *ospB*, *ospC1*, *osp D3*, *osp E1*, *osp E2*, *ospF*, *ospG*, *virA* and *phoN2*) were activated in an MxiE-dependent manner (Le Gall *et al.*, 2005). MxiE is only active when it forms a complex with the IpgC cofactor protein. IpgC is a chaperone that associates independently with the IpaB and IpaC effector proteins rendering the

chaperone protein not always available for MxiE activation. The free form of IpgC becomes available when IpaB and IpaC secretion begins and this is a signal for transcription activation by MxiE of its regulon of secreted protein genes (Mavris *et al* 2002a). Promoters activated in this way possess a so-called MxiE box with the consensus sequence 5'-GTATCGTTTTTTTAnAG-3'. This sequence is located between positions -33 and -49 with respect to the known or presumed transcription start sites of eight promoters (Marvis *et al.* 2002b). In this location the box overlaps the -35 box of the promoter and it is presumed that this promotes interaction between MxiE and RNA polymerase at these promoters. A subset of these MxiE regulated genes, *ospB*, *ospF*, *ospC1* and *virA* are also under the control of the VirB protein. At the *virA* promoter both MxiE and VirB have been shown to bind and activate transcription. The significance of this dual control is unknown at present but is thought to involve differential activation during certain periods of the invasion process (Le Gall *et al.*, 2005).

#### **1.4 Chromosomal proteins involved in regulating virulence gene expression.**

##### **1.4.1 H-NS**

The histone-like nucleoid-structuring protein, H-NS, is a pleiotropic repressor of transcription in Gram-negative bacteria (Dorman, 2004b). It is a small, basic DNA binding protein of approximately 15 kDa in size. The protein is composed of an amino terminal oligomerisation domain that is attached to a carboxyl-terminal nucleic acid

binding domain by a flexible linker peptide (Badaut *et al*, 2002, Dorman *et al* 1999). H-NS forms at least dimers in solution and these have the ability to form DNA-protein-DNA bridges both between separate DNA molecules and between different portions of the same DNA molecule (Dame *et al* 2005, 2006; Dorman and Kane, 2009). The resulting micro-loop domain formed can act to trap RNA polymerase or exclude RNA polymerase thereby silencing transcription.

Much effort has been invested in understanding the DNA binding preference of H-NS. The sequences to which it binds are often associated with regions of intrinsic curvature and are AT rich. In addition, a recent study has identified a discrete DNA sequence 5'-TCGATATATT-3' to which H-NS binds with higher affinity than other AT rich elements. This motif acts as a nucleation site from which H-NS can spread laterally along DNA forming H-NS filaments or bridges (Lang *et al.*, 2007, Rimsky *et al.*, 2001).

Genes acquired by horizontal transfer therefore become an obvious target for repression due to their unusually high A+T content. An attractive hypothesis envisages a role for H-NS in the silencing of these genes, allowing them to be incorporated into the host genome in a largely inert state, in order to limit the cost to competitive fitness (Doyle *et al.*, 2007; Lucchini *et al.*, 2006). However, if the cell is to benefit from newly acquired genetic information and expand its functional repertoire, H-NS mediated repression must be relieved either immediately or following the evolution of an appropriate and specific anti-H-NS mechanism (Navarre *et al.*, 2006).



The removal of H-NS from its target promoters involves influences acting at a number of levels (Stoebel *et al.*, 2008). The dependence of H-NS on the maintenance of DNA curvature together with a relatively degenerative DNA binding site makes it vulnerable to changes in local DNA structure (Prosseda *et al.*, 2004). Host-associated signals that alter DNA topology such as increases in temperature or osmolarity can abrogate DNA curvature, disrupting the DNA-H-NS complex. In addition, other proteins interact with H-NS via its oligomerisation domain to alter activity (Muller *et al.*, 2006). These include its many paralogues, StpA, Hha and Sfh. Finally, transcription factors can cause the displacement of H-NS, either by competitive binding or structural remodelling of the local DNA architecture to obstruct H-NS. Proteins that oppose H-NS stem from a wide variety of transcription factor families such as AraC-like and LysR-like subgroups, suggesting antagonism is not likely to utilize highly specific protein-protein interaction between repressor and anti-repressor (Beloin *et al.*, 2001; Stoebel *et al.*, 2008).

The gene coding for H-NS in *S. flexneri* was originally designated *virR* and transposon Tn10 knockout mutation in *virR* was found to de-repress virulence gene expression at 30°C (Maurelli and Sansonetti, 1988). The *virR* gene was subsequently shown to be allelic with the pleiotropic *osmZ* locus of *Escherichia coli*, and subsequent analysis revealed that both genes were identical with *hns* (Dorman *et al.*, 1990; Hulton *et al.*, 1990). DNase I protection studies have shown that H-NS binds to sites that overlap the *icsB* and *virA* promoters, as well as their activators *virB* and *virF* (Beloin *et al.*, 2002).

### 1.4.2 IHF

The nucleoid associated protein, Integration Host Factor (IHF) is a heterodimer consisting of an  $\alpha$  and a  $\beta$  subunit, encoded by the *ihfA* and *ihfB* genes, respectively (Freundlich *et al.*, 1992; Goosen *et al.*, 1995). Upon interaction with its target DNA, IHF introduces a sharp bend of more than  $160^\circ$  in the DNA (Rice *et al.*, 1996). A single IHF heterodimer has been shown to bind a region of approximately 35-bp, which contains a 13-bp consensus sequence. IHF is involved in a number of processes including chromosome and plasmid replication, transposition, site-specific recombination and transcriptional regulation (Freundlich *et al.*, 1992). A role for IHF in the positive regulation of *S. flexneri* virulence gene expression was shown, in that it is required for continued expression into stationary phase (Porter and Dorman, 1997a). For the *virF* and *icsA* genes, IHF also contributes to expression during the logarithmic phase, in contrast to *virB* and *icsB*, which are stimulated only in stationary phase, suggesting that the mechanism of regulation differs at the individual promoters (Porter and Dorman, 1997a). Direct binding of IHF to these promoter regulatory regions has been shown along with putative matches to the IHF consensus sequence. It is possible IHF aids in the antagonism of H-NS at these promoters, or in the case of *virB* and *icsB*, IHF may act to constrain negative supercoils at the promoter leading to more favorable interactions of RNA polymerase with VirF (Tobe *et al.*, 1995; Porter and Dorman, 1997).

### 1.4.3 FIS

FIS is a small homodimeric DNA binding protein which in *E. coli* participates in fundamental cellular processes such as chromosome replication (Filutowicz *et al.*, 1992), site-specific recombination (Finkel and Johnson, 1992) and DNA transposition (Weinreich and Reznikoff, 1992). FIS is involved in the regulation of a large set of genes by direct control of transcription initiation as well as by possible indirect effects on some RpoS-regulated genes (Xu and Johnson, 1995; Gonzalez-Gil *et al.*, 1996). In a process known as growth phase-dependent regulation, the intracellular FIS levels peak during early exponential growth at over 50,000 molecules per cell, and thereafter decrease until they become very low during stationary phase (Ball *et al.*, 1992).

FIS is involved in the regulation of a large set of genes by direct control of transcription initiation as well as by possible indirect effects on some RpoS-regulated genes (Xu and Johnson, 1995; Gonzalez-Gil *et al.*, 1996). In the case of the *rrnB* P1 promoter, FIS resembles a class I transcription activator that interacts with the C-terminal domain of the  $\alpha$  subunit ( $\alpha$ CTD) of RNA polymerase from position -71 (Zhi *et al.*, 2003). In the case of the RpoS-dependent *proP* P2 promoter, FIS acts more like a class II transcription activator that contacts the  $\alpha$ CTD of RNA polymerase from a binding site centered at -41 (McLeod *et al.*, 1999, 2002). In the *leuV* promoter, stimulation occurs in part by a FIS-mediated translocation of superhelical energy from upstream binding sites to the promoter region (Opel *et al.*, 2004). In the case of the *virF* promoter of

*Shigella* FIS has been shown to bind four specific sites in the promoter region and to exercise a direct and positive effect on transcription (Falconi *et al.*, 2001).

## 1.5 Two-component regulatory systems

Two-component systems are used to sense and respond to extracytoplasmic conditions. The systems are generally composed of a sensor kinase and a response regulator, communicating via phosphotransfer. Several two-component regulatory systems have an input into the regulation of gene expression in *S. flexneri*. The EnvZ/OmpR signal transduction system regulates expression of porin genes, *ompF* and *ompC*, in response to changes in osmolarity (Buckler *et al.*, 2000). In addition to these porin genes, OmpR regulates expression of many genes in *E. coli*. Disruption of the EnvZ/OmpR system in *S. flexneri* reduces its virulence and decreases expression of the virulence genes (Bernardini *et al.*, 1990). But the effect of the EnvZ/OmpR system seems to be indirect. Virulence of the *envZ/ompR* mutant is restored by multiple copies of *ompC*, a gene that is barely expressed in the mutant (Bernardini *et al.*, 1993). These mutants are impaired in their ability to spread from cell to cell and to kill epithelial cells. These results suggest that the main contribution of the EnvZ/OmpR system to virulence of *Shigella* spp. is through the unknown role of OmpC.

The CpxR-CpxA two-component system, encoded by the chromosomal genes *cpxRA*, mediates the response of the *virF* promoter to alterations in pH. Nakayama and Watanabe (1995) isolated an *E. coli* mutant that exhibited an upregulated  $\beta$ -

galactosidase activity from a *virF-lacZ* fusion plasmid at pH 6.0, with substantial activity at pH 7.4. This mutant had a Tn10 insertion in the *cpxA* gene, the sensor of the Cpx two-component system. A gene homologous to *cpxA* was conserved in *Shigella* spp. as well as in *E. coli*. Subsequently, a cognate regulator, CpxR, was identified as an essential factor for *virF* expression (Nakayama and Watanabe, 1998). Purified CpxR protein binds to the *virF* promoter between positions -37 and -100 in relation to the transcription start site, and its binding affinity is enhanced when the CpxR protein is phosphorylated. Transcription of *virF* is activated *in vitro* by phosphorylated CpxR and it is thought that under acidic conditions CpxA may act as a phosphatase to dephosphorylate CpxR, preventing it from binding to *virF* (Nakayama and Watanabe 1995, 1998).

The deletion of the *cpxA* gene also influences TTSS expression at the level of posttranscriptional processing of InvE (VirB) (Mitobe *et al.*, 2005). It has been shown that activity levels from an *invE-lacZ* transcriptional fusion in a *cpxA* mutant remains similar to those of wild type, whereas the beta-galactosidase activity levels from the translational fusion of *invE-lacZ* is reduced to 21% of that of wild type (Mitobe *et al.*, 2005).

The PhoQ/PhoP system is known to regulate transcription of *Salmonella* virulence genes (Kato *et al.*, 2008). It is composed of the PhoQ sensor kinase and the PhoP DNA-binding protein. Although a *phoP* mutant of *S. flexneri* has the same ability of cell invasion, intercellular spreading, induction of apoptosis in macrophages and resistance

to extreme acid pH, the mutant *S. flexneri* are cleared more rapidly than wild type from infected animals (Moss *et al.*, 2000). Although precise mechanisms remain unknown, the PhoQ/PhoP system is responsible for resistance to killing by polymorphonuclear leucocytes (PMN) and antimicrobial peptide (Moss *et al.*, 2000).

## 1.6 Lux quorum sensing

Quorum sensing is the regulation of gene expression in response to fluctuations in small molecules that report cell-population density (Whitehead *et al.*, 2001). Quorum-sensing bacteria produce and release chemical signal molecules called autoinducers that increase in concentration as a function of cell density. Studies of *Shigella* virulence gene expression have demonstrated that maximal expression of genes encoding the type III secretion system and its substrates and maximal activity of this virulence organelle occur at high cell density (Day and Maurelli., 2001). It was demonstrated that the expression of *ipa*, *mxi*, and *spa* invasion operons is maximal in stationary-phase bacteria and that conditioned media derived from stationary-phase cultures enhance the expression of these loci. In contrast, expression of *virB* peaks in late log phase; accordingly, *virB* expression is enhanced by a signal(s) present in conditioned media derived from late-log-phase cultures (Day and Maurelli, 2001). Autoinducer 2 (AI-2), a quorum signaling molecule active in late log phase, is synthesized by *Shigella* species and enteroinvasive *E. coli* and shown to be responsible for the observed peak of *virB* expression. However, AI-2 does not influence invasion operon expression and is not

required for *Shigella* virulence, as mutants deficient in AI-2 synthesis are fully virulent (Day and Maurelli, 2001).

## 1.7 Transcriptional regulation in prokaryotes

The ability to modulate gene expression patterns in response to environmental signals is essential to the survival of the bacterium. Initiation of transcription is the first level of gene expression and is a key step regulating the transfer of information from DNA to proteins required for all cellular functions in bacteria. Much of what we know about transcriptional regulation is the result of extensive studies of gene expression in the Gram-negative bacterium *Escherichia coli* (Haugen *et al.* 2008). For most bacterial promoters, transcription initiation involves the recognition of the promoter consensus sequence by RNA polymerase, binding of RNA polymerase holoenzyme to form a closed transcription initiation complex followed by “melting” of the DNA to form an open complex. This is followed by the elongation of the transcript and its eventual termination.

Five protein subunits,  $\beta\beta'\alpha_2\omega$ , comprise the core RNA polymerase enzyme which transcribes the DNA code into an RNA message. The  $\beta$  and  $\beta'$  subunits form the catalytic centre of this multi-subunit structure which has determinants for the binding of both template DNA and the RNA product during transcription (Korzheva *et al.*, 2000). The two identical  $\alpha$  subunits consist of two domains.  $\alpha$  NTD (amino-terminal domain) is responsible for the assembly of the  $\beta$  and  $\beta'$  subunits, and  $\alpha$  CTD (carboxy-

terminal domain) is involved in DNA-binding and interaction with transcription factors at certain promoters (Blatter *et al.*, 1994). The  $\omega$  subunit appears to function as a  $\beta'$  chaperone assisting in the large subunit's proper folding but is not required for successful transcription. The crucial step in transcription initiation is the docking of the core enzyme to the double-stranded DNA. This process requires the addition of a sixth subunit,  $\sigma$ , which confers the binding specificity of the polymerase to a region of DNA upstream of an gene, designated the promoter. *E. coli* contains seven of these  $\sigma$  factors and each of them can be used for recognition of specific promoters, although  $\sigma^{70}$  is considered to be the sigma factor controlling expression of most housekeeping genes (Ishihama, 2000). Together, the core enzyme and the  $\sigma$  subunit constitute the holoenzyme, a structure which is both necessary and sufficient for transcription initiation. Holoenzyme recruitment to the promoter of a gene is a competitive process because promoters outnumber the population of available RNA polymerases within the cell. The strength of a promoter and overall ability of a promoter to attract available RNA polymerases determines whether or not the gene is expressed.

Many of the elements that facilitate RNA polymerase binding to a promoter are already present within the promoter DNA itself. The two primary promoter elements involved in recognition and binding by the  $\sigma$  subunit and subsequently, the polymerase, are the -10 or Pribnow box (5'-TATAAT-3' is the consensus for  $\sigma^{70}$ ) and -35 (5'-TTGACA-3' is the consensus for  $\sigma^{70}$ ) hexamers, which are named according to their proximity to the transcription start site. These specific sequences are recognized and bound by domains within the  $\sigma$  subunit. Strong promoters tend to contain sequences close to this ideal



consensus, while weak promoters may have base changes in this region or may differ from the optimum spacing of 17 bp between the two motifs.

While the -10 and -35 consensus sequences are sufficient for  $\sigma$  subunit recruitment, other sequences can enhance RNA polymerase binding and therefore, transcription initiation. The extended -10 element, which consists of 3-4 bp (5'-TGTG-3' consensus) immediately upstream of the -10 element, is also recognized by a domain of the  $\sigma^{70}$  subunit (Murakami *et al.*, 2002). Additionally, AT rich sequences called UP elements consisting of approximately 20 bp (5'-AAAWWTWTTTTNNNAAANN-3' consensus; W = A or T and N = any base) and located upstream of the -35 element assist RNA polymerase binding through recognition by the C-terminal domains of the  $\alpha$  subunits. These elements can increase transcription 30-90 fold and are usually found at highly active promoter such as the promoters that drive transcription of the ribosomal RNA operons (Ross *et al.*, 2001).

Complete control over polymerase recruitment is not just restricted to these promoter proximal sequences. It can also be influenced by the architecture of the promoter DNA as well as *cis* and *trans* acting elements. *Cis*-acting elements are DNA sequences which are required for regulation, usually by acting as binding sites for *trans*-acting DNA binding proteins. The UP element (see above) is an example of a *cis*-acting regulatory sequence. Transcription factors, proteins involved in the upregulation or downregulation of transcription, play a large role in regulating which genes are expressed or not. These transcription factors also allow for fine tuning of gene

expression in response to environmental stimuli. Most transcription factors exert their influence through sequence-specific DNA binding (Perez-Rueda and Callado-Vides, 2000).

In Class I activation, the activator binds to a target that is located upstream of the promoter  $-35$  element and recruits RNA polymerase to the promoter by directly interacting with the RNA polymerase  $\alpha$ CTD. The classical example of Class I activation is the action of the cyclic AMP receptor protein, CRP, at the *lac* promoter (Ebright, 1993). The linker joining the  $\alpha$ CTD and  $\alpha$ NTD is flexible, so activators that function using a Class I mechanism can bind at several locations upstream of promoters. In Class II activation, the activator binds to a target that overlaps the promoter  $-35$  element and contacts the RNA polymerase  $\sigma$  subunit (Savery *et al.*, 1998). This contact also results in recruitment of RNA polymerase to the promoter, but other steps in initiation can also be affected. The best example of this is the activation of the bacteriophage  $\lambda$   $P_{RM}$  promoter by the bacteriophage  $\lambda$  CI protein (Nickels *et al.*, 2002). At some promoters that are subject to Class II activation, the activator contacts other parts of the RNA polymerase (for example,  $\alpha$ NTD) but still binds to a target sequence that overlaps with the promoter  $-35$  element. Because of this spatial constraint, there is very little flexibility in positioning of the activator at these promoters.

The third mechanism for simple activation is found in cases where the activator alters the conformation of the target promoter to enable the interaction of RNA polymerase

with the promoter -10 and/or -35 elements. This requires the activator to bind at, or very near to, the promoter elements, although one case has been reported of activation at a distance using a relay of conformational changes through the promoter DNA using SIDD sequences (stress induced destabilization duplexes) (Sheridan *et al.*, 1998). For promoters that are activated by members of the MerR family, the spacing between the -10 and -35 elements at target promoters is not optimal for RNA polymerase binding. MerR-type activators bind to the 'spacer' sequence and twist the DNA to re-orientate the -10 and -35 elements so that they can be bound by the RNA polymerase  $\sigma$  subunit. (Brown *et al.*, 2003). Alternatively, the conformational changes induced at the promoter by the activator could act to undermine the binding affinity of a repressor protein docked at the promoter. This would allow a greater deal of flexibility with regards to activator binding, as its positioning would be relative to the repressor protein, and not the transcription start site as such. A good example of this is the VirB activation of *icsP*, where binding almost a kilobase from the promoter allows for activation by displacement of H-NS (Castellanos *et al.*, 2009).

Repressor proteins reduce transcription initiation at target promoters. At many promoters, repression is simple and involves a single repressor. Three general mechanisms are used. Steric hindrance of RNA polymerase binding to promoter DNA is probably the simplest mechanism of repression. In these instances, the repressor-binding site is located in, or close to, the core promoter elements — for example, the Lac-repressor-binding site at the *lac* promoter (Muller *et al.*, 1996). However, in some cases, the repressor might not prevent binding of RNA polymerase to the promoter,

but instead might interfere with post-recruitment steps in transcription initiation. At other promoters — for example, the *gal* promoter, which is repressed by GalR54 — multiple repressor molecules bind to promoter-distal sites, and repression might be caused by DNA looping, which shuts off transcription initiation in the looped domain (Choy *et al.*, 1992). A similar repressive mechanism is employed by the H-NS protein at some bacterial promoters as discussed previously.

The inherent curvature of the DNA itself can greatly influence gene expression through enhancement or repression of polymerase binding. Supercoiling of the DNA within the bacterial cell is altered by changes in environmental signals such as temperature and osmolarity and can have a localized effect at promoter regions (Travers and Muskhelishvili, 2005). This effect can also be mimicked by the binding of nucleoid-associated proteins at the promoter region, for example IHF. DNA curvature is also subject to modulation through simple changes in environmental conditions such as temperature, pH, and osmolarity which can influence the chemical structure of the molecule (Dorman, 2006). Together these regulatory mechanisms combine to create a vast array of switches and potentiometers for the fine-tuning of transcription initiation and gene expression throughout the life of the bacterium. These regulatory elements and their effects on initiation of transcription are universal throughout the bacterial domain.

## 1.8 Regulation of Virulence Gene Expression in *Shigella flexneri*

### 1.8.1 Regulatory Cascade

The activation of structural virulence genes in *S. flexneri* is regulated in response to a variety of environmental signals (Figure 1.5). Optimal expression of these genes requires a temperature of 37°C (Maurelli *et al*, 1994) and an osmolarity equivalent to that of physiological saline (Porter and Dorman, 1994), and a pH of 7.4 (Nakayama and Watanabe, 1995). It is probable that this profile of conditions, signal to the bacterium that it has passed the acidic environment of the stomach and arrived in the gut of the host, the site of infection and presumably precludes expression either outside the host or in an inappropriate environment within the host. The chromosomally encoded H-NS protein targets many of the operons of the virulence genes located on the A+T rich plasmid and is largely responsible for the silencing of these genes in unfavourable conditions. Activation of the operons occurs via a tightly controlled regulatory gene cascade. H-NS exerts a negative effect on transcription at all levels in the cascade. De-repression involves the products of two plasmid encoded positive regulators, *virF* and *virB*. *VirF* activates the *virB* promoter and in turn the *VirB* protein activates transcription of the structural and effector genes under its control. The *icsA* gene is activated directly by *VirF* and so represents a branch point in the cascade (Dorman and Porter, 1998). Transcriptional regulation also requires input from other chromosomal

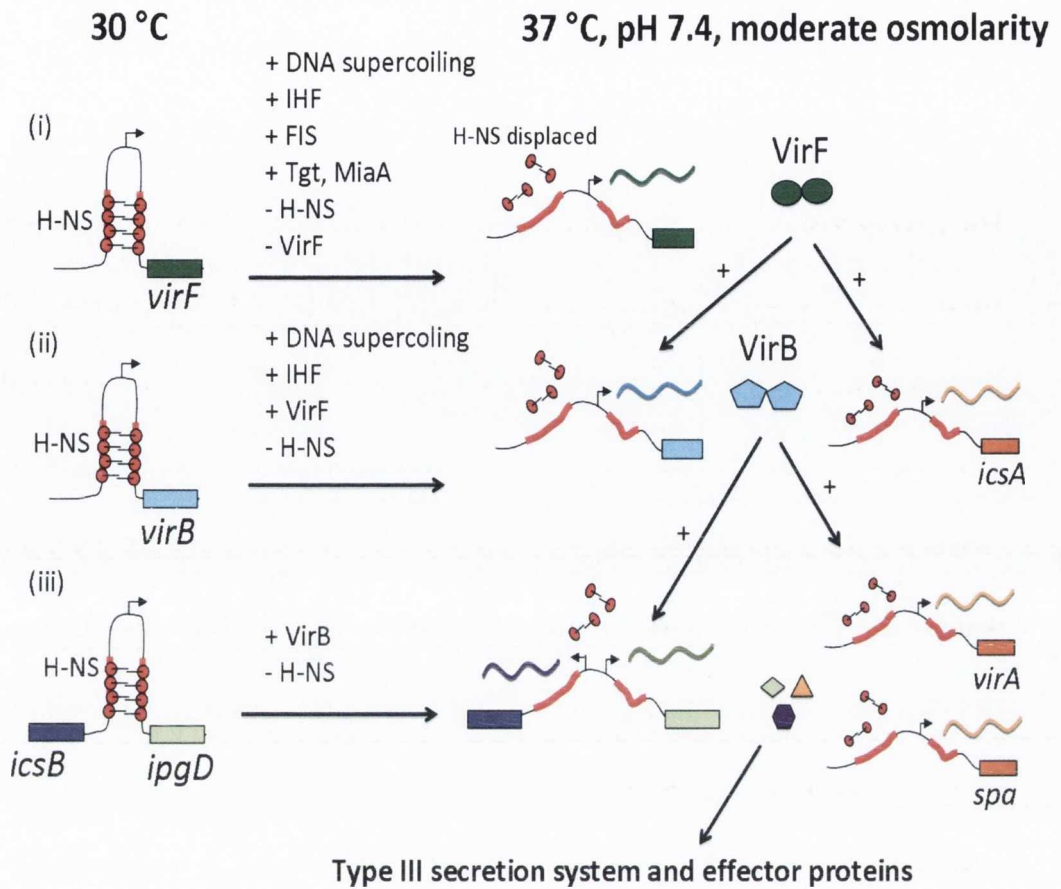


Figure 1.5 Virulence gene cascade in *S. flexneri*.

The schematic diagram portrays the H-NS repressed state of the promoters of the cascade on the left and the de-repressed and gene activated state on the right, upon the receipt of the correct environmental signals (i.e. thermal up shift to 37°C, pH 7.4 and moderate osmolarity). The cascade is divided into three main steps (i) the activation of *virF* (dark green rectangle) in response to environmental signals and transcription factors which overcome the repressive effects of H-NS (red circles) (Section 1.8.2) (ii) the VirF (dark green ovals) dependent activation of *virB* (blue rectangle) which results in the displacement of H-NS and expression of the gene (Section 1.8.3) and (iii) the VirB (blue pentagons) dependent activation of the structural and effector promoters which will ultimately lead to assembly of the type III secretion system (Section 1.8.4)

loci, such as genes encoding the DNA-binding proteins FIS and IHF which also act positively at some of these promoters (Figure 1.5).

Northern blotting data indicate that a gearing effect exists within the cascade upon thermal induction of the regulon with the *virF* gene being induced by about two fold, *virB* by 10-fold and the structural genes by 100-fold (Dorman and Porter, 1997). Therefore as one descends level by level the stringency of transcriptional control is tightened. The primary regulator *virF* displays loose regulation under standard laboratory conditions, expression of *virB* mRNA shows an intermediate level of regulation, while control of the structural genes is most stringent (Dorman and Porter 1997). In addition, a study of the genes of each level of the regulatory cascade revealed individual characteristic responses to stimuli such as osmolarity, pH, variations in DNA supercoiling and the presence or absence of H-NS. It is consistent with a system that allows the regulatory genes to be expressed sufficiently under inappropriate conditions, hence priming the system to ensure a rapid response to inducing conditions when they arise, therefore avoiding energetically wasteful expression of the structural operons. Once induced, fine tuning of the response can be achieved through the different sensitivities of the individual regulon members to external stimuli (Figure 1.5).

### **1.8.2 Regulation of *virF***

Regulation of *virF* represents the first step in full activation of the virulence cascade and assembly of the secretion machinery and is controlled in response to a variety of

environmental cues (Dorman *et al.*, 2001). In addition to negative auto-regulation, it is also subject to thermoregulation but H-NS, facilitated by the flexible DNA curvature inherent in its promoter region. The transcriptional repression of *virF* below 32°C is mediated by H-NS and depends on the presence of an intrinsically bent region whose center at 4°C is located at 137 bp upstream from the transcriptional start. This bend is flanked by two rather extended H-NS sites, one of which overlaps the promoter. H-NS through its DNA bridging activity can co-operatively bind at these two sites, rendering the promoter transcriptionally silent (Falconi *et al.*, 2001). Both extent and localization of this curvature were found to be temperature-sensitive. It was revealed that with increasing temperature there is a reduction of DNA curvature (which is maximum at 4°C) so that the bend collapses when the temperature approaches the transcription-permissive conditions (32°C) thereby allowing the transition of the *virF* promoter from the repressed to the derepressed state. Furthermore, mutations affecting the relative orientation of the two H-NS sites severely affect *in vivo* and *in vitro* H-NS binding to *virF* and the thermoregulation of its expression (Prosseda *et al.*, 2004). It has also been observed that in the absence of the promoter distal site, H-NS is unable to bind the promoter proximal site, even at low temperatures, and repression abolished. These results would suggest the two sites are not independent but act synergistically. In addition, it was shown that when the temperature increases from 4° to 60°C, the centre of the DNA bend within the *virF* promoter slides downstream by almost eight helical turns; it is also noteworthy that the sliding rate is not linear with temperature but undergoes the largest increase within the narrow range (28 °C - 32 °C) which corresponds to the transition from transcriptional repression to derepression (Pon *et*



*al.*, 2000). It is thought that biological significance of these ‘slidings’ is to reveal binding sites for FIS, of which there are four. FIS exercises a direct positive effect on the promoter at 37°C, when H-NS fails to repress transcription, as well as being able to partially counteract H-NS inhibition at the transition temperature of 32°C. Consistent with the importance of local DNA architecture at the *virF* promoter, the DNA-bending protein IHF is required for optimal expression. The protein binds downstream of the promoter (+109 to +122) and possibly aids in overcoming H-NS repression (Porter and Dorman, 1997b).

In addition, *virF* is regulated in response to pH via the CpxRA chromosomal two-component system and the signal transmitted by the CpxR DNA binding protein to the *virF* regulatory gene (Nakayama and Watanabe, 1998) (Section 1.5). Finally, translation of *virF* mRNA is modulated by the tRNA modification genes *tgt* and *miaA* loci (Durand *et al.*, 2000, 2003). Recent observations suggest that TGT may modulate the translation of VirF by base modification of the VirF encoding mRNA (Hurt *et al.*, 2007). This regulation may serve as yet another environmental sensor in the system, as levels of tRNA modification are sensitive to iron and amino acid deprivation (Urbonavicius *et al.*, 2002).

### **1.8.3 Regulation of *virB***

The activation of *virB* represents an important step in the regulatory cascade, as expression leads irreversibly to full commitment to the activation of the structural and

effector genes. Expression of the *virB* gene involves an antagonistic relationship between the activator protein VirF and the repressor protein H-NS. The interaction of the VirF protein with the *virB* promoter has been investigated in detail (Tobe *et al.* 1993). DNA footprint analysis of the promoter region with purified Male'-VirF fusion protein showed that it bound specifically to a sequence upstream of the *virB* promoter. The region encompassed nucleotides -17 to -105 with reference to the transcription start site, placing it at a position where it can interact with RNA polymerase. A protein-binding assay with P<sub>virB</sub>DNA-resin revealed another protein associated with the promoter with a molecular mass of 16 kDa, H-NS. The H-NS protein binds specifically to the *virB* promoter in the region -20 to +20 including the transcription start site which results in exclusion of RNA polymerase and hence repression of the promoter. This places H-NS protein immediately downstream of the VirF protein. *In vitro* transcription analysis revealed that transcription was blocked by the addition of H-NS. While *hns* mutant strains exhibit increased expression at 30 °C, this activation is absolutely dependent on the presence of VirF. VirF was shown to activate *virB*, but only when negatively supercoiled DNA was used as template, conditions that mimic the state of DNA upon a thermal upshift or increase in osmolarity. Thus, the topological changes to the DNA induced by these conditions possibly play an important role in the ability of VirF to activate the *virB* promoter (Dorman *et al.*, 1990; Porter and Dorman, 1994; Tobe *et al.*, 1995).

IHF contributes to the expression of *virB* in stationary phase, with a putative binding site being identified upstream of the promoter (-171 to -183). It has been proposed that

IHF may offset the inhibitory effect that relaxation of negative supercoiling in stationary phase would have on the ability of VirF to activate the promoter, therefore allowing for *virB* expression into late stages of growth (Porter and Dorman, 1997b).

Finally, several lines of evidence from analysis of *S. sonnei*, indicate that post-transcriptional regulation of InvE (VirB) expression plays a key role in the temperature-dependent regulation of virulence gene expression. Most notably the observation that while a considerable amount of *invE* mRNA continues to be transcribed under low temperatures, the production of InvE protein is tightly repressed. This is due to the stability of InvE mRNA being significantly reduced under repressing conditions (Mitobe *et al.*, 2008). Deletion of *hfq*, which encodes an RNA chaperone in Gram-negative bacteria, results in the restoration of expression of *invE* at low temperature due to an increase in the stability of *invE* mRNA. A similar study was conducted which confirmed a similar role for *hfq* in the post-transcriptional regulation of *invE* in response to low osmolarity (Mitobe *et al.*, 2009).

#### **1.8.4 Regulation of *icsB*, *ipgD*, *spa* and *virA***

The VirB dependent promoters, *icsB*, *ipgD*, *icsP*, *spa* and *virA* are all bound directly and repressed by the H-NS protein. Until recently the mechanism by which VirB activated these promoters remained elusive, probably due to the fact that VirB does not resemble any known family of transcription factors. However a detailed dissection of the interplay between both repressor and activator protein at the *icsB* promoter

provided some insight into the regulation of these promoters (Turner and Dorman, 2007).

DNase I footprinting showed VirB binding directly to the *icsB* regulatory region where it affects the structure of this DNA segment as revealed by enhanced sensitivity to cleavage by DNase I. This hypersensitivity is consistent with the wrapping of DNA around the protein with the concomitant exposure of residues, which appear periodically in the footprint. VirB protein protected the DNA from position -105 to -220. However the VirB-mediated distortion of DNA extended throughout the region -105 to +30, the region that is bound by H-NS. Analysis of the DNA revealed a *cis*-acting site that when mutated abolished VirB binding and VirB activation of the promoter. This motif, a near perfect imperfect repeat, 5'-GTTTCATcATGAAAT-3' designated box1 and box2, is located immediately upstream of the region that is bound by H-NS. It is thought to be the initial nucleation site of the VirB from which the protein polymerizes and extends upstream along the DNA. This VirB binding can completely abolish H-NS mediated DNase I protection of this sequence and is indicative of H-NS displacement by VirB. It has also been shown that VirB does not recruit RNA polymerase to the *icsB* promoter. Furthermore, VirB is not required for RNA polymerase to form an open transcription complex or initiate transcript elongation (Turner and Dorman, 2007). Thus, VirB does not perform the functions of a conventional transcription factor, acting simply as an H-NS antagonist at this promoter. This *cis*-acting motif has also been implicated in the positive control of

expression of the operon *ipaBCD* in *S. sonnei* which is under the control of the *icsB* promoter. This positive control requires the InvE (VirB) protein.

The promoter of the *icsP* gene has been documented as being repressed by H-NS and de-repressed by VirB (Catsellano *et al.*, 2009). Here, the consensus VirB binding site is located between positions -1144 and -1130 relative to the *icsP* transcription start site and is absolutely necessary for VirB-dependent derepression (Castellano *et al.*, 2009). Deletion analysis of the promoter region revealed that sequences located over 665 bp upstream of the annotated transcription start site (+1) are needed for complete H-NS-mediated repression of the *icsP* promoter. Using electrophoretic mobility shift assays (EMSAs), it was shown that H-NS binds directly to DNA sequences located both upstream of -665 and downstream of -213 and that both of these regions contain sequences predicted to display high levels of intrinsic curvature (unpublished data from the Wing group). These data support the model of H-NS repression where H-NS bound at remote sites may act co-operatively with H-NS bound at promoter sequences with the resulting micro-loop domain acting to repress transcription of the *icsP* promoter and that VirB functions to remodel the protein:DNA:protein complex, so that transcription of the *icsP* gene can occur.

## **1.9 General regulation of type III secretion systems in bacteria**

Many bacterial pathogens, including species of *Chlamydia*, *Xanthomonas*, *Pseudomonas*, *Shigella*, *Salmonella*, *Escherichia* and *Yersinia*, depend on the type III

secretion to cause disease (Hueck, 1998). Expression of type III secretion systems is coordinately regulated in response to host environmental stimuli by networks of transcription factors. Expression of type III secretion systems responds to an array of environmental conditions which usually correspond to the conditions encountered by the bacteria during infection of a host. Therefore, in many cases, transcription of type III secretion genes is controlled by multicomponent regulatory networks which integrate a diverse set of environmental cues, probably to restrict the energy-consuming expression of 20 or more proteins, unnecessarily. The transcriptional control systems are diverse in various pathogens, although as systems become dissected further recurring themes are becoming apparent. Some systems involve histone-like proteins which regulate gene expression in response to temperature and osmolarity by controlling DNA superhelicity or acting as transcriptional repressors during occupation of inappropriate environments. AraC-like transcriptional activators are also a common feature of these regulatory networks acting to initiate expression when external stimuli are correct. These AraC-like proteins usually act directly on the structural and effector gene operon promoters to activate transcription, and hence begin the assembly of the type III secretion system. In pathogenic *Salmonella enterica*, the pathogenicity island I (SPI-1) encodes a type III secretion system very similar to the one encoded in the entry region of the *S. flexneri* virulence plasmid (Golubeva *et al.*, 2012). SPI-1 encodes several activators including HilA, a member of the OmpR family, and also the AraC-like protein InvF. This protein is required for activation of *sip/sspBCDA* (Eichelberg and Galan, 1999). Another AraC-like activator in type III secretion regulation is the VirF protein of *Yersinia*. Temperature control of *yopH*, *yopE*, *lcrGVH* to *yopBD*, and

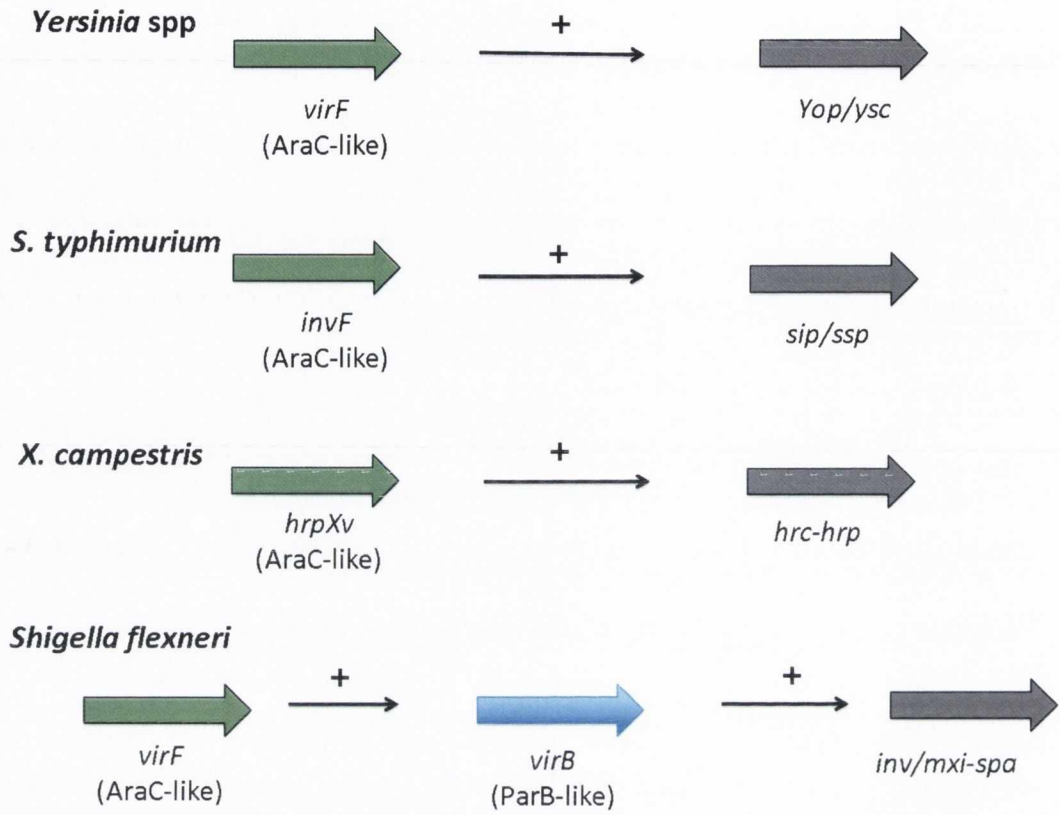


Figure 1.6. Schematic overview of selected transcriptional regulatory networks in several type III secretion systems. Green arrows represent the AraC-like transcriptional regulators and the grey arrows indicate the genes of the type III secretion system which are activated for each species.

*yscA* to *yscL* requires the transacting factor VirF in *Y. enterocolitica*. Genes which are coordinately regulated by VirF have been grouped as the *yop* regulon (Winstanley and Hart, 2001). The type III secretion systems from *R. solanacearum* and *X. campestris* are regulated in response to environmental stimuli (growth in minimal media) by AraC-like proteins HrpB and HrpXv respectively, which activate a subset of genes that encode essential proteins of the type III secretion system apparatus (Hueck, 1998) (Figure 1.6).

In contrast to these regulatory networks the *Shigella flexneri* virulence plasmid encodes an extra checkpoint of regulation prior to full commitment of secretion apparatus assembly. The AraC-like VirF protein activates an intermediate regulator VirB (a ParB-like protein), which subsequently de-represses the operons of the structural and effector genes. The reasons for this additional step have been subject to much speculation (Turner and Dorman, 2007) (Figure 1.5) (see section 1.11).

### **1.10 The ParB family of plasmid partitioning proteins.**

The stable inheritance of DNA at cell division, including both plasmids and the chromosome, is an essential process in bacteria. Eukaryotic cells rely on the mitotic apparatus including the spindle, kinetochore and the centromere to ensure faithful segregation of plasmids (Miyazaki and Orr-Weaver, 1994). In prokaryotes high copy number plasmids do not generally require active partitioning mechanisms, as the concentration of the plasmid in the cell ensures that at least one copy is likely to pass



to each daughter cell. However low copy number plasmids and chromosomal DNA in bacteria require a system that physically connects the DNA molecule to the components of the partition machinery resulting in stable maintenance (Wake and Errington, 1995). The virulence plasmid of *Shigella* has two active and fully functioning plasmid-partitioning systems, in keeping with its mosaic structure, located on opposite sides of the plasmid.

The first of these partitioning systems consists of the *stbA* and *stbB* genes encoding the StbA and StbB protein. Sequence analysis of the plasmid identified a putative *cis*-acting site exhibiting a strong A C/T G bias and several repeats (Tabuchi *et al.*, 1988) are present upstream from *stbA* (Buchrieser, 2000).

The second system uncovered from sequencing is a homolog of the partition region of the plasmid prophage of bacteriophage P1. The P1*par* (partition) region consists of an operon containing *parA* and *parB* genes and a centromere analog site, *parS*, that lies downstream. The *parA* gene encodes a Walker-type ATPase essential for plasmid movement during partition and is thought to supply energy (Funnell, 1991). Although, incapable of binding to *parS*, ParA binds to the DNA at the *par* promoter in a sequence specific manner to auto-repress the operon (Bouet and Funnell, 1999). The *parB* gene encodes a protein that can bind tightly to the partition site, *parS* forming a nucleoprotein complex with the chromosomally encoded IHF protein and is required for capture of the plasmid at the cell centre prior to partition. It is thought that plasmids are paired via interaction between ParB-*parS* complexes on two plasmids and

subsequently plasmids are positioned at the poles of the dividing cell such that both daughter cells receive a plasmid each (Nordstrom and Austin, 1989; Jensen and Gerdes, 1997).

Phylogenetic analysis of the sequence divergence of the ParB-like proteins divides the family into three distinct groups, chromosomal ParB-like proteins, phage/plasmid SopB-like proteins and phage/plasmid ParB-like proteins, all of which diverged from a common ancestor. In general these proteins are transcribed as part of an operon, which includes a ParA-like protein and a *cis*-located ParB binding site (i.e. *parS*). The phage/plasmid SopB-like proteins are all located on episomal DNA and show most homology to the SopB protein, responsible for the partitioning of the F plasmid in *E. coli* (Mori *et al.*, 1989). This group is closely related to the phage/plasmid ParB-like proteins, which are also episomally located.

It should be recalled that the VirB protein is a member of the phage/plasmid ParB-like group of proteins. The presence of two seemingly functional partition systems (*parAB* and *stbAB*) and of the remnant of a third one (*virB*) suggests that parts of the *Shigella* virulence plasmid come from three plasmids. Moreover, the G+C content of the replication system (*repA*, 58%) is different from those of the above mentioned partition systems, *virB* (34%), *stbAB* (39%) and *parAB* (43%), which suggests that the virulence plasmid contains elements that were initially carried by four plasmids (Buchrieser, 2000). VirB is the only one of the entire ParB family of proteins, characterized to date that is not associated with the partitioning of DNA at cell division. Thus VirB

represents a unique member of the ParB family of proteins, a protein that retains a high level of sequence homology but appears to have evolved toward a different function, that of virulence gene regulation (Sergueev *et al.*, 2005, Beloin *et al.*, 2002) (Figure 1.3).

### **1.11 VirB; plasmid partitioning protein co-opted into a regulatory role**

The amino acid sequence homology between VirB and the plasmid partitioning proteins (Figure 1.3) and the nucleotide sequence similarities between plasmid partitioning *parS* elements and the box1/box2 sequence to which VirB binds (Figure 1.4) strongly suggests that VirB is a former plasmid partitioning protein that has been redirected into a regulatory role through evolution. The *Shigella* virulence plasmid is a mosaic structure which contains two fully functioning plasmid partitioning systems, and while there is no evidence to suggest VirB contributes to the maintenance of the virulence plasmid, it is tempting to speculate that it is the vestige of a third system, redundant in the context of the modern plasmid. When the genetic maps of VirB and ParB genes are aligned, a few stark differences are apparent. Firstly, neither the region surrounding the *virB* gene, nor any of the virulence plasmid, seems to encode a partner protein equivalent to that of ParA, an ATPase that interacts with ParB when plasmid partitioning occurs.

Perhaps the loss of the gene encoding the putative VirB partner protein that is essential for plasmid partitioning to function effectively, coupled with the existence of two systems already in place, resulted in VirB becoming redundant as a partitioning protein. This would have allowed VirB to then become available for other roles. The second difference between the genetic maps relates to the positioning of the VirB and ParB

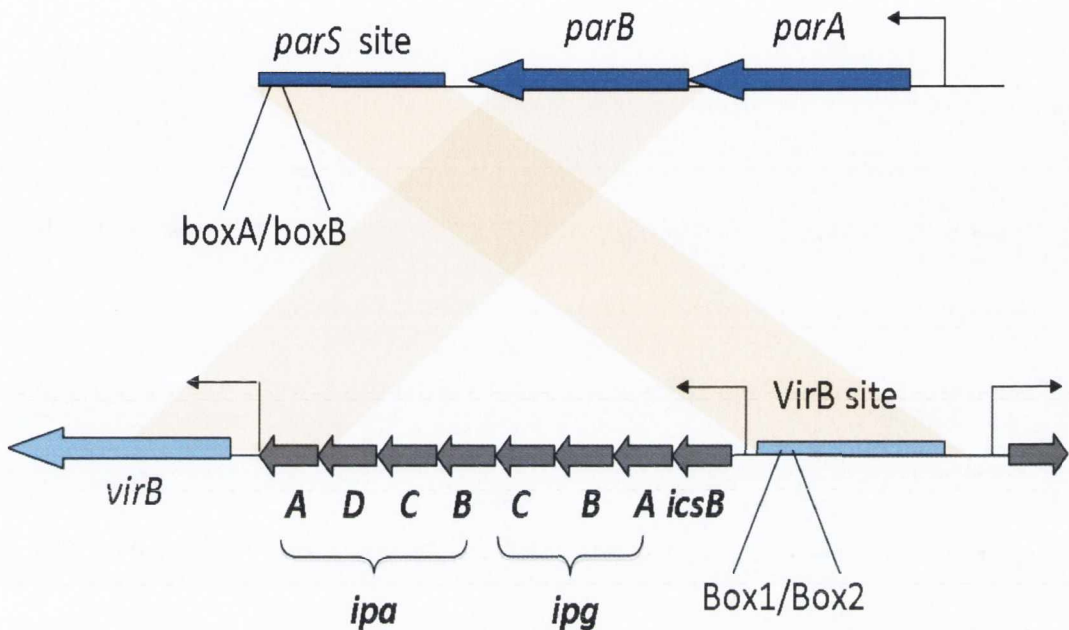


Figure 1.7. The alignment of the regulatory regions of the *parAB* operon and the *virB* gene and its binding site at the *icsB* promoter. The illustration indicates the position of the *parS* binding site in relation to the *parA* and *parB* genes. Below is the genetic map of a portion of the entry region showing the relative locations of the *virB* gene and the regulatory sequences of the *icsB*-*ipg*-*ipa* operon. The angled arrows represent promoters. The blue arrows represent genes while the filled blue rectangles represent binding sites. Also represented is the essential box1/box2 nucleation site within the VirB site. The pale pink parallelogram is indicative of regions of homology between the nucleotide sequences of VirB binding site and *parS* and amino acid homology between *virB* and *parB* genes.

binding sites. The VirB box1/box2 *parS*-like binding site is located at some distance from the *virB* gene. In contrast, the *parS* sequences are located immediately downstream of their respective genes. If the nucleation 'box1/box2' site was originally in the promoter region of *virB* then it has become disconnected from its origin by the insertion of the *icsB/ipg/ipa* operon, an event that would place the motif at its current position at the *icsB* promoter. This evolutionary step would have placed the VirB binding site at a position where, due to its DNA binding and wrapping abilities, it could now act to antagonize H-NS, and de-repress the *icsB* operon (Figure 1.7). As the overwhelming negative effects of H-NS on the expression of genes of the virulence plasmid had to be overcome in order to benefit from the genetic material being silenced. Therefore this redirecting of an apparently redundant DNA plasmid partitioning protein would be advantageous to the cell and so would be conserved. Through point mutations, binding sites for VirB could then be selected for if they appeared in the promoter of an H-NS repressed virulence gene.

The genetic rearrangements discussed here are simple and analogous to modification happening routinely in bacteria. Pathogenicity Islands, in particular, are characterized by a mosaic-like structure, association with mobile genetic elements, and genetic instability. Both the *Shigella* chromosome and the virulence plasmid reflect these dynamics by containing numerous insertion sequences (IS) and markers of genomic rearrangements. The study of VirB hints at the flexibility of bacterial regulatory circuits and their ability to co-opt DNA binding proteins for new regulatory roles. It also highlights the danger in assigning unique biological properties to DNA-binding proteins.

## 1.12 Summary

In summary, the expression of several structural virulence genes in *Shigella flexneri*, including a type III secretion system and invasions, is crucial to the ability of the bacteria to adapt to and infect the host. Expression of these genes in an inappropriate environment, with no potential for invasion, is energetically costly and wasteful, and so the control of expression must be stringent. The VirF and VirB proteins act in a regulatory cascade to overcome the overwhelming negative effect of H-NS on transcription of the virulence genes. The VirB protein, an unusual transcription factor, is a homologue of the ParB plasmid partitioning family, but instead serves as an H-NS antagonist in this virulence regulatory cascade. Previous work has suggested fortuitous placement of an appropriate *cis*-acting site in the regulatory region of the *icsB* region, and the redundant nature of VirB protein as a partitioning protein, has allowed the development of a new regulatory switch. We hypothesize here that the insertion of the *icsB* operon, during evolution, immediately upstream of *virB* gene, may have disconnected VirB from its binding site while simultaneously placing this *parS*-like site in a position where it can interfere with H-NS repression. The gene rearrangements involved are simple and analogous events may occur routinely in bacteria. The purpose of this study was to investigate how easily H-NS antagonism can arise, by the development of new genetic switches. The hypothetical evolutionary step proposed here was repeated by insertion of a binding site for VirB upstream of an H-NS promoter, not currently subject to VirB activation. This serves to expand our

knowledge of the development of regulatory networks, and provide further insight into the evolution and regulation of *Shigella*'s virulence cascade.

## **Chapter 2**

### **Materials and Methods**



## **2.1 Chemicals and growth media**

### **2.1.2 Chemicals and reagents**

All chemicals and reagents used in this study were purchased from BDH Chemicals Ltd, Bio-Rad, Clontech, Invitrogen™, Novagen, PALL, Pfizer, Pierce, Promega, Quiagen, RBA Bioscience, Roche Molecular Biochemicals and Sigma Chemical Company. DNA restriction and modifying enzymes were obtained from Roche, New England Biolabs (NEB), Promega, Stratagene or Fermentas. In addition, several molecular biology 'kits' were used during the study. The basic principle of each kit is described in brief without giving a complete protocol.

### **2.1.3 Growth media**

Materials for preparing media were obtained from Difco (BD Diagnostic systems). All media were sterilized by autoclaving at 121°C for 15-20 min prior to use. Aqueous solutions which would be damaged by autoclaving e.g. amino acid or antibiotic solutions were sterilized by filtration through 0.2 µm SupaTop Syringe Filters (Anachem). Media were supplemented with the appropriate antibiotics as required. All quantities listed below are for the preparation of 1 litre of medium in distilled, de-ionised water (ddH<sub>2</sub>O). Solid media were prepared by the addition of 15 g of agar per litre of medium prior to autoclaving .

L 10 g bacto-tryptone, 5 g yeast extract, 5 g NaCl, pH 7

LO 10 g bacto-tryptone, 5 g yeast extract, pH 7

### **2.1.3 Antibiotics**

All stock antibiotics were stored in aliquots at -20°C and those prepared in water were sterilized by filtration through 0.2 µm SupaTop Syringe Filters (Anachem).

Carbenicillin was prepared as 100 mg ml<sup>-1</sup> stock solution in ddH<sub>2</sub>O and used in media at a concentration of 100 µg ml<sup>-1</sup>, Kanamycin was prepared as 50 mg ml<sup>-1</sup> stock solution in ddH<sub>2</sub>O and used at a final concentration of 50 µg ml<sup>-1</sup> Tetracycline was prepared as 10 mg ml<sup>-1</sup> stock solution in ethanol and used at a final concentration of 10 µg ml<sup>-1</sup>

## **2.2 Bacterial strains and culture conditions**

### **2.2.1 Bacterial strains.**

All bacterial strains used in the study were derivatives of *Shigella flexneri* 2a or *Escherichia coli* K-12. These are listed in Table 2.1 together with source and genotype.

Bacterial strains were maintained as permanent stock in 15% glycerol at -80°C.

### **2.2.2 Bacterial culture conditions**

Bacterial strains were routinely cultured in L broth unless otherwise stated. L agar plates were commonly used throughout this study for culturing of strains, for selection of transformants and transductants. Typically bacterial cultures were grown aerobically, at 200 rpm, at 30°C or 37°C. Stationary phase cultures were obtained by inoculating single colonies into 2 ml broth in a sterile test tube and grown overnight. Exponential

Table 2.1 List of Strains

Strain	Relevant characteristic(s)	Reference
<i>S. flexneri</i> 2a 2457T		
BS184	<i>mxiC::mudI1734</i> Km <sup>R</sup>	Maurelli, A.T., 1984
BS185	BS184 <i>hms::Tn10</i> Tc <sup>R</sup>	Maurelli, A.T., 1984
CJD1018	BS184 <i>virB::pMEP151</i>	Porter, M.E., 1997
BS184 <i>proU-tet</i>	BS184 <i>proU::tetRA</i>	This Study
BS184 <i>proU-gfp</i>	<i>proU</i> promoter (-270 to +330), <i>gfp</i> gene, <i>cat</i> gene from pZep1 inserted into <i>tetRA</i> locus of BS184 <i>proU-tet</i>	This Study
BS184 <i>icsBR-proU-gfp</i>	<i>icsB</i> region, <i>proU</i> promoter (-270 to +330), <i>gfp</i> gene, <i>cat</i> gene from pZep2 inserted into <i>tetRA</i> locus of BS184 <i>proU-tet</i>	This Study
BS184 <i>icsBF-proU-gfp</i>	<i>icsB</i> region, <i>proU</i> promoter (-270 to +330), <i>gfp</i> gene, <i>cat</i> gene from pZep3 inserted into <i>tetRA</i> locus of BS184 <i>proU-tet</i>	This Study
BS184 <i>icsBM-proU-gfp</i>	<i>icsB</i> region, <i>proU</i> promoter (-270 to +330), <i>gfp</i> gene, <i>cat</i> gene from pZep2M inserted into <i>tetRA</i> locus of BS184 <i>proU-tet</i>	This Study
<i>E. coli</i> K-12		
XL1-Blue	<i>recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac ΔFΔ proAB lacIqZ_M15 Tn10</i> (Tc <sup>R</sup> )	Stratagene
MG1655	F <sup>-</sup> Δ <i>ilvG rfb-50 rph-1</i>	Lab stocks

phase cultures were obtained by inoculating broth 1/100 with bacterial culture and monitoring growth to logarithmic phase by spectrophotometry. Volumes of 2-3 ml were grown in test tubes, while larger volumes 10-25 ml were grown in 250 ml flask, so as not to limit the available oxygen concentration.

## **2.3 Plasmids, bacteriophage and oligonucleotides**

### **2.3.1 Plasmids**

Plasmids used in this study are listed in Table 2.2 together with genotype and source. Any details of plasmid construction will be described in the appropriate sections of the results chapter.

### **2.3.2 Bacteriophage**

The bacteriophage used in this study was bacteriophage P1 *vir*.

### **2.3.3 Oligonucleotides**

The sequences of all primers used in the study are listed in Table 2.3. Oligonucleotides were purchased from MWG-Biotech or Integrated DNA Technologies (IDT). Primers were delivered as lyophilized DNA and were resuspended in ddH<sub>2</sub>O to a final concentration of 100 pmol/μl before use.

Table 2.2 List of Plasmids

Plasmid	Relevant Characteristic(s)	Reference
pZep08	GFP reporter plasmid Ap <sup>R</sup> Cm <sup>R</sup>	Hautefort, I., 2003
pZep- <i>proU</i> 1	-270 to +230 P <i>proU</i> cloned into pZep08	This study
pZep- <i>proU</i> -2	-240 to -80 of <i>PicsB</i> cloned into pZep- <i>proU</i> -1 in reverse orientation	This study
pZep- <i>proU</i> -2M	-240 to -80 of mutated <i>PicsB</i> cloned into pZep- <i>proU</i> -1 in reverse orientation	This study
pZep- <i>proU</i> -3	-240 to -80 of <i>PicsB</i> cloned into pZep- <i>proU</i> -1 in forward orientation	This study
pZep- <i>proU</i> -3M	-240 to -80 of mutated <i>PicsB</i> cloned into pZep- <i>proU</i> -1 in forward orientation	This study
pZep- <i>proU</i> -4	-150 to +230 P <i>proU</i> cloned into pZep08	This study
pZep- <i>proU</i> -5	-240 to -80 of <i>PicsB</i> cloned into pZep- <i>proU</i> -4 in reverse orientation	This study
pZep- <i>proU</i> -5M	-240 to -80 of mutated <i>PicsB</i> cloned into pZep- <i>proU</i> -4 in reverse orientation	This study
pZep- <i>proU</i> 5- <i>ipgD</i>	-240 to -80 of <i>PicsB</i> cloned into pZep- <i>proU</i> -4 in forward orientation	This study
pZep- <i>proU</i> -6	-60 to +230 P <i>proU</i> promoter cloned into pZep08	This study
pZep- <i>proU</i> -7	-240 to -80 of <i>PicsB</i> cloned into pZep- <i>proU</i> -6 in reverse orientation	This study
pZep- <i>proU</i> -7M	-240 to -80 of mutated <i>PicsB</i> cloned into pZep- <i>proU</i> -6 in reverse orientation	This study

pZep- <i>proU7-ipgD</i>	-240 to -80 of <i>PicsB</i> cloned into pZep- <i>proU-6</i> in forward orientation	This study
pZep- <i>proU-8</i>	-240 to -80 of <i>PicsB</i> cloned into pZep- <i>proU-6</i> in reverse orientation with the URE upstream	This study
pZep- <i>proU-8M</i>	-240 to -80 of mutated <i>PicsB</i> cloned into pZep- <i>proU-6</i> in reverse orientation with the URE upstream	This study
pZep- <i>proU8-ipgD</i>	-240 to -80 of <i>PicsB</i> cloned into pZep- <i>proU-6</i> in forward orientation with the URE upstream	This study
pZep- <i>proU-9</i>	210 bp <i>yaeT</i> coding region cloned into pZep- <i>proU-6</i>	This study
pZep- <i>proU-10</i>	-240 to -80 of <i>PicsB</i> cloned into pZep- <i>proU-9</i> in reverse orientation	This study
pZep- <i>proU10-ipgD</i>	-240 to -80 of <i>PicsB</i> cloned into pZep- <i>proU-9</i> in forward orientation	This study
pZep- <i>virB-310</i>	-310 to +20 <i>PvirB</i> cloned into pZep08	This study
pZep- <i>virB-310</i> SDM	pZep- <i>virB-310</i> with putative <i>VirB</i> binding site mutated	This study
pZep- <i>virB-150</i>	-150 to +20 <i>PvirB</i> cloned into pZep08	This study
pZep- <i>virB-80</i>	-80 to +20 <i>PvirB</i> cloned into pZep08	This study
pZep- <i>virB-80-icsB</i>	-240 to -80 of <i>PicsB</i> cloned into	This study
pZep- <i>virB-80-icsBX2</i>	-240 to -80 of <i>PicsB</i> cloned into pZep- <i>virB-80</i> twice	This study
pZep- <i>virF-380</i>	-380 to +64 <i>PvirF</i> cloned into pZep08	This study
pZep- <i>virF-280</i>	-280 to +64 <i>PvirF</i> cloned into pZep08	This study
pZep- <i>virF-280</i> SDM	pZep- <i>virF-280</i> with putative <i>virF</i> binding site mutated	This study

---

Table 2.3 List of Primers

Oligonucleotides	Sequence (5' to 3')
<i>icsB</i> fw.TAG	tgtacctcgtgagcatatgtagt
<i>icsB</i> rev.TAG	ggggcattgattgcagtttt
<i>proU</i> rev +230.XbaI	ttttctagatatctcgccttctcaatgg
<i>proU</i> fw -270.NotI	tttgccggccgcctaccgcagcagggaaata
<i>proU</i> fw -150.NotI	tttgccggccgcctcgtatctttgacaaaaatatac
<i>proU</i> fw -60.NotI	tttgccggccgcgccacatttgccatcaggg
<i>proU</i> SDM FseI fw	ttgctcgcatacaatattcaGGCCGGCCttgccatcaggggttgctc
<i>proU</i> SDM FseI rev	gaggcaaccctgatggcaaGGCCGGCCtgaatattgatgcgcaaa
<i>yaeT</i> fw.SmaI	tttcccgggaagaacgtaacactggtagct
<i>yaeT</i> rev. NotI	tttgccggccgcgacaggtcggcgtcatctg
<i>tet</i> .40 bp <i>proU</i> fw	<u>aagactggaatttctgagggtgttat</u> tttcaaaatatactcattaagaccc actttcacatt
<i>tet</i> 40 bp <i>proU</i> rev	<u>cgcggcactgtacccgctggcgtggatcccacggattactaagcag</u> ttgtctctctg
<i>icsB</i> .fw.40 bp <i>proU</i>	<u>aagactggaatttctgagggtgttat</u> tttcaaaatatactgtacctcgtga gcatatgt
<i>icsB</i> .rev.40 bp <i>proU</i>	<u>aagactggaatttctgagggtgttat</u> tttcaaaatatacgggggattgatt gcagtttt
<i>cat</i> .rev.40 bp <i>proU</i>	<u>cgcggcactgtacccgctggcgtggatcccacggattacacttattca</u> ggcgtagcag
<i>rho</i> .RT.fw	gttgtagtcgatagccgcga
<i>rho</i> .RT.rev	tacaggcaacatggaactgc
P1.RT.fw	cttatatacagggaaattcc
P2.RT.fw	gggtaatatatcgacatagac
<i>gfp</i> .RT.rev	agtgaaaagtcttctctct
<i>proU</i> .fw.BIO	BIO-ctacccgcagcagggaaata
<i>proU</i> .rev	tatctcgccttctcaatggcc
<i>icsB</i> .fw.BIO	BIO-cctcgtgaccatagtagtg
<i>virB</i> .rev.+190.XbaI	ttttctagattttggttgacgaagggttaaatc

<i>virB</i> .fw.-80.NotI	ttt <u>g</u> cg <u>g</u> ccgcttgcacatccagctattaa
<i>virB</i> .fw.-310.NotI	ttt <u>g</u> cg <u>g</u> ccg <u>g</u> ccgattctctttctctgatte
<i>virF</i> .rev.+160.XbaI	tttctagaaatgacggttagctcaggca
<i>virF</i> .fw.-280.NotI	ttt <u>g</u> cg <u>g</u> ccgctggagcctccagctcgaag
<i>virF</i> .fw.-380.NotI	ttt <u>g</u> cg <u>g</u> ccg <u>g</u> ccttcacgatcgcaatatgga
<i>virB</i> .RT.fw.+2	actgcatttaactttgtcaata
<i>virB</i> .RT.rev	tttgggtgacgaaggttaaate
<i>virF</i> .RT.fw	ggttcgcttgcataactata
<i>virF</i> .RT.rev	ataaaagcaatttgccttc
<i>virB</i> .fw.BIO	BIO-gccgattctctttctctgatt
<i>virB</i> .rev.+1	gagcaagattgtatgtaactgt
<i>virF</i> .fw.BIO	BIO-tggagcctccagctcgaag
<i>virF</i> .rev.EMSA	accgctattgaacaccttt
<i>virB</i> .fw.SDM	gattctctttctctgGCTGGGCgctggatatttag
<i>virB</i> .rev.SDM	ctaatcatatccagcGCCAGCagagaaagagaate
<i>virF</i> .fw.SDM	tatggttatagtcccGGGcagtgcGGGCacttagcttgttga
<i>virF</i> .rev.SDM	tgaacaagctacgtGCCGcactgCCCgggactataaccata

---

<sup>a</sup> Restriction enzymes used for cloning, and 40 bp regions of homology used for Lambda Red recombination are underlined

<sup>b</sup> BIO denoted 5' biotin tags

<sup>c</sup> Capitalised base pairs indicate mutated nucleotides used for site-directed mutagenesis



## 2.4 Genetic techniques

### 2.4.1 Generalized transduction using p1 phage

Phage transduction using bacteriophage P1 *vir* was used to move selectable genetic markers, in this case mutant genetic alleles, between *Shigella* strains. Phage P1, which encapsulates ~90 kb of DNA, is commonly used as a transducing agent because it is a generalized transducer (it can package random sections of the host chromosome instead of its own genome) giving rise to "transducing particles". P1 *vir* is a mutant phage that enters the lytic cycle upon infection, ensuring replication and lysis, and preventing the phage from generating P1 lysogens (circular phage molecules that are capable of self-replicating) among transductants (Sternberg and Hoess, 1983). Once a phage population has been generated from a donor host, the phage are used to infect a recipient host. Most of the bacteria are lysed by phage that packaged P1 genomes, but a fraction of the phage inject a genome segment derived from the donor host. Homologous recombination then allows the incoming genomic segment to replace the existing homologous segment. Calcium is required for infectivity.

### 2.4.2 Preparation of high titre P1*vir* lysate

A high titre lysate of the donor strain (strain containing the allele for transduction) was made as outlined previously (Thomason et al., 2007). An overnight culture of donor strain (grown with antibiotic selection) was diluted 1/100 in fresh LB supplemented with 5 mM CaCl<sub>2</sub> and 0.2% glucose (3 ml per lysate), and was grown with aeration at 37°C 200 rpm for 1-2 hours. This was done in duplicate, with one serving as a no

phage control. When cells are in early log phase ( $OD_{600}$  0.4-0.6) 100  $\mu$ l of high titre P1 vir ( $10^9$  phage/ml) was added to one of the cultures. Lysis could be monitored by comparison of turbidity with the phage free culture. Typically lysis was visible 1 hr after addition of the phage, which is characterized by visible clearing of the culture and accumulation of cell debris clumping. However this may take longer for strains with a slower growth rate. Several drops (50-100  $\mu$ l) of chloroform were added to sterilize the lysate which was then vortexed and centrifuged at 13,000 rpm for 3 min to pellet cell debris and any remaining whole cells. The supernatant was then passed through a 0.45  $\mu$ M filter into 2 ml sterile Eppendorfs tubes and stored at 4°C. If the lysate was to be used immediately the lysate was incubated with the caps open at 37°C for 20 min to allow for any residual chloroform to evaporate. It could then be added to the recipient cells without the risk of killing them.

To determine the titre of the resulting phage lysate stock, an agar plate containing 5 mM  $CaCl_2$  was covered with a strain of interest and allowed to dry. A series of 10 fold dilutions of the lysate were made and 10  $\mu$ l of each spotted on the agar and incubated upright at 37°C overnight. Each infectious particle produces a circular zone of infected cells called a plaque, which can be counted and the titer calculated in plaque forming units (PFU) per ml.

### **2.4.3 Transduction with P1 vir**

Recipient strains were grown in 2 ml cultures overnight in LB medium. Cells were harvested by centrifugation at 6000 rpm for 3 min and re-suspended in 2 ml L broth

containing 10 mM MgSO<sub>4</sub> and 5 mM CaCl<sub>2</sub>. The reaction mixtures were set up as follows in 2 ml Eppendorfs; A. 100 µl phage (negative control), B. 1 µl phage + 100 µl cells, C. 10 µl phage + 100 µl cells, D. 100 µl phage + 100 µl cells, and E. 100 µl cells (negative control). The tubes were mixed gently and incubated at 37 °C for 30 min in a static waterbath. To stop the infection, 200 µl of 1M Sodium citrate (pH 5.5) was added to the tubes, as Sodium citrate chelates the Ca<sup>+</sup> ions required for phage adhesion and thus inhibits further infection. 1 ml of LB was then added to each tube and incubated with shaking for 1 hr at 37°C to allow expression of the antibiotic resistance gene being selected for. Cells were centrifuged and the supernatant discarded. The pellets were re-suspended in the residual supernatant and plated into agar containing 10 mM Sodium citrate and the appropriate antibiotic containing plates. Transductants were streaked to single colony purity and screened by PCR for the allele of interest. Permanent stocks of the new strain were then stored.

## **2.5 Purification of Nucleic Acids**

### **2.5.1 Isolation of plasmid DNA**

Plasmid DNA was routinely purified from *E.coli* and *S. flexneri* strains using the HiYield plasmid mini-kit (RBC Biosciences) from 3 ml cultures according to the manufacturer's instructions. The procedure is based on a modified alkaline lysis method of Ish Horowicz and Burke 1981, where bacteria are lysed and an RNase treatment is used to get cleared cell lysate with minimal genomic DNA and RNA contaminants. The denaturation of the DNA into single stranded molecules separates

the chromosomal DNA from plasmid DNA, as the plasmid being small closed circular DNA, easily reforms. Only double stranded DNA is bound by the silica matrix spin column, provided in the kit, thus isolating the plasmid. Contaminants are washed with an ethanol wash buffer. In the final step, the purified plasmid DNA is eluted by 50  $\mu$ l of low salt elution buffer or distilled water.

### **2.5.2 Isolation of chromosomal DNA**

Purified genomic DNA was required as template for PCR or as standards controls in qRT-PCR reactions. Chromosomal DNA was isolated using the Puregene DNA Purification Kit (Gentra) which uses a method modified from Bubbone, 1985; Davis, 1980, which uses salt as a substitute for toxic organic solvents in the deproteination step. Briefly, 500  $\mu$ l of a turbid cell culture was placed in a sterile Eppendorf tube on ice. Cells were pelleted at 13,000 rpm for 5 seconds, and the supernatant removed. The bacterial pellet was re-suspended in Cell Lysis Solution and heated at 80°C for 5 min to lyse cells. The Cell Lysis Solution is an anionic detergent containing DNA Stabilizers that work by limiting the activity of DNases that are contained in the cell and elsewhere in the environment. The samples are stable for at least two years at room temperature. Contaminating RNA was then removed, by adding RNase A Solution, an RNA digesting enzyme, and heating at 37°C for 30min. After cooling the cell lysate on ice for 1 min, other contaminants, such as proteins, were removed by salt precipitation with the addition of Protein Precipitation Solution. The precipitated protein will form a tight pellet when centrifuged at 13,000 rpm, and the supernatant containing genomic DNA was recovered by precipitation with 100% Isopropanol (2-propanol). 70% Ethanol

was used to wash the DNA and then dissolved by incubation at 65°C for 1 hr in DNA Hydration Solution, a buffered solution containing a DNA stabilizer. The DNA was then stored at 4°C.

### **2.5.3 Purification of linear DNA**

Linear DNA fragments (PCR products or cleaved DNA) were purified for cloning or for the preparation of labeled probes using the HiYield gel/PCR DNA fragments extraction kit (RBC BioScience). The DNA was recovered after electrophoresis through a 1 x TAE (40 mM Tris, 1 mM EDTA, 0.114% (v/v) glacial acetic acid) agarose gel containing 1 g/ml of ethidium bromide. The DNA was visualized by a brief exposure to short wave (254 nm) UV light. The desired fragment was cut out using a surgical blade and purified following the guidelines supplied. Briefly, the method uses a chaotropic salt, guanidine thiocyanate to dissolve agarose gel and denature enzymes. DNA fragments in chaotropic salt solution bind to the glass fiber matrix of the spin column. Following washing off of contaminants with 80% ethanol, the purified DNA fragments are eluted by addition of low salt Tris-HCl (pH 8.0) elution buffer or water.

### **2.5.4 Isolation of RNA**

The purity and integrity of RNA isolated from tissue or cultured cells are critical for its effective use in applications, in this study, reverse transcription PCR (RT-PCR). The successful isolation of intact RNA requires four essential steps: effective disruption of cells or tissue, denaturation of nucleoprotein complexes, inactivation of endogenous ribonuclease (RNase) activity and removal of contaminating DNA and proteins. Total

RNA was isolated from cultures using the SV Total RNA Isolation System (Promega). OD<sub>600</sub> 2.0 units of bacterial culture was harvested (e.g. 4mls of a culture with an OD<sub>600</sub> 0.5). A solution of ice-cold 5% (v/v) phenol pH 4.3, 95% (v/v) ethanol, of a volume of 2/5 that of the bacterial culture was added to the cells (e.g. 4mls culture to 1.6 phenol/ethanol mix). This mixture was kept on ice for 30 min to stabilize the RNA and prevent degradation. Samples were then centrifuged at 4000 rpm at 4°C for 10 min and the supernatant discarded. At this point the pellets were frozen at -80°C if not used immediately. To aid in the lysis of the cells, the pellet was re-suspended in 100 µl TE buffer (10 mM Tris-HCl, 1 mM EDTA) containing 50 mg/ml lysozyme and incubated at room temperature for 5 min. 75 µl of RNA Lysis Buffer was then added to the mix. The RNA Lysis Buffer combines the disruptive and protective properties of guanidine thiocyanate (GTC) and β-mercaptoethanol to inactivate the ribonucleases present in cell extracts (Chergwin, 1979). GTC, in association with SDS, acts to disrupt nucleoprotein complexes, allowing the RNA to be released into solution and isolated free of protein. Subsequently 350 µl RNA Dilution Buffer was added and mixed well by inversion, followed by heating to 70°C for 3 min. Dilution of cell extracts in the presence of high concentrations of GTC causes selective precipitation of cellular proteins to occur, while the RNA remains in solution. After centrifugation at 13000 rpm for 10 min, to clear the lysate of precipitated proteins and cellular debris, the supernatant was transferred to a clean tube and 200 µl of ethanol added. The sample was transferred to a spin column and centrifuged for 30 sec. The ethanol selectively precipitates the RNA and it is bound to the silica surface of the glass fibers found in the spin column. The binding reaction occurs rapidly due to the disruption of water

molecules by the chaotropic salts, thus favouring adsorption of nucleic acids to the silica. The bound RNA was washed with 600  $\mu$ l wash buffer to further remove contaminating salts, proteins and cellular impurities. RNase- Free DNase I was applied directly to the silica membrane to digest contaminating genomic DNA. A Dnase stop mix was added after 15 min and the wash step preformed twice. Finally, the total RNA was eluted from the membrane by the addition of 100  $\mu$ l Nuclease-Free Water. This procedure yields an essentially pure fraction of total RNA and quality was assessed by electrophoresis in 1% agarose (1x TAE). Once purity was confirmed the concentration was then determined using a spectrophotometer. Typically stationary phase samples yielded less RNA, however still sufficient quantities for analysis.

### **2.5.5 Purification of RNA by DNase treatment**

Despite Dnase I treatment of the RNA during the extraction process a second treatment was performed to remove any residual DNA contamination, with TURBO DNA-free (Ambion DNA-free kit). TURBO DNase has a markedly higher affinity for DNA than conventional DnaseI and thus is more effective in removing trace quantities of DNA. For routine DNase treatment, 1  $\mu$ l TURBO DNase (2U) is added to up to 10 g of RNA, with 1/10 volume of 10X TURBO DNase buffer, in a 50  $\mu$ l volume. This mix was incubated at 37°C for 20-30 min and re-suspended in DNase Inactivation Reagent (typically 1/10 volume). This was subsequently incubated for 5 min at room temperature with occasional mixing. The DNase Inactivation Reagent is pelleted by centrifugation at 13,000 rpm for 2 min and the RNA containing supernatant transferred to a sterile tube.

### **2.5.6. Synthesis of cDNA by Reverse Transcription**

cDNA templates were synthesized by random priming 1 µg of RNA in a 20 µl reaction using the GoScript Reverse Transcription System (Promega), to be used in qPCR reactions. Briefly, the RNA and random primers (0.5 µg/reaction) were mixed in a final volume of 5 µl, and thermally denatured at 70 °C for 5min, then chilled on ice for at least 5 min. A reverse transcription reaction mixture was assembled on ice to contain 1 µl GoScript Reverse Transcriptase, 1 µl PCR nucleotide mix (final concentration 0.5 mM each dNTP), 1.2 l MgCl<sub>2</sub> (final concentration 1.5 mM), 4 l GoScript 5X reaction buffer, and Nuclease free water to a final volume of 15 µl. The 5 µl RNA/primer mix was then added to the reaction mixture to make a final volume of 20 µl. The mixture was incubated at 25°C for 5 min as an initial annealing step. This was followed by first strand synthesis with an extension temperature of 42°C for 1hr. The mixture was then heated to 70°C for 15min to thermally inactivate the reverse transcriptase, prior to use in amplification reactions.

## **2.6 Manipulation of DNA *in vitro***

### **2.6.1 Restriction endonuclease cleavage of DNA**

Throughout this study restriction endonucleases were used to analyse the physical structure of DNA and as a tool in the construction of hybrid promoters. Typically 0.5-2.0 µg of plasmid or linear DNA was cut with 1-2 U of restriction enzyme in a 20 µl volume containing the appropriate reaction buffer. For double digests involving



simultaneous cleavage of DNA by two endonucleases, a suitable buffer was chosen in which both enzymes had >75% activity according to the manufacturers guidelines. For double digests where no compatible buffer was available, sequential digests were performed in suitable buffers with the purification of the DNA between digestions. Reactions were incubated at the recommended temperature for 1-2 hr. Following digestion the reactions were usually heated to >60C for 20 min to heat kill (denature) the enzyme.

### **2.6.2 Phosphatase treatment of restriction-endonuclease cleaved DNA**

Cleavage of DNA by restriction enzymes leaves a 5'-protruding and 3'-recessive "sticky" ends or 5'-blunt ends phosphoryl group on the DNA. These phosphoryl termini are required for formation of phosphodiester bonds between ends of DNA. Shrimp alkaline phosphatase (SAP; Roche) has become an important tool in molecular biology laboratories by removing these 5'-phosphoryl groups (dephosphorylation) which prevents the DNA from ligating (the 5' end attaching to the 3' end), thereby keeping DNA molecules linear until the next step of the process for which they are being prepared. This is useful for blunt ended cloning or cloning into a single restriction site by decreasing the empty vector background contamination. Typically, DNA sufficient for one ligation reaction was incubated with 2 U SAP in a 20 µl volume in 1 x dephosphorylation buffer (supplied with the enzyme) and incubated at 37°C. SAP was irreversibly inactivated by heat treatment for 15mins at 65°C. DNA was recovered at described in 2.5.3.

### **2.6.3. Ligation of DNA molecules**

Bacteriophage T4 DNA ligase (Roche) was used to clone DNA inserts in to the plasmid vectors. T4 DNA ligase catalyzes the ATP-dependent formation of phosphodiester bonds between DNA molecules involving the 3'-hydroxy and the 5'-phosphate termini. It also catalyzes the covalent joining of two segments to one uninterrupted strand in a DNA duplex, provided that no nucleotides are missing at the junction (repair reaction). For its catalytic activity the enzyme requires the presence of ATP and  $Mg^{++}$ . Reactions were performed by incubating an estimated molar ratio of purified vector:insert DNA in a 30  $\mu$ l volume with 2  $\mu$ l of T4 ligase. While it is recommended that a 1:3 (vector:insert) ratio is used for sticky ended and a 1:5 ratio required for blunt ends, typically a series of ratios from 1:3 to 1:50 was used to ensure success. This series included negative controls in which the reaction mixture contained dephosphorylated vector DNA with fragment (a control for the efficiency of phosphorylation), and fragment DNA without vector. The mixtures were incubated for 16 hr at 4°C for sticky end ligations and 16° C for blunt end ligations. A standard *E.coli* cloning strain (usually DH5 $\alpha$ ) was transformed with 10  $\mu$ l of the ligation reaction by the calcium chloride method.

### **2.6.4. Amplification of DNA by PCR**

The polymerase chain reaction (PCR) was used for the *in vitro* amplification of DNA products, for preparation/confirmation of fragments during cloning strategies and generating probes for electrophoretic mobility shift assays. The PCR method is based

on the ability of a thermostable DNA polymerase to amplify DNA, primed from oligonucleotides annealed to denatured single-stranded templates (Saiki et al., 1988). The procedure involves the initial thermal denaturation of double stranded template to separate the strands from each other. Because DNA polymerase can add a nucleotide only onto a preexisting 3'-OH group, it needs a primer to which it can add the first nucleotide. These oligonucleotides hybridise specifically to complementary sequences on opposite strands of the DNA flanking the region of interest. A thermostable DNA polymerase, in the presence of dNTPs and Mg<sup>2+</sup>, carries out second-strand synthesis from the annealed primers. The denaturation, annealing and extensions steps are repeated for 20-35 cycles with each newly-synthesized strand serving as template in further cycles of annealing and extension, resulting in exponential amplification of the DNA region.

#### **2.6.4.1 Amplification of DNA using Phusion Polymerase**

Phusion DNA polymerase is a *Pyrococcus*-like enzyme fused to a unique dsDNA-binding domain, which increases the affinity of the polymerase for double-stranded DNA, allowing incorporation of more nucleotides per binding event and decreasing the number of binding events required for elongation. The processivity of Phusion DNA Polymerases is approximately 10-fold greater than that of *Pyrococcus furiosus* (Pfu) DNA polymerase and twice that of *Thermus aquaticus* (Taq) DNA polymerase. This increased processivity results in shorter extension times and more robust amplification with minimal optimization. The error rate is approximately 50-fold lower than that of Taq polymerase, and 6-fold lower than that of Pfu polymerase. Calculation of

annealing temperatures for Phusion PCR does not follow traditional formula methods (Deighan, 2001) but instead can be determined using the Finnzyme online calculator. PCR reactions were carried out by mixing 0.5 U Phusion polymerase, 10  $\mu$ l of 5X Phusion buffer, 200  $\mu$ M dNTPs, 0.5  $\mu$ M of each primer and 10-50 ng DNA template, with ddH<sub>2</sub>O added to a final volume of 50  $\mu$ l in a 500  $\mu$ l thin-walled PCR tube (Sarstedt). These tubes were placed in a Peltier Thermal Cycler, model PTC-200 (MJ research) and amplification run as follows; Initial denaturation 98 °C for 3 mins, followed by 34 cycles (98°C for 10 sec, annealing temperature for 30 sec, 72°C for 15-30 sec/kb) and a final extension of 10 min at 72°C.

DNA sequences amplified used templates of purified chromosomal or plasmid DNA, but also direct amplification from cultures and single colonies were performed. In the case of a single colony, a scraping using a sterile pipette tip, was re-suspended in 20  $\mu$ l ddH<sub>2</sub>O, and 2  $\mu$ l of this used in the PCR reaction. Turbid cultures were diluted 1:1000 and 2  $\mu$ l of this added to the PCR reaction. In the case of some strains such as *hms* mutants that can form mucoid colonies, the re-suspended cultures can be boiled for 5 mins prior to the PCR reaction to aid in denaturation.

#### **2.6.5. Agarose gel electrophoresis**

Electrophoresis through agarose gels was routinely used in this study to separate DNA and RNA molecules for size and concentration analysis and for the purification of DNA fragments. The gel was prepared using electrophoresis grade agarose (Roche Molecular Biochemicals) in 1 x TAE to concentration of between 0.7% (for good separation of large 5–10 kb DNA fragments) and 2% (for good resolution of small 0.2–

1kb fragments). Ethidium bromide, which intercalates into DNA, was added to a concentration of 1 ml just prior to pouring into a cast. Samples were mixed with 6X loading dye and loaded alongside a known DNA size ladder. Gels were typically run at 50-100 Volts in 1 x TAE buffer and visualized as fluorescent bands under UV.

### **2.6.6 Site-directed mutagenesis**

Site-directed mutagenesis was performed using the QuickChange II SDM kit according to the manufacturer's recommendations (Stragene). The QuickChange II kit is used to exchange DNA sequences and to delete or insert single or multiple nucleotide base pairs into DNA targets contained on a small plasmid (pJET). Primers, complementary to the region of the vector to be mutated, contained 15-20 bp of homology on each side of the non-homologous nucleotides. 10-50 ng of plasmid DNA was used as a substrate with 125 ng of each oligonucleotide. A standard annealing temperature of 55 °C was used for all primers; if SDM was unsuccessful this was lower to 53 °C. Following amplification, the product was digested for 2 hr with the restriction enzyme DpnI to remove the methylated parental DNA template. Only newly synthesized DNA containing the required mutation should remain, which is then transformed in to an appropriate *E.coli* strain. Typically 1 µl of the 30 µl reaction was transformed by the calcium chloride method.

### **2.6.7 Quantitative PCR**

PCR reactions were carried out in duplicate with primer set on an ABI 7500 Sequence Detection System (Applied Biosystems) using FastStart SYBR Green Master with

ROX (Roche). SYBR Green Master is a ready to use, 2X concentrated master mix that contains all the reagents (except primers and template) needed for running quantitative, real-time DNA detection assays. It contains FastStart Taq DNA Polymerase, Reaction Buffer, Nucleotides (dATP, dCTP, dGTP, dUTP), and SYBR Green I. SYBR Green I is a DNA double-strand-specific dye. During each phase of DNA synthesis, the SYBR Green I dye, which is included in the reaction mix, binds to the amplified PCR products; the amplicon can be detected by its fluorescence. A master mix for each primer set was prepared on ice, and avoiding direct sunlight, which contained 25 l (1X) FastStart SYBR Green Master, 300 nM forward primer, 300 nM reverse primer and PCR-grade water to the volume of 45 µl. To this mix, 5 µl cDNA (1 ng) was added and incubated at 95°C for 10 min (to activate FastStart Taq polymerase) before 40 cycles of (95°C for 15 sec, 60°C for 60 sec, 72°C for 30sec). Standard curves were generated for each primer set using five serial 10-fold dilutions of chromosomal DNA, and included in every run. As an internal control expression of a gene not regulated by H-NS, VirB, or temperature (*rho*) was also monitored using primer set *rho*.RT.fw and *rho*.RT.rev (Table 2.3) All experimental data is expressed relative to *rho* expression.

### **2.6.8 Allelic replacement using lambda Red**

All knockout mutations and chromosome integrations were made by homologous recombination using the lambda Red recombination system of Datsenko and Wanner. The lambda Red recombination system facilitates integration into the chromosome of PCR products containing 40 bp of DNA sequence homology at both the 5' and 3' ends of the region of the chromosome where the PCR product is to be integrated. To

generate a *proU::tet* mutant the *tetRA* resistance cassette was PCR-amplified with primer set *tet.40 bp proU fw* and *rev* and spin column purified using HiYield PCR DNA Fragment Extraction Kit (RBC Bioscience). It was then transformed into electrocompetent *S. flexneri* BS184 containing the Red helper plasmid as previously described. The structure of the *proU::tet* lesion was confirmed by PCR and DNA sequencing (GATC Biotech), and then used as a template for insertion of the modified *icsB-proU-gfp-cat* promoters using primer set *icsB.fw.40 bp proU* and *cat.rev.40 bp proU* or *icsB.rev.40 bp proU and cat.rev.40 bp proU*.

## **2.7 Transformation of bacterial strains with plasmid DNA**

### **2.7.1 Preparation of calcium chloride competent cells**

An overnight culture of the strain to be made competent for transformation was used to inoculate 100 ml of LB broth and grown to an OD<sub>600</sub> 0.2-0.4. The cells were pelleted by centrifugation at 4000 g for 10 mins and the bacterial pellet resuspended in 20ml of ice-cold CaCl<sub>2</sub> solution (100mM CaCl<sub>2</sub>, 10% glycerol v/v). This step was repeated, after which cells were resuspended in 2ml of cold CaCl<sub>2</sub> solution, Cells were subsequently incubated on ice for between 1-4 hr, then distributed into 200 µl aliquoted and stored at -70°C.

### **2.7.2 Transformation of calcium chloride competent cells**

DNA to be transformed (10ng - 100ng in a volume not exceeding 10 µl) was added to a sterile 2 ml Eppendorf tube containing competent cells (100-200 µl) and incubated on

ice for 30 min, thereby allowing the DNA to contact the bacterial surface. Tubes were then transferred to a 42 °C water bath for 2 min before returning immediately to ice for a further 2 min. This heat-shock treatment allows uptake of the plasmid DNA through the CaCl<sub>2</sub>-induced competent membrane. L broth was added to a final volume of 1 ml and the cultures incubated at 37 °C for 1 hr with shaking to allow phenotypic expression of the plasmid borne antibiotic resistant marker. Subsequently, 10 µl and 100 µl samples of the transformation mix were plated onto the appropriate selection plates. As a negative control, cells to which no DNA had been added (but were otherwise treated the same) were plated to screen for any contamination. Following overnight incubation at 37 °C, single colony transformants were purified on fresh selective agar plates.

### **2.7.3. Preparation of electro-competent cells**

Typically 1 ml of an overnight culture of the strain to be made electrocompetent for transformation was used to inoculate 100 ml of LB and grown to an OD<sub>600</sub> 0.2-0.4. The bacterial culture was chilled on ice for 15 min after which the cells were pelleted by centrifugation at 3,000 rpm for 10 min and the pellet resuspended in 10 mls of ice cold sterile ddH<sub>2</sub>O. After 30 min incubation on ice the centrifugation step was repeated and the cells resuspended in 5 ml ice cold sterile ddH<sub>2</sub>O. After a final centrifugation step the cells were resuspended in 2 ml of cold 10% (v.v) glycerol. The electrocompetent cells were stored at -70°C in 50 µl aliquots.



#### **2.7.4 Transformation of electro-competent cells**

DNA (100ng – 2 µg) to be electroporated, in a volume not exceeding 5 µl, was added to the 50 µl aliquot of electrocompetent cells and incubated on ice for 2 mins. The mixture was transferred to a pre-chilled electroporation cuvette (EquiBio) with a gap width of 2 mm. The cuvette was then left on ice for 5 min and then placed in the Gene Pulser chamber (Bio-Rad). A pulse, typically 4-5 msec duration, was delivered with the Gene Pulser set to deliver 2.5 kV from the 25 F capacitor and a Pulse Controller resistance setting of 200 ohms. 1 ml of pre-warmed LB broth was added to the cuvette and the contents transferred to a sterile tube and incubated at 37 °C with aeration and shaking for 1 hr. 10 µl and 100 µl of the transformation mix were plated onto appropriate selection plates. Cells to which no DNA was added, but had been treated in the same manner, were also plated to serve as a control for contamination. Following overnight incubation at 37 °C, transformants were single colony purified.

### **2.8 Assays based on spectrophotometry**

#### **2.8.1 Monitoring bacterial growth**

The growth of bacterial cultures was monitored by measuring the optical density of the culture at a wavelength of 600 nm ( $OD_{600}$ ). This was done by a spectrophotometer (Genesys 10uv; ThermoSpectronic) which takes measurements of the amount of light passing through the tube, or conversely the amount of light absorbed. For routine measurements, the bacterial cell fraction assayed was 1ml. For dilute cultures ( $OD_{600} < 0.6$ ) neat samples of cultures were assayed directly, however for *E. coli*, the relationship

between OD<sub>600</sub> and cell density is typically approximately linear up to OD<sub>600</sub> of around 0.6. Therefore more turbid cultures (OD<sub>600</sub> > 0.6) were diluted in sterile broth before the assay. It should be noted, that these are not direct measurements of bacterial numbers, but an indirect measurement of cell biomass that includes both living and dead cells.

### 2.8.2 Assay of $\beta$ -galactosidase activity

Transcription of the *mxiC::lacZ* fusion was measured in overnight cultures of wild type, *hns* and *virB* mutant strains grown in LO or LB broth at 37°C.  $\beta$ -galactosidase activity was measured as described by Miller (1992). A sample of 500  $\mu$ l of cells was permeabilised with Z buffer, 100  $\mu$ l chloroform and 50  $\mu$ l 0.1% SDS in a 2ml Eppendorf. The samples were vortexed and 20  $\mu$ l of these permeabilised cells added to 180  $\mu$ l of chromogenic  $\beta$ -galactosidase substrate o-nitrophenyl- $\beta$ -D-galactopyranoside (ONPG; 4mg/ml in Z buffer), in a 96-well flat-bottomed microtiter plates.  $\beta$ -galactosidase activity was monitored kinetically using a mutiscan ascent plate reader (Thermp labsystem). The kinetics of substrate hydrolysis at 37 °C was measured at 30 second intervals after an initial 3 min lag period. The measurements were plotted (OD<sub>414</sub> vs Time) and the slope of the line used to determine the  $\beta$ -galactosidase activity according to the equation;

Slope (OD<sub>414</sub> vs Time) / (OD<sub>600</sub> x volume (ml) of cells used)

Z-Buffer:

60 mM Na<sub>2</sub>HPO<sub>4</sub>·2H<sub>2</sub>O

40 mM NaH<sub>2</sub>PO<sub>4</sub>·2H<sub>2</sub>O

50 mM β-mercapatoethanol

10 mM KCl

1 mM MgSO<sub>4</sub>·7H<sub>2</sub>O, pH 7

### **2.8.3. Determination of nucleic acid concentrations**

The concentration of DNA samples was determined using the NanoDrop device (Thermo Scientific). DNA concentration is measured by the absorbance at 260nm (A<sub>260</sub>) and reported in ng/μl. ddH<sub>2</sub>O the machine which then also measures the absorbance of the buffer to calibrate a base line.

### **2.8.4 Determination of protein concentration by the Bradford assay**

Protein concentration was determined using a Bio-Rad Protein Assay, which is based on the method of Bradford (1976), and measures the differential colour change (shift in absorbance from 465 to 595 nm) of Coomassie Blue G-250 upon protein binding. The concentrations of His-tagged VirB were determined using several serial dilutions of protein, along with a known standard of BSA (bovine serum albumin), and the resulting OD<sub>595</sub> measurements plotted for comparison. On occasion, to verify the concentration determined by the Bradford assay, the serial dilutions of protein of interest, along with a standard protein, were ran through a SDS PAGE gel and stained with Coomassie Blue. Band intensities were quantified by densitometry using ImageJ analysis software.

## 2.9. Flow cytometry

The bacterial culture to be assayed was harvested and immediately fixed in 4% formaldehyde in phosphate-buffered saline (PBS) and then stored at 4°C in the dark. Before analysis samples were diluted to a concentration of approximately  $10^6$  bacteria per ml and then analyzed with an EPICS-XL flow cytometer (Beckham Coulter). Approximately 10,000 bacteria per sample were assayed and the results were expressed as mean channel fluorescence after analysis with EXPO-XL software. Each sample was measured in duplicate and the mean values were determined from the results of at least three independent experiments. The following detection settings were used;

	Volts	Gain
Forward scatter	55	100
Side scatter	400	500
F1	650	7.5

## 2.10 Proteomic analysis

### 2.10.1 Preparation of total cellular extracts

Crude cell protein extracts were prepared for SDS-PAGE and western immunoblotting as described below. A sufficient volume of cells was harvested to provide 50  $\mu$ l samples after equalization to an OD<sub>600</sub> of 0.2. Fifty microliters of a 2X Laemmli buffer

(Sigma) was added to each sample, mixed and boiled for 10 min. Laemmli 2X buffer contains 4% SDS, 20% glycerol (v/v), 10% 2-mercaptoethanol (v/v), 0.0004% bromophenol blue (w/v) and 0.125 M Tris-HCl, pH approx. 6.8. Samples were stored at -20°C or 10 µl loaded immediately and separated through a 12% SDS-PAGE gel.

### **2.10.2 Separation of proteins by SDS-PAGE.**

Purified proteins and purified proteins from total cell extracts were separated by discontinuous electrophoresis through a 12 % polyacrylamide gel as described in Sambrook et al, (1989). The discontinuous buffer systems employ different buffer ions and pH in the gel. Samples are loaded onto a non-restrictive large pore gel, called a stacking gel, which contains a low percentage of acrylamide (5%), which focuses the protein into narrow bands. This stacking gel overlays a smaller pore resolving gel with a higher percentage acrylamide matrix (12%) in which the proteins will migrate according to their molecular mass. Both gels are prepared with 0.1% SDS, in most proteins, the binding of SDS to the polypeptide chain imparts an even distribution of charge per unit mass which leads to identical charge ratios for denatured proteins. This allows the fractionation of proteins by approximate size during electrophoresis and not charge. The electrophoresis apparatus used was a Mini-Protean II ® verticle electrophoresis cell supplied by Bio-Rad and assembled in accordance with the manufacturer's instructions.

Gel plates were cleaned before use and a 12% (w/v) separating gel mix was made by mixing 2 ml Protogel (National Diagnostics), 1.25 ml 1.5 Tris-HCl (pH 8.8), 50 µl 10%

SDS and 1.7 ml ddH<sub>2</sub>O. Prior to pouring, polymerization was catalyzed by the addition of 50 µl 10% (w/v) ammonium persulphate (made fresh) and 5 µl TEMED (N,N,N',N'-tetramethyl-ethylenediamine). The gel mix was then poured between the plates until 1 cm below the tip of the comb and immediately overlaid with isopropanol (100%) to exclude oxygen and eliminate air bubbles. This was allowed to polymerize for 20 min, before inverting the plates to pour off the isopropanol and blotting dry with absorbent paper (Protran). The stacking gel was made by mixing 0.833 ml Protogel (National Diagnostics), 1.25 ml 0.5 M Tris-HCl (pH 6.8), 50 µl 10% (w/v) SDS and 2.7 ml ddH<sub>2</sub>O, 50 µl 10% (w/v) ammonium persulphate and 5 µl of TEMED. Following the insertion of appropriate sized well forming combs the gel was allowed to polymerize for a further 20 min.

The polymerized gels were assembled in the electrophoresis chamber with the wells facing inwards to complete the circuit. The apparatus was filled with 1x Tris-glycine running buffer [25 mM Tris-HCl, 250 mM glycine (pH 8.3) 0.1% (w/v) SDS]. 10 µl of sample was loaded into flushed out wells along with 5 µl Laemmli buffer in empty wells, and a prestained protein marker (NEB) to estimate the molecular mass of proteins. Electrophoresis was performed at 90V for 90 min.

### **2.10.3 Visualisation of proteins in SDS-PAGE gels.**

Gels were stained using Coomassie Brilliant Blue (0.1% (w/v) Coomassie R250, 10% glacial acetic acid, 40% (w/v) methanol). Gels were covered in Coomassie stain and microwaved at high power for 30 sec then placed on a rocking bellydancer for 10 min.

After this, the stain was removed and replaced with destain solution (10% glacial acetic acid, 40% (w/v) methanol in ddH<sub>2</sub>O). This was left with agitation for 1 hr, with repeated solution changes.

#### **2.10.4 Transfer of proteins to nitrocellulose membrane.**

If the gel is to be analysed by western immunoblot, following electrophoresis, the gel is removed from the electrophoresis plates and covered with 0.2 mM Protran nitrocellulose membrane (Schleicher and Schuell). The gel/nitrocellulose was sandwiched between 2 sheets of absorbent paper (Protran) and clean sponges, and placed in a Mini Trans-blot electrophoretic transfer cell (BioRad). Transfer was at 90V for 1 hr in ice cold buffer (25 mM Tris, 192 mM glycine, 20% methanol) ensuring that the gel was between the cathode (-) and the nitrocellulose membrane. Nitrocellulose membranes were then stained with Ponceau (0.2% Ponceau dye, 3% trichloroacetic acid) to check the efficiency of transfer and also to ensure equal loading of samples which was followed by extensive washing with distilled water.

#### **2.10.5 Western Immunoblot analysis**

After the proteins are irreversibly transferred to a nitrocellulose membrane, it can then be incubated with antigen specific primary antibody to detect the proteins of interest. The membrane was blocked on 5 % Marvel (low fat milk powder in PBS) for 2 hr. The membrane was then incubated with shaking for 1 hr with primary polyclonal anti-VirB antiserum (1:500 in 5 % Marvel/PBS). The primary antibody/marvel solution was then decanted into a sterile contained and stored at -20°C for re-use. The membrane was

washed 3 x 10 min with PBS, then incubated for 1 hr with blocking buffer containing HRP (horseradish peroxidase-conjugated)-linked goat anti-rabbit IgG secondary antibody (1:10,000). The blot was washed again for 3 x 10 min periods in PBS before incubation for 5 min with a chemiluminescent Pierce West Pico Super Signal kit. The HRP mediated enzymatic reaction results in a luminescent signal that allows the antigen-antibody complex to be visualized using autoradiography. Detection of emitted light was performed using developer and fixer solutions (Kodak) and Hyperfilm (Amershan Biosciences), typically re-using a 30 sec exposure. In some cases equal protein loading to gels used for immunoblotting was ensured by monitoring the presence of DnaK protein using rabbit monoclonal anti-DnaK antiserum at a dilution of 1:200,000 (Enzo Life Sciences).

#### **2.10.6 Over-expression of VirB protein**

Plasmid pET22b*virB* is a pBluescript plasmid which harbours an N-terminal ten histidine fusion to the *virB* gene cloned downstream of the T7 polymerase promoter. Induction of expression was performed in the strain BL21(DE3), DE3 lysogen contains phage T7 gene 1 (T7 polymerase) upon IPTG induction. This strain is deficient of *lon* and *ompT* proteases and is therefore suitable for expression of non-toxic genes. The plasmid encodes the LacI repressor that completely silences the *lacUV5* promoter prior to induction, giving tight control of *virB* expression. Upon addition of IPTG, the LacI-mediated repression of the promoter is effectively removed and T7 RNA polymerase is produced, which can then initiate selective transcription from the plasmid borne T7 promoter sequence upstream of the *virB* gene.



250 ml of culture was grown at 30°C to mid exponential phase in LB broth, then 1 mM IPTG was added to the culture and growth allowed to continue for a further 2 hr. Cells could then be pelleted by centrifugation and stored at -20°C for further use.

Alternatively, the pellet was resuspended in binding buffer [60 mM imidazole, 500 mM NaCl, 20 mM Tris-HCl (pH7.9), 1 mM PMSF (phenylmethylsulfonyl fluoride)] supplemented with 200 g/ml lysozyme. Complete lysis was then achieved by several bursts of sonication. Cell debris was pelleted by centrifugation for 10 min at 10,000 g and the soluble extract (15 ml) retained.

#### **2.10.7 Protein Purification**

His-tagged VirB protein required for EMSA and DNaseI footprinting was purified from the crude extracts by metal affinity chromatography using His-Bind magnetic beads (Novagen). The His-Bind magnetic beads are coupled to immobilized Ni<sup>2+</sup> ions that can retain the His-tag sequence by electrostatic interaction. Imidazole was then used as a competitor for Ni<sup>2+</sup> binding to elute target protein.

Briefly, the soluble extract was passed over a pre-equilibrated His-Bind 900 column (Novagen). The column was washed with binding buffer and then wash buffer (100 mM imidazole, 0.5 M NaCl, 20 mM Tris-HCl pH 7.9, 1 mM PMSF) to remove proteins weakly associated with the resin. His-tagged VirB was then eluted with 500 mM imidazole and deemed to be 98% pure by SDS-PAGE followed by Coomassie staining. Purified VirB was dialysed in storage buffer (60 mM NaCl, 50 mM Tris-HCl pH7.9, 1

mM EDTA, 0.5 mM DTT, 0.1 mM PMSF, 50% glycerol) and stored in aliquots at -20°C.

## **2.11 EMSA**

Electrophoretic mobility shift assays were used in this study to analyze the interaction of VirB protein with DNA in vitro. EMSA analysis is based on the reduced electrophoretic mobility of nucleic acids bound by protein through polyacrylamide gels, compared with identical free DNA. The method was first described by Garner and Revzin (1981). Approximately 0.5 ng of probe DNA was incubated with increasing concentrations of purified His-tagged VirB, for 30 min at room temperature in a 20 µl reaction cocktail of 20 mM Tris-HCl (pH 7.5), 80 mM NaCl, 1 mM EDTA, 100 ng of bovine serum albumin (BSA), 25 µg/ml poly(dI-dC), 10% glycerol, and 1 mM DTT. The protein-DNA complexes were then resolved by electrophoresis through 6% nondenaturing polyacrylamide gels (Invitrogen), in 0.5% TBE, or 3 h at room temperature. After electrophoresis, the Biotin-labeled DNA was transferred onto a Biotinylated nylon membrane (Pall) and then immobilized by UV cross-linking (GS GeneLinker, BioRad). The DNA was then visualized using the chemiluminescent-nucleic acid detection kit per the instructions of the manufacturer (Pierce).

## **2.12 DNase I footprinting using FAM labelled primers**

The use of fluorescently labeled primers eliminates the need for radioactively labeled nucleotides, as well as slab gel electrophoresis, and takes advantage of commonly available automated fluorescent capillary electrophoresis instruments. With

fluorescently labeled primers and dideoxynucleotide DNA sequencing, we have shown that the terminal base of each digested fragment may be accurately identified with a capillary-based instrument. Polymerase chain reaction (PCR) was performed with a 6FAM-labeled primer to amplify a typical target promoter region. This PCR product was then incubated with a transcriptional activator protein, or bovine serum albumin as a control, and then partially digested with DNase I. Briefly, bait DNA was prepared using primers sets pZec.6FAM.F and pZec.comfirm.R (Table 2.3). The reactions were then conducted in a 15  $\mu$ l reaction volume consisting of 1x DNase buffer (Roche) (40 mM Tris-HCl, 10 mM NaCl, 6 mM MgCl<sub>2</sub>, pH 7.9), 0.001 mM dithiothreitol, 100 ng/ $\mu$ l BSA, 50 nM bait DNA and 50  $\mu$ M VirB-His. This mixture was allowed to equilibrate at 37 °C for 15 mins, then 1  $\mu$ l (0.0015 units) of Dnase I was added and mixed gently, then incubated at 37 °C for 10 min. The digestion reaction was stopped by addition of 2  $\mu$ l EDTA (100 mM) followed by vigorous vortexing and heat denaturation at 95 °C for 10min. Digestion products were desalted using MicroSpin G-50 columns (GE Healthcare) and analyzed on a ABI 3130 Genetic Analyzer along with GeneScan 500-LIZ size standards (Applied Biosystems).

## **Chapter 3**

### **The rational design of a new genetic switch**

### 3.1 Introduction

The creation of new genetic regulatory networks is something routinely predicted in bioinformatics studies of evolution (van Hijum et al., 2009). Gene regulatory circuits in bacteria can be adapted to meet the changing needs of the organism that harbours them. While transcriptional control is mediated through a wide variety of sophisticated processes, some very effective genetic switches operate through relatively simple mechanisms. Some of the most basic involve mechanisms that relieve repression exerted by the nucleoid-associated protein H-NS.

H-NS is a pleiotropic repressor of transcription in Gram-negative bacteria. The removal of H-NS from its target promoters can be achieved via a variety of mechanisms (Stoebel *et al.*, 2008). Transcription factors can cause the displacement of H-NS, either by competitive binding to DNA or by structural remodelling of the local DNA architecture to interfere with H-NS binding. Proteins that oppose the DNA binding activity of H-NS stem from a wide variety of transcription factor families, suggesting antagonism is not likely to utilise highly specific protein-protein interaction between repressor and activator.

The expression of the type III secretion system encoded by genes on the large virulence plasmid of *Shigella flexneri* is a target for H-NS-mediated repression (Porter and Dorman, 1994; Beloin *et al.*, 2003). Antagonism of this repression is thought to have involved the evolution of an elaborate regulatory cascade. Upon receipt of correct

environmental signals, expression of an AraC-like protein VirF results in derepression of an intermediate regulatory gene *virB*. The VirB protein is required to derepress the principal promoters governing the expression of the structural and effector genes encoding the type III secretion system (Tobe *et al.*, 1991). It is thought that derepression of the *virB* gene represents a regulatory checkpoint that must be passed prior to full commitment to expressing the elaborate type III secretion system and its effector proteins.

VirB antagonism of H-NS has been dissected in detail for the de-repression of *icsB*. The mechanism of de-repression involves VirB remodelling the DNA within the H-NS-DNA nucleoprotein complex (Turner and Dorman, 2007). This can be observed in DNase I footprinting assays with the appearance of VirB-induced hypersensitivity bands in regions known to be protected by H-NS. Such hypersensitivity arises when DNA winds around a protein, making particular bases even more susceptible to cleavage by DNase I. At the *icsB* promoter this hypersensitivity originates at an essential nucleation site, a 15-bp asymmetric inverted repeat (5'-GTTTCATcATGAAAT- 3') denoted Box 1/Box 2, which is located immediately upstream on the cusp of the region bound by H-NS. It then extends through the region under H-NS protection for approximately 120 bp. The current model of VirB action involves binding of the protein to the Box1/ Box 2 nucleation site, followed by propagation of a VirB multimeric complex along the DNA with associated wrapping of the DNA duplex by the protein (Figure 3.1). This action is detrimental to the maintenance of the H-NS-DNA repression complex and results in displacement of the H-NS protein, allowing for initiation of transcription by RNA

polymerase. *In vitro* transcription assays show that VirB does not act as a conventional transcription factor. It does not recruit RNA polymerase to the promoter nor does it isomerize the closed *icsB* transcription initiation complex to its open form. Instead, VirB acts solely as an antagonist to the repressive activity of H-NS. Furthermore, VirB belongs to no known family of transcription factors.

Instead VirB resembles DNA binding proteins involved in plasmid partitioning (Beloin *et al.*, 2002; Watannabe *et al.* 1990). There is strong amino acid sequence similarity between VirB and the well-characterised SopB and ParB proteins, expressed by the F factor and the P1 plasmid/bacteriophage, respectively. Secondly, the *cis*-acting Box1/Box2 nucleation site required for VirB to bind and activation the *icsB* virulence gene exhibits a highly similar nucleotide sequence to the *parS* sequences (5'-ATTTCAAgTGAAAT-3') used by the partitioning proteins to bind and polymerise (Figure 1.4).

The *S. flexneri* virulence plasmid has two functional partitioning systems, including *parAB* which consists of the ParB protein, its partner ATPase ParA and the *cis*-acting *parS* site. The *parS* motif is bound by the ParB protein, facilitated by Integration Host Factor (IHF) which induces bends in the DNA up to 180° (Vecchiarelli *et al.*, 2007). Interestingly, IHF is also involved in modulating expression of virulence genes within the entry region (Porter and Dorman, 1997). There is no evidence that VirB plays any direct role in plasmid partitioning in the modern virulence plasmid. Nor is there any evidence for a ParA-like partner for VirB. It is tempting to speculate that the loss of the

gene coding for the putative VirB partner protein combined with the existence of other partitioning systems resulted in the VirB protein becoming redundant in terms of plasmid maintenance. This would have allowed VirB to become available for other tasks within the plasmid.

Alignment of the *virB* gene and VirB protein binding site with the genetic map of its homologue *parB* and its *parS* site suggests that insertion of the *icsB-ipg-ipa-acp* operon disconnected the *virB* gene from its *cis*-acting binding site while simultaneously positioning this *parS*-like site at a region where it can act to antagonise H-NS-mediated silencing of *icsB* gene (Figure 3.1).

Many lines of evidence indicate plasticity in the organisation of the bacterial genome with genes and parts of genes moving within and between genomes (McAdams *et al.*, 2004). Bioinformatic studies suggest that positively acting regulators and their binding sites are evolutionarily uncoupled, allowing target genes to join and leave regulons controlled by the protein as they gain and lose appropriate binding sites (Berg *et al.*, 2004) With this in mind, analysis of VirB suggests that the fortuitous placement of an appropriate *cis*-acting site in the regulatory network of *icsB* virulence gene may have allowed the evolution of a new regulatory switch and a functional reassignment of the VirB protein. These re-arrangements are simple and analogous to modifications that happen routinely in bacteria. But how easily can this proposed evolutionary step occur? Addressing this question could provide some insight into how gene-regulatory circuits have evolved, and may evolve in the future. I speculate that the introduction of the VirB



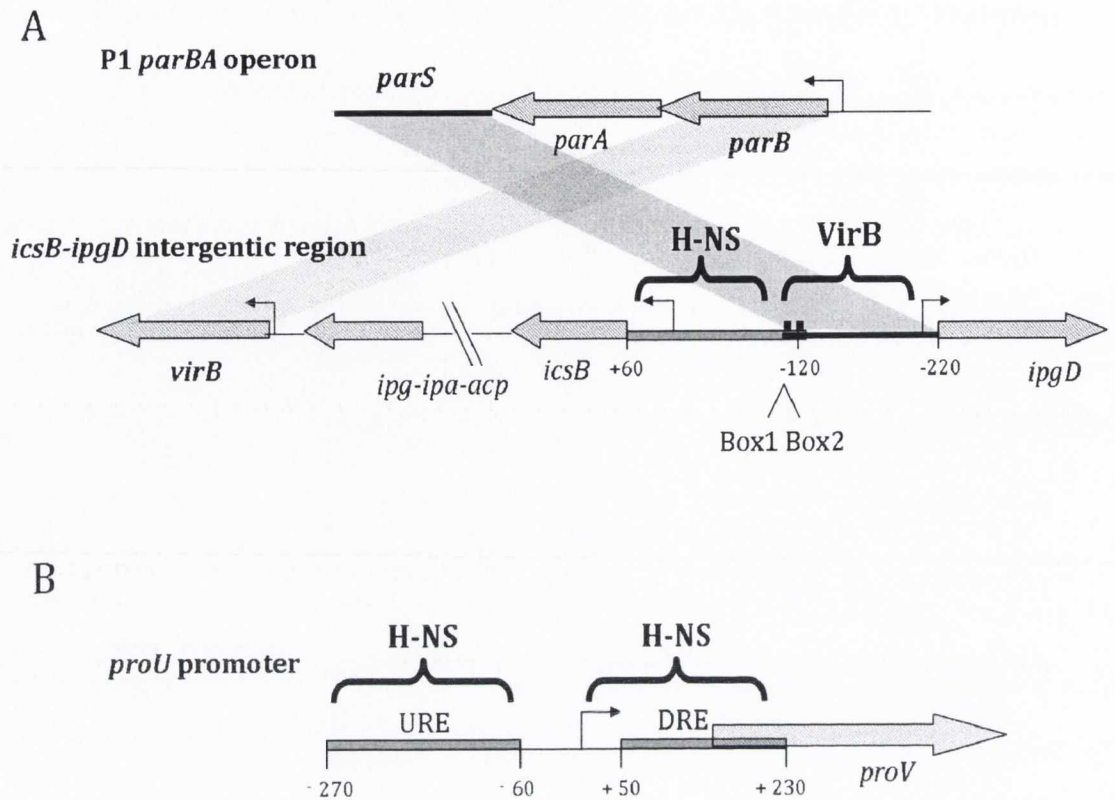


Figure 3.1. (A) The alignment of the regulatory regions of *parAB* operon and the *icsB-*ipgD** intergenic region. The illustration indicates the position of the *parS* binding site in relation to the *parA* and *parB* genes. Below the genetic map of a portion of the entry region showing the relative locations of the *virB* gene and the regulatory sequences of the *icsB-*ipg-ipa-acp** operon. The angled arrows represent promoters. The filled grey box indicates the region to which H-NS binds at *icsB*, and dark black line the region to which VirB binds. Also represented is the essential Box1/Box 2 nucleation site shown as 2 black squares within the VirB site. The grey parallelograms are indicative of regions of homology between the nucleotide sequence of the VirB binding site and *parS* and the amino acid sequence homology between *virB* and *parB* genes. (B) An illustration of the regulatory region of *proU* representing the P2 promoter (angled arrow) and the H-NS binding site, URE and DRE (grey boxes).

binding site from the *icsB* gene into the regulatory regions of an unrelated H-NS repressed promoter would theoretically allow for H-NS antagonism in a VirB dependent manner (Figure 3.1A).

To test this hypothesis the chromosomally-encoded house keeping gene, *proU*, was exploited, which possesses a well-characterised example of an H-NS repressed promoter.

The *proU* operon encodes an uptake system for proline and the osmoprotectant glycine-betaine and is upregulated in the presence of hyperosmotic shock (Gowrishankar and Manna, 1994). The operon is controlled by a promoter (P2) whose transcription start site is approximately 60 nucleotides upstream of the initiation codon of the first structural gene *proV* and is recognized by the  $\delta^{70}$  - RNA polymerase holoenzyme (E  $\delta^{70}$ ). Another promoter, P1 located 250 nucleotides upstream of *proV* has been shown to be RpoS ( $\delta^S$ ) dependent. However, its role is uncertain as deletion of P1 or mutation of *rpoS* does not affect *proU* expression *in vivo* (Gowrishankar and Manna, 1994).

The promoter region contains two regulatory domains, the upstream regulatory element (URE) and the downstream regulatory element (DRE), both of which are known H-NS binding sites. These sites are high-affinity binding sites and include the archetype from which the H-NS binding site consensus sequence was identified (Figure 3.1B). Under conditions of low osmolarity H-NS is bound in these regions and represses

transcription. Upon addition of salt, the promoter is de-repressed as H-NS becomes displaced allowing for transcription, with full expression occurring at 300 mM NaCl (Gowrishankar and Manna, 1994). The transcription silencing mechanism is thought to involve the creation of a repressive complex formed by the bridging of H-NS bound at the URE and DRE (Nagarajavel *et al.*, 2007). However, the DRE alone has been shown to be necessary and sufficient to maintain repression under low salt conditions (Nagarajavel *et al.*, 2007).

Hypo-osmotic conditions can be maintained easily in the laboratory by altering the salt concentration of LB medium. With the exception of HU, which does not bind the promoter but affects P2 indirectly, (Rajkumari and Gowrishankar, 1996; Manna and Gowrishankar, 1994) no trans-acting factors are known to be involved in the de-repression of *proU* promoter. Because of the relatively straightforward and easily manipulated regulation of *proU*, this system was an ideal model for modification to test the ease with which a heterologous promoter, known to be silenced by H-NS, could be made dependent upon VirB for the relief of silencing.

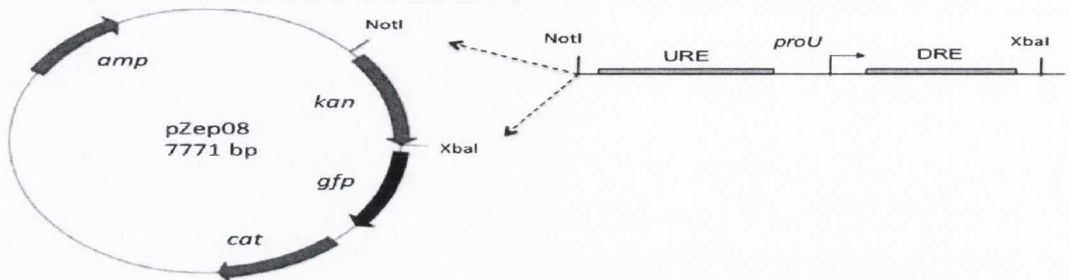
In this study the previously-characterised binding site for VirB at the *icsB* promoter was introduced into the promoter region of *proU* at different locations and the synthetic promoters were monitored for VirB-dependent de-repression under normally repressive growth conditions.

## 3.2. Results

### 3.2.1. Creation of *proU-gfp* expression vectors.

To monitor *proU* expression, the pZep08 green fluorescent protein (GFP) plasmid based system was used (Table 2.2). The *proU* promoter, including the surrounding regulatory region encompassing the H-NS sites URE and DRE, was amplified with oligonucleotides *proU*fw-270.NotI (a forward primer annealing to position -270 and tagged with restriction site NotI) and *proU*rev+230.XbaI (a reverse primer annealing to position +270 and tagged with restriction site XbaI) (Table 2.3). This PCR product was digested with the respective enzymes. This cut product was then inserted into the NotI- and XbaI-cut pZep08 plasmid, placing the *gfp* gene under direct control of the *proU* promoter (Figure 3.2). Expression from this promoter region could then be monitored by flow cytometry. The construct was transformed into wild type BS184 (Table 1.1), *hns* BS185 (Table 1.1) and *virB* CJD1018 (Table 1.1) all previously published and confirmed strains. The *proU-gfp* hybrid was tested at increasing salt concentrations, 0 mM, 172 mM and 300 mM NaCl. The results (Figure 3.2) confirmed the characteristic osmosensitive promoter response. Under conditions of low osmolarity (0 mM) the promoter remained repressed throughout growth. When monitored at 172 mM NaCl the promoter activity increased 60 fold as compared to the repressed promoter. Activation occurred almost immediately with a 20-fold increase observed after the first hour of growth. Full activation was observed at 300 mM NaCl, where the promoter was up-regulated 120 fold in comparison to its repressed state, as

A



B

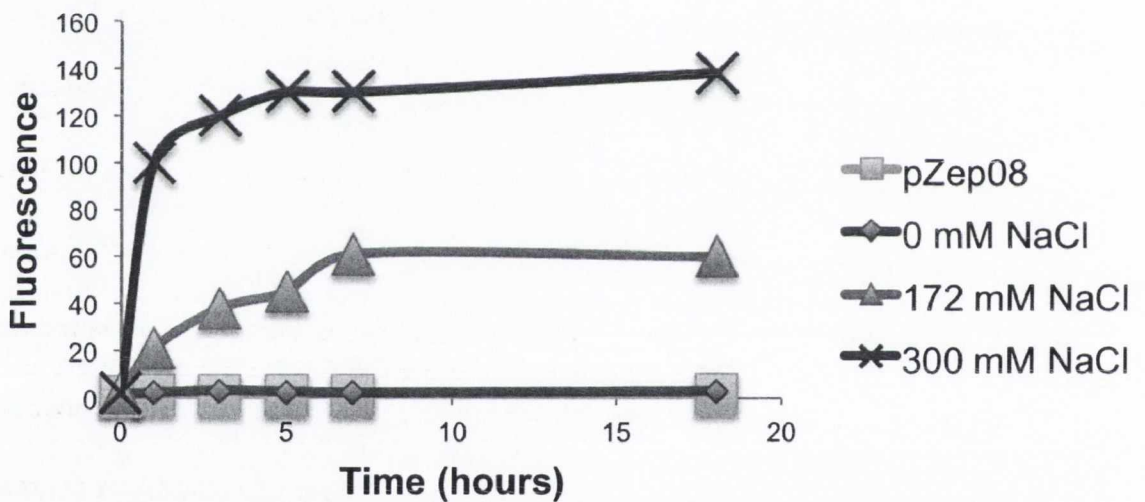


Figure. 3.2. Creation of a *proU-gfp* fusion. (A) Strategy used for the construction of *proU-gfp* on the 7771-bp plasmid pZep08. Promoter was amplified from the *E. coli* chromosome and inserted into pZep08 in place of the kanamycin resistance cassette (grey arrow labelled *kan*) by digestion with restriction enzymes NotI and XbaI. Also present on the plasmid is the *cat* gene conferring chloramphenicol resistance and the *bla* gene conferring ampicillin resistance. (B) Expression data confirming *proU-gfp* osmo-regulation in increasing increments of NaCl, with the empty vector pZep08 acting as a negative control.

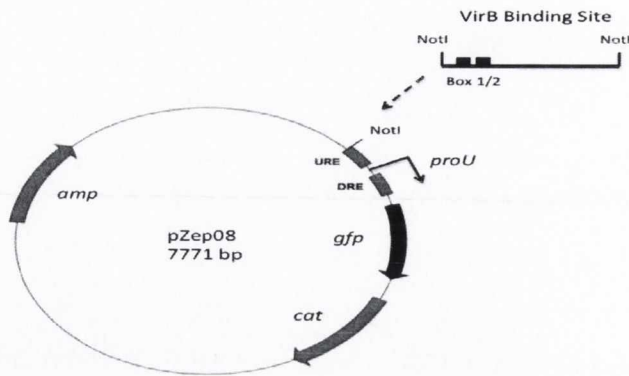
previously described (Gowrishankar and Manna, 1996). These results (Figure 3.2) confirm that the plasmid *gfp* system is a reliable method of monitoring expression.

### **3.2.2. Creation of VirB dependence at the *proU* promoter.**

#### **3.2.2.1. Insertion of the VirB binding site from the *icsB* promoter.**

The VirB binding site in its native location at *icsB* is found upstream of the promoter, bordering the region that is bound by H-NS (Turner and Dorman, 2007). To mimic this arrangement, the VirB binding site, including the essential Box 1/Box 2 motif, was inserted immediately upstream of the URE at the *proU* promoter. The region to which VirB binds at the *icsB* promoter (as determined by DNase I footprinting) was amplified using NotI-tagged oligonucleotides pairing *icsB.fw.NotI* and *icsB.rev.NotI* (Table 2.3). These primers amplified the DNA region -80 (with respect to the transcription start site of *icsB*) to -220. This product was then inserted into the NotI digested, previously constructed, *proU-gfp* vector (3.2.1). This produced a hybrid *icsB-proU* promoter with the *icsB* region inserted upstream of the URE (at position -270 with respect to the transcription start site of *proU*) in both the forward and reverse orientation. The two forms of the modified promoter were produced with approximately equal efficiencies suggesting that the two insertions were benign from the perspective of plasmid vector stability. A PCR based assay with primers specific for each orientation of the inserted DNA, was followed by DNA sequencing to confirm the orientation of the VirB binding

A



B

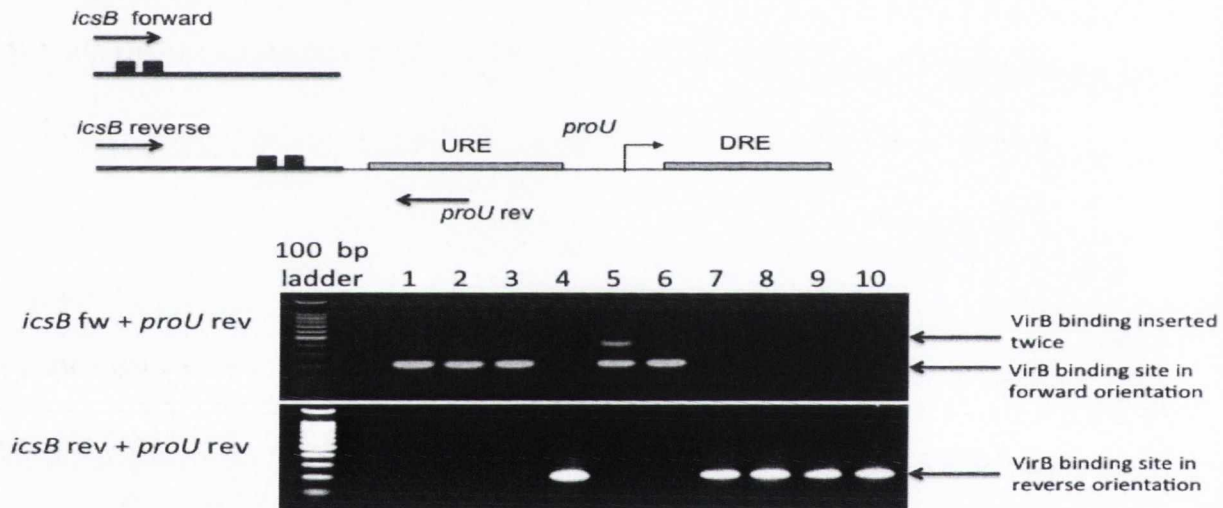


Figure 3.3. Insertion of VirB binding site. (a) The region of DNA bound by VirB at the *icsB* promoter was amplified and inserted into a single digest immediately upstream of the URE on the plasmid *proU-gfp*. (b) PCR-based assay for determining orientation of VirB binding site by coupling of primers specific to the VirB binding site in the forward orientation (*icsB* forward) or reverse orientation (*icsB* reverse) and a reverse primer with the *proU* promoter region (*proU rev*). Screening of 10 plasmids revealed 5 (1, 2, 3, 5, 6) in the forward orientation and 5 (4, 7, 8, 9, 10) in the reverse orientation. In one screen (sample 5) because of only a single restriction digest used the binding site inserted twice producing a PCR product of around 400 bp.

site (Figure 3.3). These constructs were used to address the possibility that VirB binding site directionality might influence derepression.

### 3.2.2.2. Purification of VirB protein and EMSA of hybrid promoters

In order to confirm that the *icsB-proU* hybrid promoters now displayed affinity for the VirB protein, C-terminally His-tagged VirB protein was over-expressed from the plasmid pET22bVirB (Table 1.2) and purified by nickel affinity chromatography. The protein bound the column, as little protein was observed in the flow-through fraction, and eluted in 3 separate fractions with 500 mM imidazole which out-competes the C-terminal histidines for Ni<sup>2+</sup>. The protein was judged to be >98% pure by SDS PAGE and densitometric analysis (Figure 3.4A).

Electrophoretic mobility shift assays were used to assess the ability of the VirB protein to bind specifically to the *icsB-proU* promoter DNA *in vitro*. The results showed shifting of the DNA probe with increasing concentrations of VirB protein (Figure 3.4B). The probes were amplified using a Biotin labelled forward primer *proU.fw.BIO* coupled with reverse primer *proU.rev* (Table 2.3). The DNA templates used in the PCR reactions were either the *icsB-proU* hybrid promoter (on pZep) or the *icsB-proU* hybrid promoter with the essential box1/box2 has deleted, therefore abolishing VirB binding at this region (Turner and Dorman, 2007) (2.6.6). Over 90 % of the DNA was shifted using a concentration of 20  $\mu$ M VirB. By comparison, when using the mutated VirB binding site (Turner and Dorman, 2007) as the DNA probe a concentration of



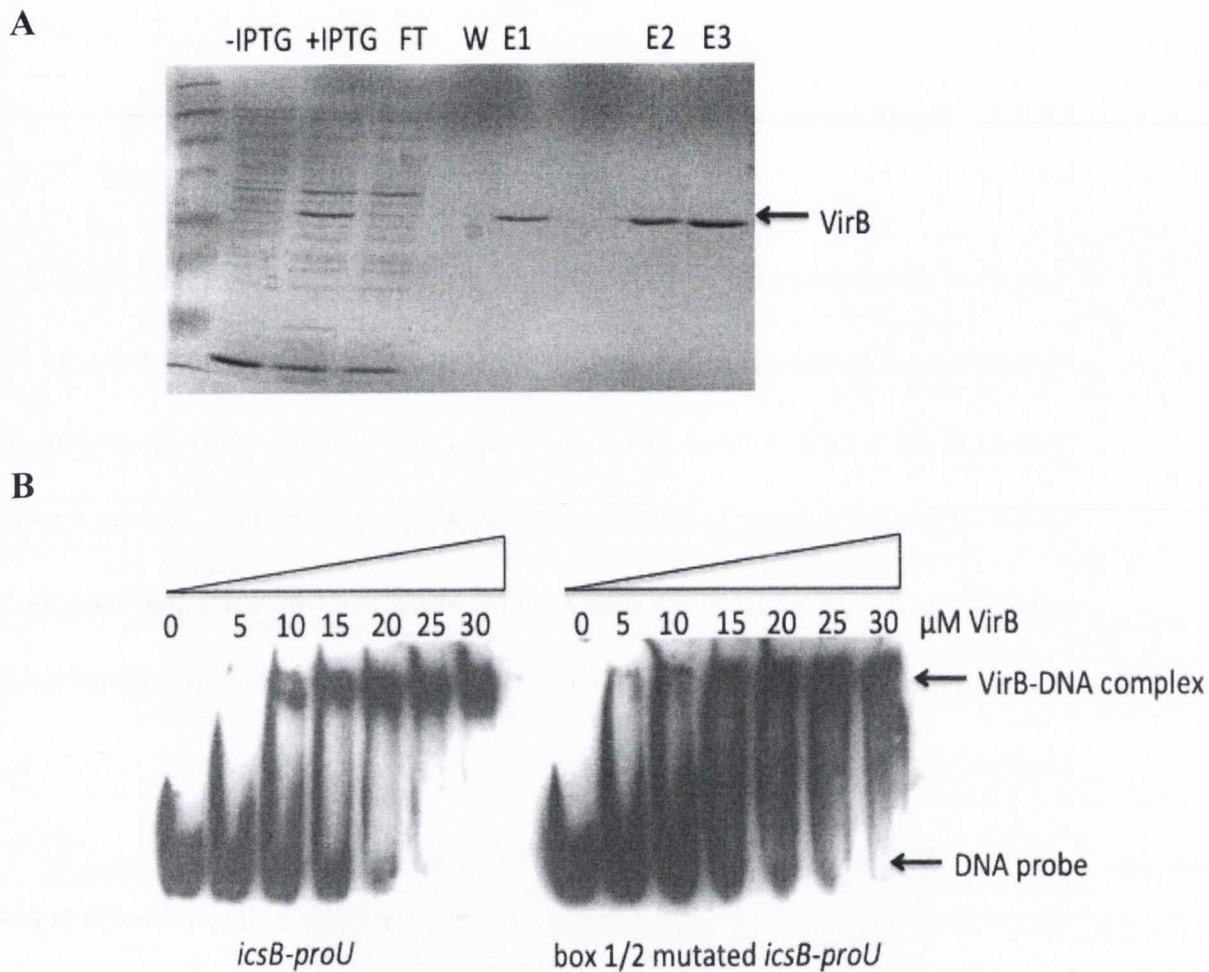


Figure 3.4. Purification and binding affinity of the VirB protein. (A) SDS PAGE analysis of overexpression and purification of VirB. Gel shows cells in which the plasmid pET22bvirB were uninduced (-IPTG), induced (+IPTG), the flow through fraction (FT), the wash flow through fraction (W) and three elution fractions (E1, E2, E3), which isolate the His-tagged VirB protein. (B) Electrophoretic mobility shift assay displaying binding of the VirB protein to the modified *proU* promoter (*icsB-proU*) and the *proU* promoter including a deleted Box1/2 VirB binding site (Box 1/2 mutated *icsB-proU*), with increasing concentrations of VirB protein. DNA probes were amplified with Biotin labelled primers (Table 2.3) for detection (Section 2.11).

30  $\mu$ M was required to produce a shift. The smearing of the DNA in the lanes with a mutated binding site probe is indicative of a DNA-protein complex bound with poor affinity which dissociate from the complex during electrophoresis (Kaer et al., 2008).

### 3.2.2.3. VirB protein levels and VirB function at low osmolarity.

VirB protein levels are reduced in the low osmolarity growth conditions under which these experiments were to be conducted (Porter and Dorman, 1997). To verify this, VirB protein levels were monitored by Western blots in bacteria exposed to increasing concentrations of NaCl. The results showed a clear increase in protein levels with increasing osmolarity (Figure 3.5 A). Westerns blots were then used to monitor variations in VirB protein concentrations in wild type and mutant strains grown in 0 mM NaCl conditions or in standard 172 mM NaCl LB. As expected, VirB protein levels increased in an *hns* mutant BS185 (the *virB* promoter is repressed by H-NS) and no VirB was detected in a *virB* null mutant CJD1018 (Table 2.1) (Figure 3.5 A). To determine whether quantities of VirB were sufficient to exert an effect, expression of a *lacZ* reporter fusion to *mxiC*, a VirB-regulated gene encoding a structural component of the type III secretion system, was analysed. The wild type BS184 (Table 2.1) strain contains this *lacZ* fusion on the large virulence plasmid. Beta-galactosidase activities were assessed at both osmolarities in wild type, *virB* and *hns* strains. The results showed that although VirB protein levels were depleted at 0 M NaCl compared to standard NaCl (172 mM) in wild type, there was sufficient protein to de-repress *mxiC*

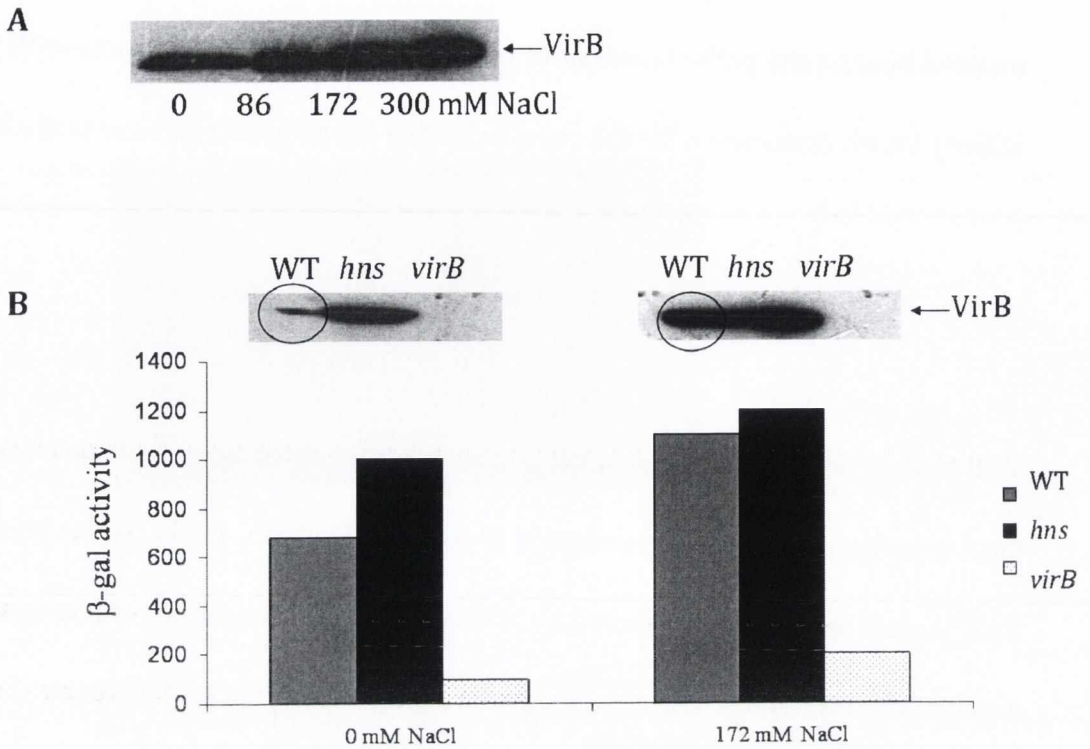


Figure 3.5. VirB protein levels are reduced in low salt but are still sufficient to derepress promoters. (A) Western blot analysis of VirB protein levels in bacteria grown with increasing increments of salt. (B) Western blot analysis of VirB levels coupled with expression data from a *mxiC-lacZ* fusion (a gene expressed from a VirB-dependent promoter) in wild type, *hns*, and *virB* bacteria grown in the presence of 0 mM or 172 mM NaCl.

transcription. (Figure 3.5 B). This confirmed the presence of a sufficient concentration of VirB protein in the cell to exert its de-repressive effects.

#### **3.2.2.4 Insertion of a binding site for VirB confers VirB-dependent regulation at modified *proU* promoters.**

Expression of *icsB-proU-gfp* was monitored in wild type, *virB* mutant and *hns* mutant backgrounds, under growth conditions that were normally non-permissive for *proU* transcription, i.e. LB without NaCl added. Control constructs that completely lacked a VirB binding site (i.e. the *proU* promoter) were also monitored to determine the level of native *proU* expression under non-permissive conditions (Figure 3.6 A). Further controls that included constructs with a modified VirB binding site where the essential Box 1/Box 2 motif has been deleted were also monitored to examine *proU* expression in constructs where the antagonist could not bind (Turner and Dorman, 2007).

Expression of *proU-gfp* in the control constructs, either completely lacking the binding site for VirB (Figure 3.6 A) or containing the version with a mutated Box 1/Box 2 motif (Figure 3.6 D & E), was repressed, and repression was maintained regardless of the presence or absence of VirB protein. Repression was only alleviated when the *hns* gene was inactivated. However, the modified *proU* promoters that contained a functional binding site for VirB were derepressed in the presence of a functioning *hns* gene (Figure 3.6 B and C). These *proU-gfp* expression

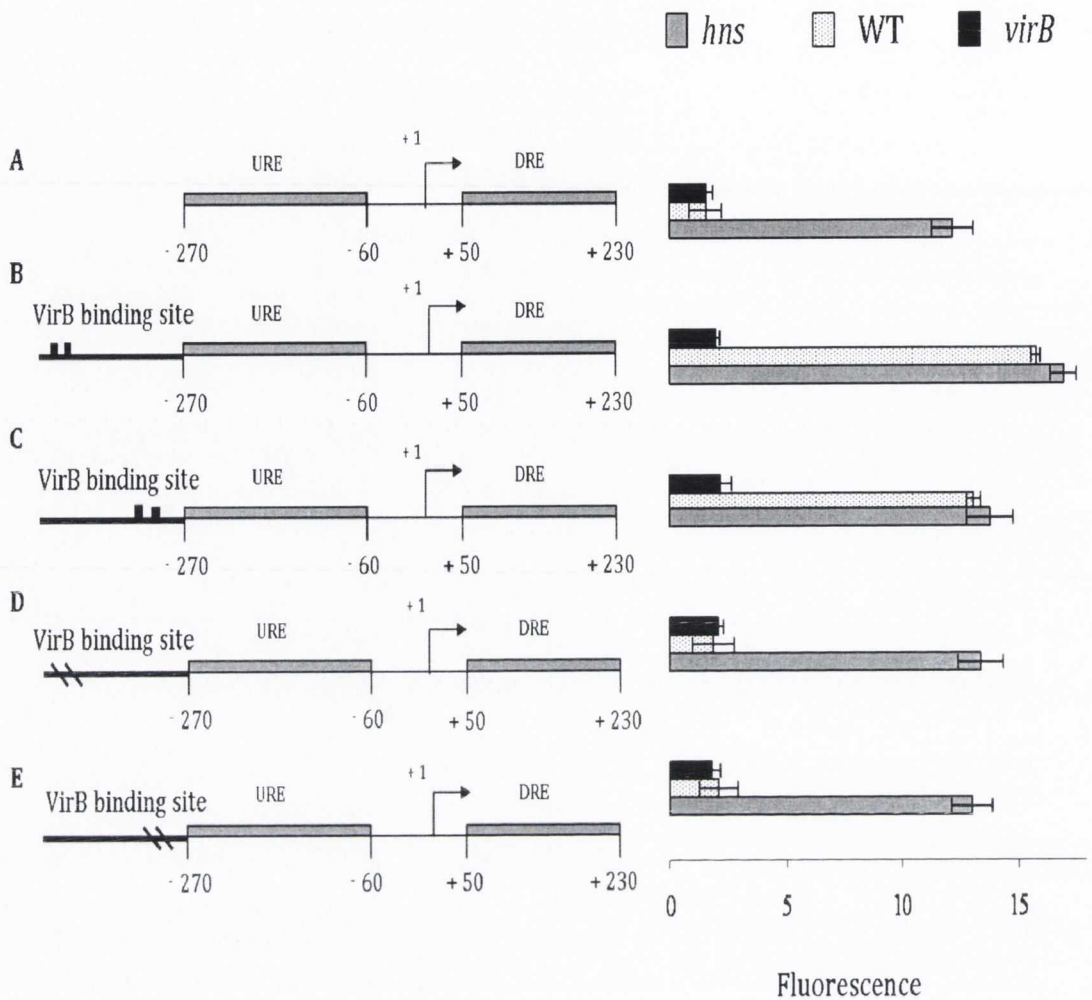


Figure 3.6. De-repression of the *proU* promoter by VirB. Summaries of the structures of various derivatives of the *E. coli proU* promoter region are presented at left; levels of expression of the *proU-gfp* transcriptional reporter fusion are given at right (diagrams not to scale). The native *proU* promoter is shown in (A); *proU* incorporating functional VirB binding and nucleation sites (denoted by two black boxes) shown in the forward (B) reverse orientations (C); *proU* with inactivated VirB binding sites (denoted by two dashes in the VirB site) in the forward (D) and reverse (E) orientations (drawings not to scale). The constructs were assessed for *proU-gfp* expression in a *virB* mutant (black bars), the wild type (white bars) and the *hns* mutant (grey bars) under conditions normally repressive for *proU* transcription

levels were similar to those seen in the *hns* mutant, suggesting that up-regulation of the *proU* promoter involved relief of H-NS-mediated repression. This de-repression of *proU* in low salt was contingent on the presence of the VirB protein. The ability of VirB to de-repress the promoter was not influenced by the orientation of the VirB binding site (Figure 3.6 B and C)

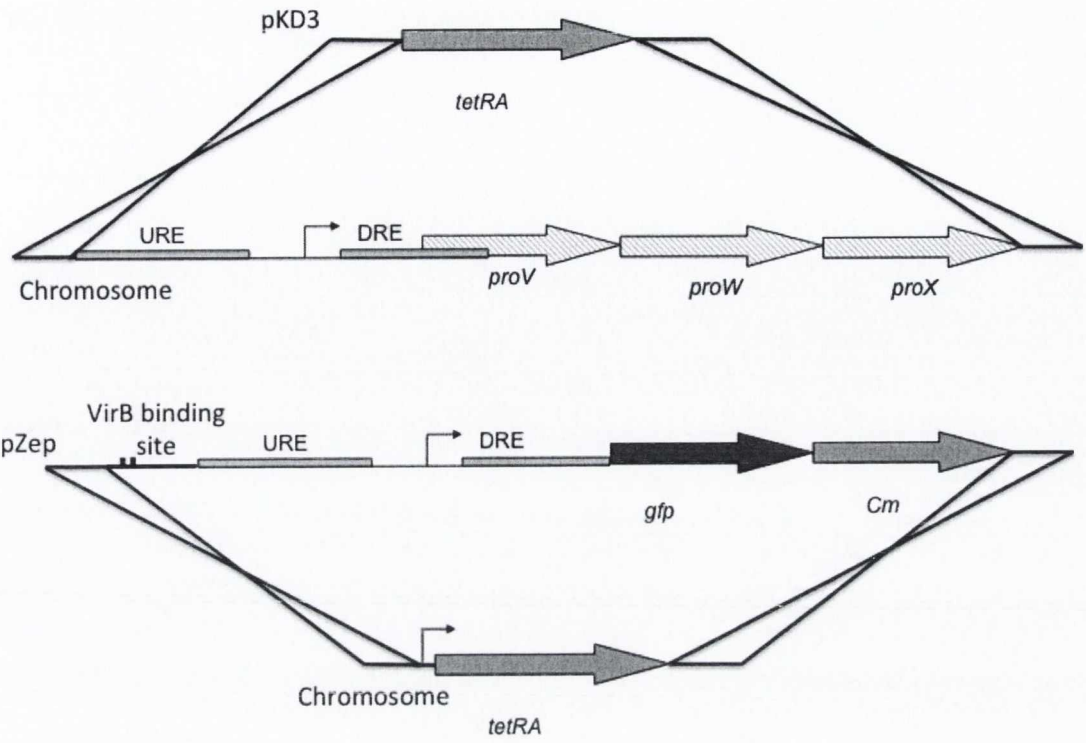
### **3.2.3 Chromosome integration of modified *proU* promoters.**

#### **3.2.3.1 Expression of modified *proU* promoter on the chromosome mimics plasmid based experiments.**

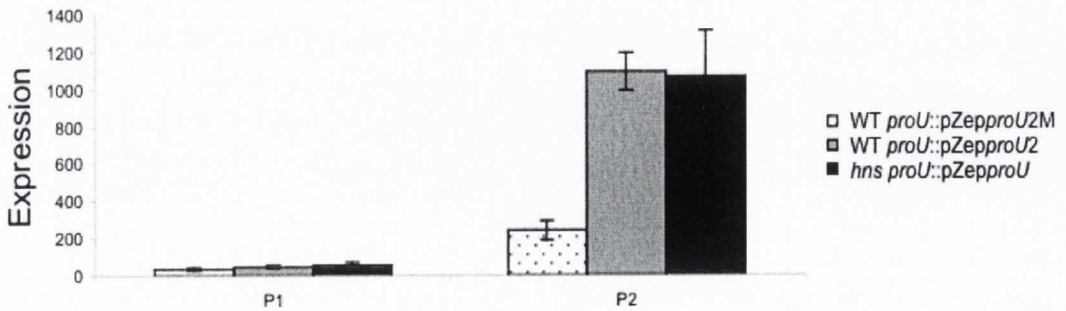
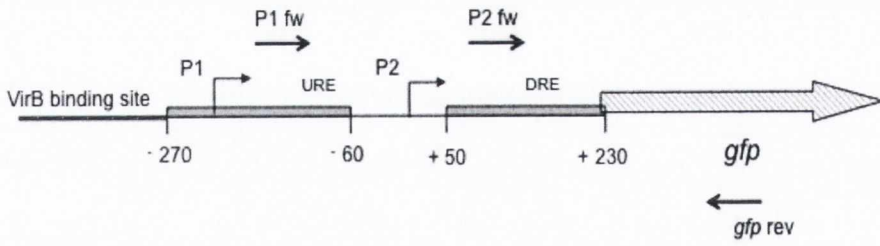
To address the possibility that plasmid-associated artefacts might have influenced the results, the *proU-gfp* fusion constructs with the various forms of the VirB binding site were integrated into the chromosome at the native *proU* locus. These strains were constructed first by insertion of the tetracycline resistance cassette in place of the native *proU* locus by lambda-Red recombination (2.6.8) using the primer set *tet.40bp.proU.fw* and *tet.40bp.proU.rev* (Table 2.3) Using this *proU::tet* strain (Table 2.1), and the reverse primer *cat.rev.40bp.proU* coupled with either *icsB.fw.40bp.proU* or *icsB.rev.40bp.proU* as forward primers, lambda-Red recombination was employed again to insert the modified *icsB-proU-gfp-cat* DNA segment into the native *proU* locus. (Figure 3.7A) This avoided the possibility of recombination occurring between the native *proU* promoter region and the *proU* promoter region on the plasmid resulting in

Figure 3.7 (A) Strategy for construction of a single-copy *proU-gfp* fusion on the chromosome. *tetRA* was amplified from plasmid pKD3 by primers that included 40-nucleotide tails exhibiting perfect homology to regions outside the *proU* locus on the chromosome where it was to be inserted. Using this *proU::tet* knockout mutation as template, the *icsB-proU-gfp-cat* region amplified from pZep08 could be inserted into the native *proU* locus replacing the *tetRA*. (B) Transcription from derepressed *proU* promoters is mediated by promoter P2, not P1. Schematic of pZep-*proU*-2, the modified *proU* promoter that was integrated into the *proU* locus on the chromosome. Primers P1 fw, P2 fw and *gfp* rev (Table 2.4) are indicated with arrows. Quantitative PCR measurements of gene transcript levels (arbitrary values relative to control gene *rho*) from primer sets P1 fw and *gfp* rev, and P2 fw and *gfp* rev. Measurements were taken from wild type cells with the pZep-*proU*-2 (functional VirB binding site at position -270) inserted, wild type cells with pZep-*proU*-2M (mutated binding site at position -270) inserted, and *hns* cells also harbouring pZep-*proU*-2 in place of native *proU*. All experiments were conducted with bacteria grown under repressive conditions in 0 mM NaCl growth medium.

A



B





the loss of the adjacent upstream VirB binding site. Results obtained with these strains were similar to those observed from the plasmid-based experiments (data not shown).

### **3.2.3.2 VirB dependent de-repression of *proU* involves promoter P2 not P1.**

To verify further the results from the plasmid based experiments, the constructs pZep-*proU*-1, pZep-*proU*-2M and pZep-*proU*-2 (Table 2.2) were integrated onto the chromosome at the native *proU* locus in both wild type and *hns* backgrounds (Figure 3.8). The level of transcription was then monitored by qRT-PCR which also facilitated differentiation between transcripts originating at the *proU* P1 and P2 promoters. As before, with the VirB binding site at position -270 (pZep-*proU*-2) the *proU* promoter was derepressed in the *E. coli* wild type to levels similar to those seen in the *hns* mutant. With the mutated, non-functional binding site for VirB inserted at the same position, VirB dependent derepression was not observed (pZep-*proU*-2M). Using forward primers specific for detection of transcripts from either P1 (P1.RT.fw) or P1 and P2 combined (P2.RT.fw) (Table 2.3) it was concluded that VirB mediated de-repression occurred at the P2 promoter *not* at the adjacent P1 promoter (Figure 3.7 B).

### **3.2.4 VirB-dependent derepression is conditional on binding site location.**

The VirB binding site was inserted into different locations upstream of the *proU* promoter to investigate the importance of its position relative to the URE and DRE, the *cis*-acting negative regulatory elements that are bound by H-NS (Figure 3.1). The DRE

found downstream of the *proU* promoter is known to be imperative for the maintenance of *proU* repression by H-NS (Bouffartigues et al., 2007). In contrast, the URE has been shown to be less essential for H-NS-mediated repression in low osmolarity growth conditions (Lucht et al., 1991, Nagarajavel *et al.*, 2007). To assess the significance of the URE for the VirB-mediated derepression mechanism, two new *proU* promoter constructs were made, one with the VirB binding site inserted into the URE and the other with the site located between the URE and DRE (Figure 3.8).

Insertion of the VirB binding site at position -150 with respect to the *proU* transcription start site interrupted the URE at approximately its mid-point (Figure 3.8A-C). A forward primer *proU.fw.-150.NotI* (Table 2.3) annealing to position -150 was coupled with *proUrev+230.XbaI*. This product was inserted into pZep08 and the VirB binding site from *icsB* could then be ligated upstream of the *proU* region as before (Section 3.2.2.1). Insertions were made at -150, of the native VirB binding site (Figure 3.8 B) and of the site with the Box 1/Box 2 deletions (Figure 3.8 C). The mutated site was unable to support derepression of the *proU-gfp* fusion in the presence of VirB while the presence of the native binding site resulted in only partial VirB-dependent depression. In both cases, derepression was assessed by comparison with the level of *proU-gfp* expression seen in the *hns* knockout mutant (Figure 3.8).

The VirB binding site was next inserted at position -60, placing it between the intact URE and DRE elements. This insertion also had the effect of displacing the URE further upstream from the transcription start site; its promoter proximal boundary was now at



position -160 rather than -60 (Figure 3.8). Insertions were made at this position of both the intact binding site (Figure 3.8D) and the mutant derivative with the Box 1/Box 2 base pair substitutions (Figure 3.8E). As before, the mutant binding-site could not support VirB-dependent derepression of *proU-gfp* transcription, whereas the intact binding site did (Figure 3.8 B and C). However, the level of derepression achieved was not as great as that seen in the *hns*-deficient strain. Taken together, these data indicated that successful antagonism of *proU* repression was affected by both the presence and the location of an intact URE.

### **3.2.5 Removal of the URE abolishes VirB-dependent derepression of *proU*.**

The *proU* promoter region was modified so that the URE was removed completely and replaced by the VirB-binding-site (Figure 3.9A and B). This was achieved by performing PCR using the primer set *proU.fw-60*.NotI, which anneals to position -60 with respect to the *proU* transcription start site, and the reverse primer *proU.rev+230*.XbaI (Table 2.3) This product was inserted into pZep08 and the VirB binding site from *icsB* could then be ligated upstream of the *proU* region as before (Section 3.2.2.1). A control construct was produced in which the altered form of the VirB binding site containing the Box 1/Box 2 deletion had replaced the URE (Figure 3.9C). Neither of these constructs supported VirB-dependent derepression of the *proU-gfp* fusion, although the fusion was derepressed when introduced to an *hns* mutant strain (Figure 3.9).

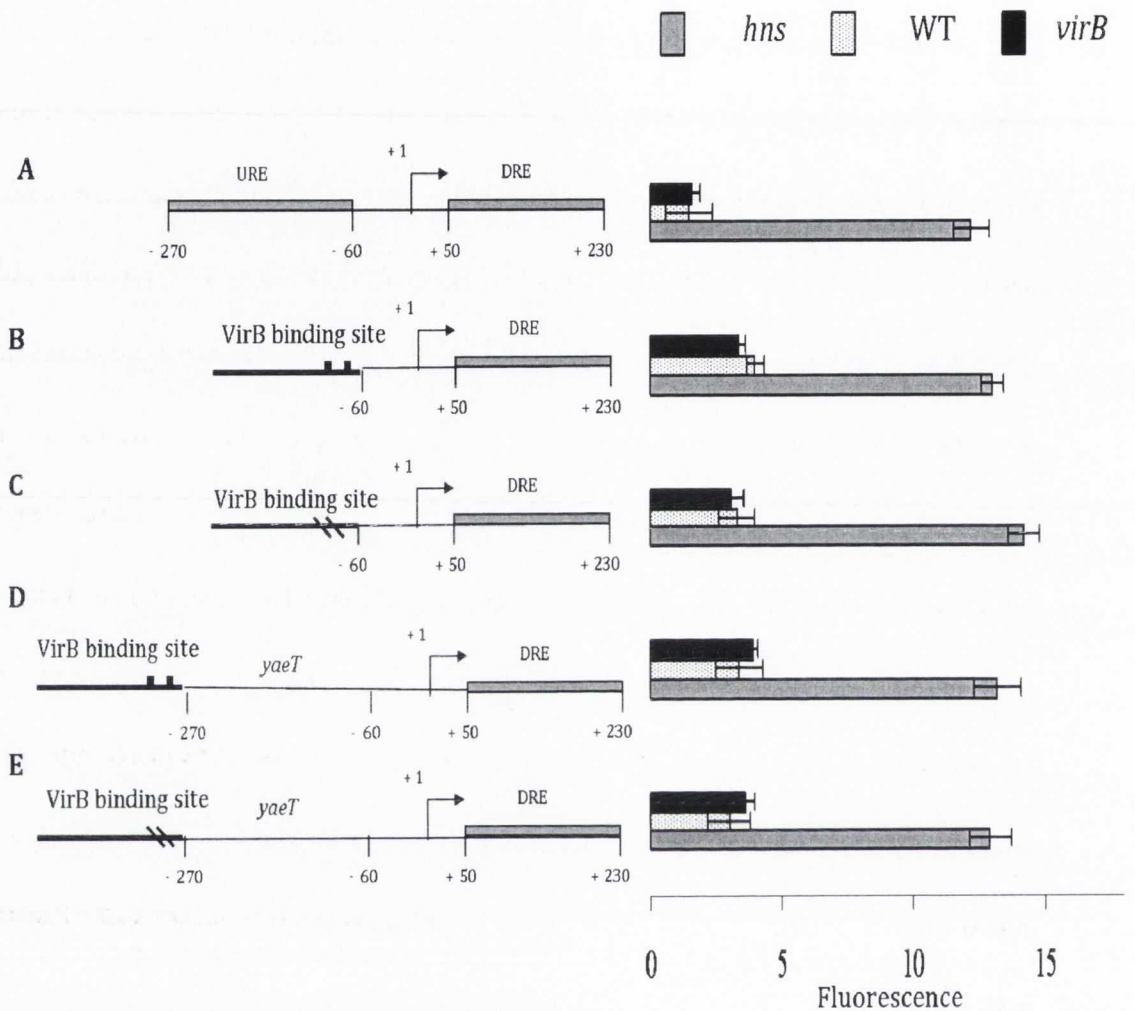


Figure 3.9. The URE and VirB-dependent *proU* de-repression. Summaries of the structures of various derivatives of the *E. coli proU* promoter region are presented at left; levels of expression of the *proU-gfp* transcriptional reporter fusion are given at right (diagrams not to scale). The native *proU* promoter is shown in (A); *proU* with a functional VirB binding site at position -60 (B) or a mutated VirB binding site at the same position (C); *proU* with a functional VirB binding site at position -270 with the URE removed (D) or with an inactivated VirB binding site at the same position (E). All constructs were assessed for *proU-gfp* expression in a *virB* mutant (black bars), the wild type (white bars) and the *hns* mutant (grey bars) under conditions normally repressive for *proU* transcription.

Full VirB-dependent derepression of *proU* transcription had been achieved when the VirB binding site had been located immediately upstream of the URE (Figure 3.9B and C). In this construct, the promoter-proximal boundary of the binding site had been at position -270. The constructs where the URE had been replaced by the VirB binding site (Figure 3.9B and C) were modified by the insertion of a DNA sequence from the *yaeT* gene. The primer couple *yaeT*.fw.SmaI and *yaeT*.rev.NotI (Table 2.3) amplifies DNA from the *yaeT* coding region downstream of the promoter, that is known not to bind the H-NS protein (Dillon et al., 2010). This insertion had the effect of moving the promoter-proximal boundary of the VirB binding site to position -270 with respect to the transcription start site, restoring the distance (but not the H-NS binding activity) associated with the presence of the URE (Figure 3.9D and E). The presence of an intact VirB-binding site did not result in VirB-dependent derepression of *proU-gfp* transcription in the absence of the URE, despite the placement of the binding site at a location from where it was effective when the URE was present (compare Figure 3.9B and D).

Both sets of constructs in which the URE was removed showed a slight alleviation of repression, but maintained a 3- to 4-fold repressive state as compared to an *hns* mutant. This is consistent with the hypothesis that while the URE and DRE may act synergistically, the DRE alone can maintain significant repression (Nagarajavel et al., 2007).

### 3.2.6 Modified *proU* promoter constructs retain osmo-sensitivity.

To ensure that the *icsB-proU* hybrid promoters retained the osmotic sensitivity that is characteristic of *proU*, the hybrid promoters were tested for activity under conditions of increasing osmolarity. Both the native *proU* promoter and those derivatives that had been modified to accept a binding site for the VirB protein showed strong sensitivity to increasing osmolarity, showing that the underlying *proU* promoter function had not been altered by the introduction of DNA containing the *S. flexneri icsB* VirB binding site (Figure 3.10).

### 3.2.7. URE is a functional transcription regulatory element

Although previous work had indicated that the URE is less essential than the DRE for H-NS mediated repression of *proU* transcription at low osmolarity (Nagarajavel et al., 2007), the experiments described in the present study had revealed that it influenced the efficacy of the VirB-mediated derepression mechanism. The functionality of the URE was examined in more detail by exploiting the presence of the divergently oriented *ipgD* promoter that is located immediately upstream of the *icsB* promoter (Figure 3.1A). When this promoter was placed immediately upstream of the full-length URE, its presence resulted in a detectable elevation in *proU-gfp* transcription (Figure 3.11A and B). This level of transcription increased when the URE was truncated (Figure 3.11C) and reached values similar to those detected in an *hms* knockout mutant when the URE was eliminated completely (Figure 3.11D). These data showed that the URE could

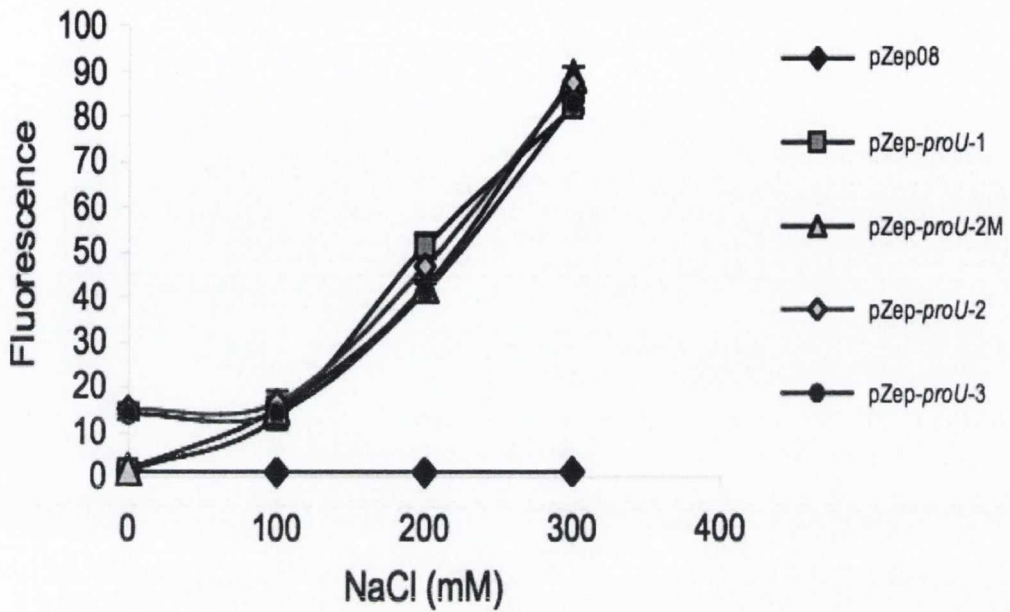


Figure 3.10. Osmoregulation of modified *proU* promoters. Expression of Gfp was measured in bacteria growing in increasing concentrations of NaCl from pZep08 (promoter-less *gfp*), pZep-*proU*-1 (native *proU* promoter), pZep-*proU*-2M (*proU* with mutated VirB binding site at -270), pZep-*proU*-2 (*proU* with functional VirB binding site at -270 in the reverse orientation), pZep-*proU*-3 (*proU* with functional VirB binding site at -270 in the forward orientation).



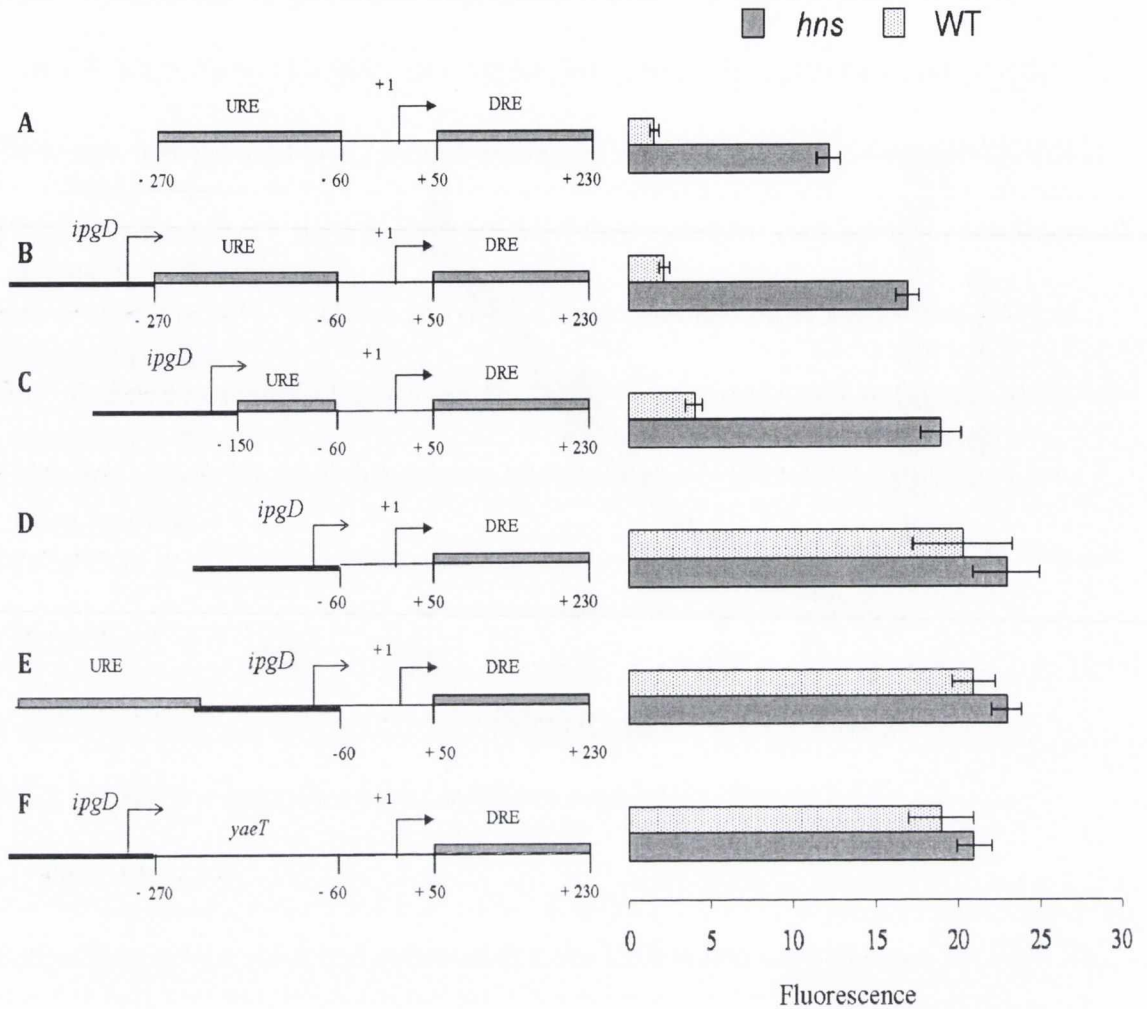


Figure 3.11. The URE imposes a barrier to transcription read-through. Summaries of the structures of various derivatives of the *E. coli proU* promoter region are presented at left; levels of expression of the *proU-gfp* transcriptional reporter fusion are given at right (diagrams not to scale). The native *proU* promoter is shown in (A); the *proU* promoter with the *ipgD* promoter located at position -270 (B); the *proU* promoter with the *ipgD* promoter located at position -150 (C); the *proU* promoter with the *ipgD* promoter at position -60 (d); the *proU* promoter with the *ipgD* promoter at position -60 with the URE immediately upstream (E), and the *proU* promoter with the *ipgD* promoter at position -270 with the URE removed (F). The constructs were assessed for *proU-gfp* expression in the wild type (white bars) and the *hns* mutant (grey bars) under conditions normally repressive for *proU* transcription.

inhibit the ability of the *ipgD* promoter to influence *proU-gfp* expression. Moreover, this inhibitory activity was dependent on the presence of H-NS, because it was abolished in an *hns* mutant (Figure 3.11). Reintroduction of the URE to a position upstream of the *ipgD* promoter did not impair the positive influence of the promoter on *proU-gfp* expression (Figure 3.11 E). This showed that the presence of the URE *per se* was not responsible for its inhibitory effect on *ipgD* promoter influence, but rather it was the presence of the URE between the *ipgD* promoter and the *proU-gfp* reporter fusion that was significant. This positional requirement, together with the need for an intact *hns* gene suggested that the URE was acting as a transcriptional barrier by virtue of its ability to bind H-NS. This hypothesis was supported by the observation that replacement of the URE by a DNA sequence from the *yaeT* gene that does not bind H-NS failed to impede the ability of the *ipgD* promoter to up-regulate *proU-gfp* transcription (Figure 3.11 F).

### 3.3. Discussion

The creation of new genetic regulatory networks is something routinely predicted in bioinformatics studies of evolution. Activators not native to a particular regulon may be incorporated into the regulatory cascade by the acquisition of an appropriated positioned binding site.

Modifying the regulation of a system by insertion of a binding site for a DNA binding protein has been studied for the *proU* and *bgl* regulatory regions. These previous

attempts at altering regulation have relied on a more sophisticated mechanism of activation. The TyrR protein and its specific TYR R box binding site was introduced adjacent to the P2 promoter in *proU* regulatory region and conferred TyrR mediated activation of *proU* under non-permissive low osmolarity conditions. TyrR acts as a class I transcription factor by increasing the affinity of binding of RNA polymerase to the promoter and stimulating open complex formation (Gowrishankar and Pittard, 1998). In that case, the TyrR binding site must be placed in the proximity of the promoter taking into account spatial constraints that might interfere with physical interactions between the transcription factor and RNA polymerase.

The *bgl* regulatory region exhibits parallels to *proU* in that it also possesses two high affinity binding sites for H-NS that flank the promoter. However, in contrast to *proU* both the upstream silencer and downstream silencer are required for maximal H-NS repression, and deletion of either silencer sequence results in alleviation of repression (Nagarajavel *et al.*, 2007). Caramel and Schnetz (JMB 1998) introduced the *lac* operator sites at various locations within the upstream H-NS binding site and were able to demonstrate activation of *bgl* in the presence of the Lac repressor. This is achieved presumably by creating direct competition between the Lac or lambda repressor and H-NS for access to *bgl* DNA, thus disrupting H-NS binding upstream, undermining the nucleoprotein complex formed by H-NS which requires both sites to be occupied.

The mechanism demonstrated in this study however, does not involve anti-repression by recruitment of RNAP or disruption of repressor sites. The system created involves

the removal of H-NS from a promoter by VirB, a ParB homologue. Here, we have recreated a proposed evolutionary step, in which it may be possible to co-opt VirB into regulatory roles at many H-NS regulated promoters.

However, interpreting the results, they suggest that this re-directing of DNA-binding proteins into other roles may not be straight-forward and this ‘wiring’ of regulation undoubtedly requires fine tuning. Introducing a binding site for VirB immediately upstream of the URE in the full *proU* promoter, allowed for full de-repression of the promoter. The model that is most likely to account for the VirB-mediated derepression of the *proU-gfp* fusion under normally non-permissive growth conditions is one that assumes that both the URE and the DRE participate in an H-NS-dependent transcription repression complex. This repressive nucleoprotein complex here is presumably a linking of URE and DRE, and the formation of an URE–DNA–DRE bridge. The bridges formed by H-NS are considered somewhat fragile, requiring 7 pN to break, when compared to another DNA bridging molecule LacI, which require 500 pN (Dame *et al.*, 2006). The binding of VirB upstream of this site of inhibition, with the associated remodelling of the complex induced by wrapping of the DNA by VirB might be expected to reduce its stability (Figure 3.12.1).

It is clear that positioning of the binding site is crucial and this suggests the limitations of de-repression are perhaps dependent on the proximity of the VirB antagonist to the H-NS repressor. At the *icsB* promoter the evidence of DNA wrapping by VirB extends from the VirB nucleation site to 120 bp downstream, despite H-NS binding continuing

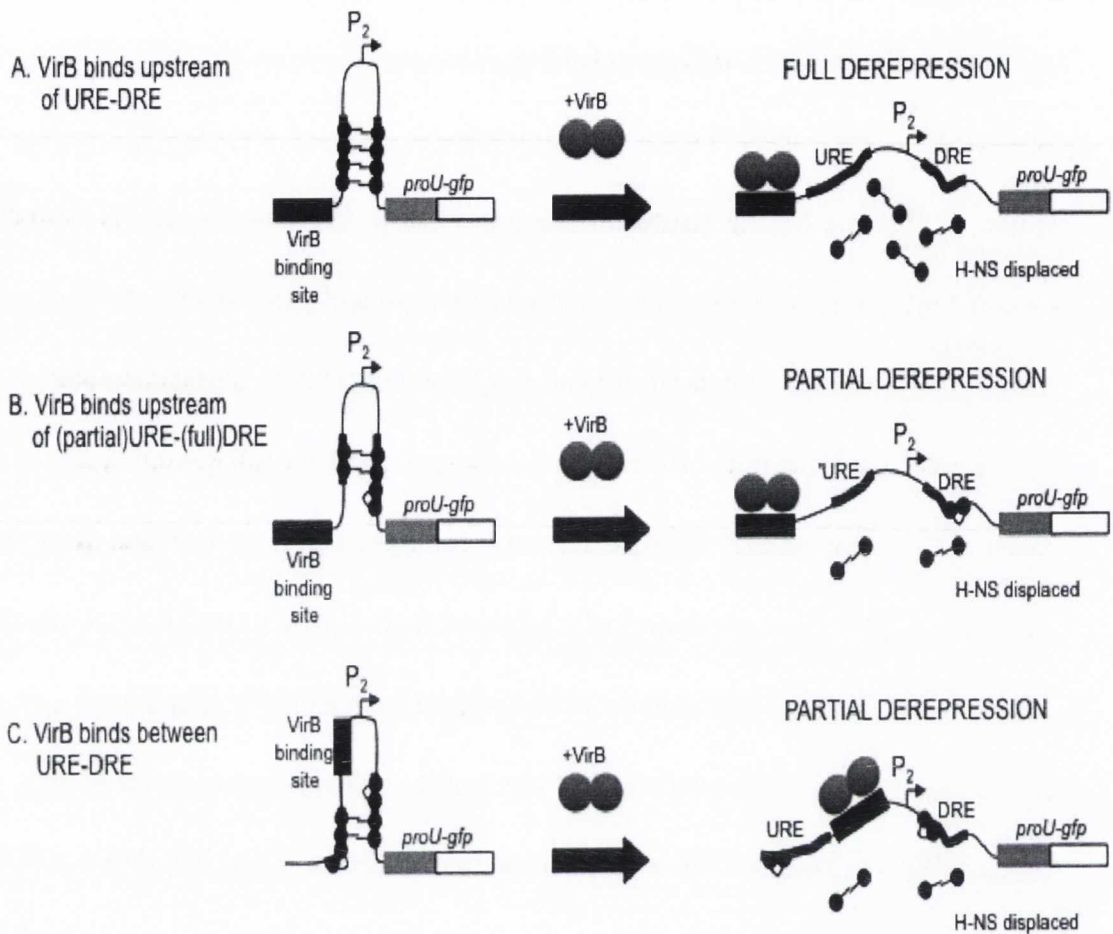


Figure. 3.12.1 VirB-dependent de-repression of the *proU* promoter. The model is based on differential sensitivity of the known *trans*-binding (i.e. bridging) and *cis*-binding (non-bridging) modes of DNA binding exhibited by the H-NS protein to antagonism by the VirB protein. Full derepression is achieved when the URE and DRE are intact and the binding site for the VirB protein is located immediately upstream of the URE (a). This mimics the situation in *icsB*, a natural target for VirB in *S. flexneri*. Here, VirB acts to antagonize DNA-H-NS-VirB bridging. If the URE is truncated (b) the degree of bridging is reduced, allowing H-NS to adopt a non-bridging binding mode over part of the DRE, imposing a partial barrier to full transcriptional derepression. When the VirB binding site is located between the URE and DRE (c), derepression is also inefficient.

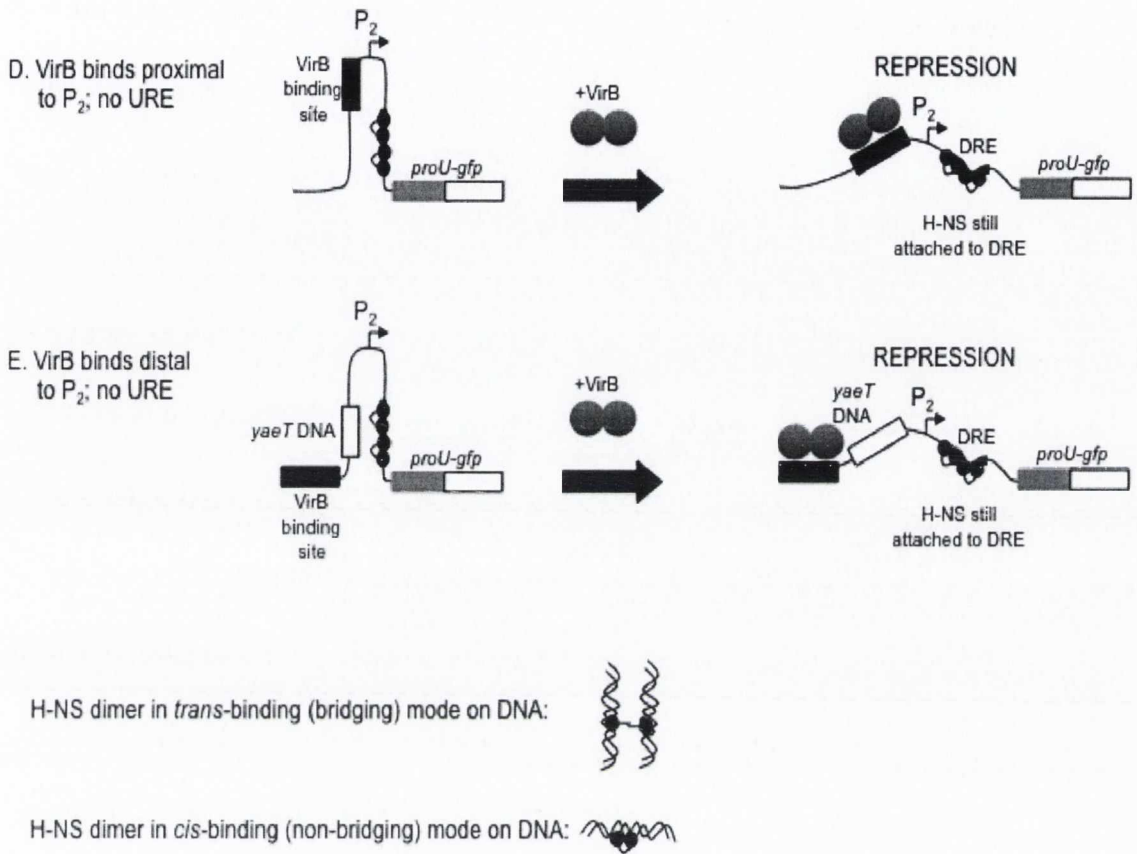


Figure. 3.12.2 Repression of the *proU* promoter. Removal of the URE (d) abrogates bridging by H-NS and the resulting *cis* binding mode over the DRE produces a partially effective barrier to *proU* derepression. Similarly, displacement of the *virB* binding site to a position at  $-270$  but with an H-NS non-binding DNA sequence (*yaeT*) in place of the URE, also permits the development of the *cis*-bound H-NS complex at the DRE that is partially effective at preventing *proU* derepression.

another 40 bp (Turner and Dorman, 2007). This may demonstrate an upper limit to the range of DNA that can be distorted by VirB. This would imply VirB may only be capable of displacing H-NS bound in the vicinity of its nucleation site. With this in mind, in the full length *proU* promoter (Figure 3.6) VirB displaces the URE, which simultaneously removes the DRE bound *in trans*, and alleviates repression (Figure 3.12.1 A). With part of the URE removed or the position of the URE altered (Figure 3.7), the H-NS molecules bound at the DRE and URE may be forced to adopt a *cis*-binding mode rather than a *trans*-binding mode. Again VirB displaces the adjacent URE which only now partially disrupts the DRE (Figure 3.12.1 B, C). With the URE removed (Figure 3.8), the VirB nucleation site is either 110 bp (VirB site at -60) or 320 bp (VirB site upstream of *yaeT* region) from the nearest H-NS site located at the DRE, which may be beyond the limits of VirB wrapping (Figure 3.12.2 D, E).

A binding site for VirB has recently been uncovered at the VirB-dependent *icsP* promoter, almost 1 kb from the transcription start site (Castellanos *et al.*, 2009).

However, recent findings have shown H-NS binding to also extend this distance from the promoter, placing VirB immediately upstream of the H-NS protein at this regulatory region. Therefore, while the literature describes this VirB binding site as “remote”, this is only true in terms of its distance from the de-repressed promoter. As the VirB protein is not a conventional transcription factor, unlike TyrR described above, this distance is somewhat irrelevant. The positioning of VirB should be thought of in terms of its proximity to the repressor H-NS.

A role for the apparently dispensable URE H-NS site has also been revealed. With H-NS bound in the URE, read-through transcription from a neighbouring promoter driving into *proU* was incapable of disturbing expression (Figure 3.11). With the DRE holding the primary role as the H-NS repression site, perhaps the URE became redundant in terms of maintaining repression at low salt, however it remained intact because of its role as a protection barrier against local topological disturbances. This hints at another case of redundant DNA binding proteins being re-directed into other functional roles.

The modified *proU* constructs all retained osmoregulation and expression was highly induced under osmotic upshock, comparable to that of the original wild type promoter (Figure 3.10). The VirB dependency offered an additional level of regulation to the promoter without altering its native expression pattern. Also the titre of VirB protein in the cell was that of wild type, if not slightly below as *virB* is downregulated in low salt conditions (Porter and Dorman, 1997). Western blot analysis confirmed a slightly lower concentration of VirB in the cell as compared to a standard osmolarity, however monitoring expression of a *mxiC::lacZ* fusion on the virulence plasmid (a VirB dependent gene) we were able to demonstrate the concentration of VirB was sufficient to exert an effect (Figure 3.5). The constructs were not bombarded with overexpression of VirB protein, rather natural levels of the protein were used to monitor de-repression. The modified promoters were also integrated into the native *proU* loci on the chromosome to address any artefacts that may arise when using a plasmids based expression system. The results obtained verified our initial data. Therefore, although the constructs tested here are artificial the system does mimic that of wild type, in the



physiologically conditions used, the concentration of VirB protein, and the gene position and osmoregulation of *proU*. This serves to support the hypothesis of this event occurring naturally in evolution.

Overcoming the overwhelming negative effect on gene expression fortified by H-NS promoted the evolution of a variety of anti-silencing mechanisms. The pattern emerging is one of *ad hoc* solutions to H-NS-mediated repression in which anti-repression often involves the re-deployment of the existing DNA-binding protein, rather than the evolution of specific dedicated antagonists. For example, DNA-wrapping proteins or DNA bending proteins that distorts local DNA topology, undermining H-NS bridging, could all be employed as H-NS antagonists. The story of VirB, a presumptive former plasmid partitioning protein, is an elegant example of this evolution in play. The strong amino acid similarities between VirB and ParB, along with the nucleotide sequence similarities between VirB binding site and *parS*, are strong evidence that VirB may once have functioned as a member of the plasmid partitioning family but became re-directed into a regulatory role. We have demonstrated that insertion of a VirB binding site can confer VirB-mediated de-repression on an H-NS repressed promoter. This supports the hypothesis that the VirB binding site may once have originated in the vicinity of its gene, analogous to its homologue ParB and its *cis*-acting *parS* site. The insertion of the *icsB-ipg-ipa-acp* operon disconnected the VirB site from its gene, while simultaneously placing it at a position in the *icsB* promoter region whereby it can act to de-repress H-NS, creating a new role for the protein. The results also illustrate the fine-tuning required for this regulatory network innovation to be successful. Correct positioning of the

binding sites for the antagonist and repressor are essential for de-repression to occur. This would imply that perhaps these rearrangements of binding sites and co-opting of proteins are events happening routinely in bacteria, however only those appropriately positioning and functional are selected for and are apparent in modern day regulatory systems. This may offer the bacteria the potential to explore and/or exploit novel environmental niches. It is clear studies of H-NS antagonism have much to teach us about the evolution of regulatory switches, but also highlights pit-falls of assigning unique biological properties to DNA-binding proteins.

## **Chapter 4**

**Auto-regulation and positive feedback are features of  
the *Shigella* virulence cascade**

#### 4.1. Introduction

Invasion of host cells by *S. flexneri* requires expression of virulence genes located on a 230-kb plasmid. The products of these genes include the Ipa invasins, which mediate the invasion of the gut epithelium and macrophage apoptosis, the Mxi and Spa proteins which form a type III secretion system, and the IcsA, IcsB and VirA proteins which are responsible for intercellular spreading of bacteria in the lower gut (Buchrieser et al., 2000). The virulence genes are expressed only under conditions approximating to those found at the site of infection, i.e. at 37°C, a pH of 7.4 and moderate osmolarity (Porter and Dorman, 1994, 1997a; Dorman et al., 1990). The chromosomally encoded H-NS protein controls much of this expression through repression of many of the virulence promoters under conditions that are inappropriate for invasion (Maurelli and Sansonetti, 1987; Beloin and Dorman, 2003). Virulence gene activation occurs via a regulatory cascade mediated by the plasmid encoded VirF and VirB proteins (Dorman and Porter, 1998).

The transcriptional cascade is initiated with the activation of *virF*. At this promoter H-NS protein is bound at 2 sites, one overlapping the transcription start site and another centered upstream at position -250, separated by a region of intrinsically curved DNA (Falconi et al., 1998; Prosseda et al., 1998). Repression involves the formation of a nucleoprotein complex whereby H-NS proteins bound at these two sites are thought to act co-operatively (Prosseda et al., 2004). The nucleoprotein complex is temperature sensitive, forming at temperatures below 32°C but being disrupted by the change in

DNA topology upon a temperature upshift to 37°C. This disruption in local DNA architecture simultaneously reveals a binding site for the transcriptional activator FIS. FIS binds to four specific sites in the promoter region of *virF* with two of the four sites overlapping one of the two H-NS sites, responsible for thermoregulation. Results indicate that FIS can exert some antagonism to H-NS repression by partially counteracting H-NS inhibition at 32°C. However the principal role of FIS at the *virF* promoter is to exercise direct positive transcriptional control at a permissive temperature 37°C, when H-NS fails to repress *virF* (Falconi *et al.*, 2001). Expression of *virF* is also facilitated by the IHF architectural protein, which bind to a region localized downstream of the promoter at position +51. *virF* is stimulated by IHF in both logarithmic and early stationary phases of growth (Porter and Dorman, 1997b).

With the repression of *virF* overcome, the AraC-like protein VirF activates *virB*. VirF interacts directly with the *virB* promoter in a region spanning -17 to -105 (Figure. 4.1) (Tobe *et al.*, 1993). These VirF binding sites are located immediately upstream of the region bound by H-NS which coats the promoter, targeting -20 to +20. Removal of *hns* from the cell through mutation of the *hns* gene leads to an increased expression of the *virB* gene at 30°C, suggesting that H-NS plays a role in controlling the thermal regulation of *virB* expression. However, there is an absolute requirement for the presence of the VirF protein, as removal of H-NS from the cell will not permit transcription in the absence of VirF. The current model predicts the interaction of VirF with RNA polymerase to induce transcription, and this is dependent on an increase in negative supercoiling (Tobe *et al.*, 1993, 1995). This occurs due to the shift from 30°C to 37°C

and under conditions of high osmolarity. Thus at 30°C VirF activation is insufficient to overcome repression by H-NS but at 37°C or under high salt conditions topological changes at the promoter make interactions between VirF and RNA polymerase more favorable, leading to transcription activation. *virB* expression is also enhanced by the direct binding of IHF to the region -171 to -183, upstream of the DNA that is bound by VirF (Porter and Dorman 1997). It acts positively on *virB* transcription in stationary phase (Figure. 4.1).

While this is a relatively simplified synopsis of activation of virulence expression on the Large Virulence Plasmid, it is clear that the process is intricate and tightly regulated. By employing its own specific activators, the overwhelmingly negative effects of the chromosomally encoded H-NS can be overcome in a timely and efficient manner. The removal of H-NS from the promoter of virulence genes requires a rapid surge in the expression of the activators VirF and VirB after penetration of the bacteria into the host to ensure prompt assembly of the essential type III secretion system. It is not clear why regulation occurs via a *virB* intermediate step; perhaps this provides an additional checkpoint prior to the irreversible activation of a large number of structural genes.

Chapter 3 documents a detailed analysis of the co-option of transcription factor binding sites in the generation of new regulatory circuits, which antagonize H-NS repression. Analysis of VirB and its paralogue ParB suggests that the VirB binding site at the *icsB* promoter may once have originated in the vicinity of the *virB* gene and become disconnected by insertion of the *icsB-ipg-ipa-ACP* operon. This hypothesis suggests *virB*

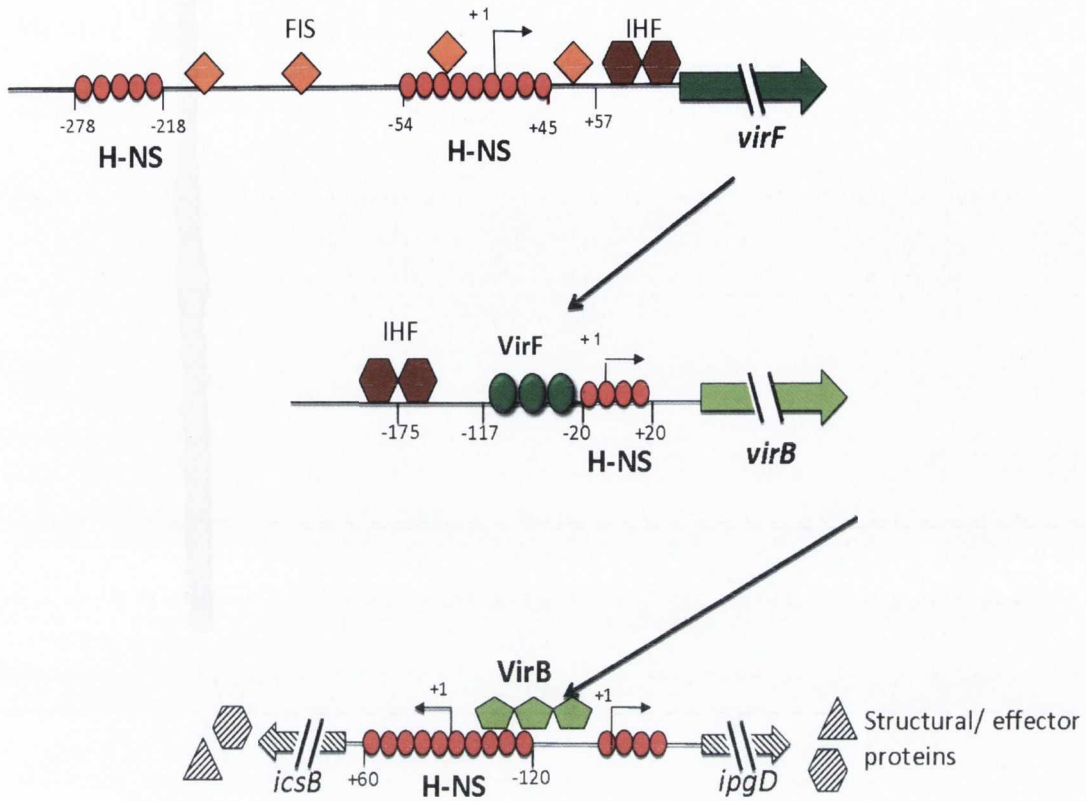


Figure 4.1. Schematic overview of the interplay of transcription factors in the virulence gene regulatory cascade. The *virF* promoter is repressed by H-NS (red ovals). Following a temperature upshift to 37°C which disrupts H-NS binding, the positively acting regulators FIS (orange diamonds) and IHF (brown hexagons) act to up-regulate the promoter. The VirF protein (dark green oval) in turn activates the H-NS-repressed *virB* promoter which is also facilitated by IHF. VirB (light green pentagon) then acts to derepress the H-NS repressed *icsB-ipgD* promoter (among others) by binding in the intergenic region and inducing DNA topological changes, which act to displace H-NS.

may have, at one time, auto-regulated its expression, perhaps acting as its own H-NS antagonist. With this in mind, inserting a binding site for *virB*, upstream of H-NS, at the *virB* promoter, would recreate the hypothetical original arrangement. The *virF* promoter, with its clearly defined H-NS binding site, also provides an interesting target for manipulation as insertion of a correctly positioned VirB binding site at this location would create a positive feedback loop in the regulatory cascade.

The results presented in this chapter however, demonstrate how *virB* auto-regulation and positive feedback to *virF* are already essential features of the virulence regulatory cascade. It also shows how manipulation of this native regulation by insertion of additional activator binding sites can disturb the delicate mechanism of gene activation.

## 4.2 Results

### 4.2.1 Artificial expression of VirB protein at 30°C can alleviate the repressive effects of H-NS at the *virB* and *virF* promoters.

Thermoregulation of the *virB* and *virF* promoters involves transcriptional repression by H-NS at non-permissive temperatures, i.e. 30°C. VirB dependent promoters can be de-repressed under these conditions by artificial over-expression of the VirB protein *in trans* (Beloin and Dorman 2003). The effect of VirB on its own promoter and on that of *virF* was tested using a similar approach. The P<sub>BAD</sub> vector system (Table 2.2) was used to manipulate the levels of VirB protein in the cell. In the plasmid P<sub>BAD</sub>*virB*<sup>+</sup>, the *virB*



gene is placed under the control of the arabinose inducible promoter  $P_{BAD}$ . Therefore titre of VirB protein can be altered with the addition of either glucose or arabinose and the protein levels determined by Western blot (Figure. 4.2 A) The transcript levels specified by the *virB* and *virF* promoters (on the Shigella virulence plasmid) were analyzed by qRT-PCR at 30°C, using primers specific for each promoter (Table 2.3). In the case of the *virB* transcript a forward primer (*virB*.RT.fw+2) was designed to anneal immediately after the native *virB* promoter but preceding the start codon of the *virB* open reading frame. This is a region not included when constructing the  $P_{BAD}virB+$  expression plasmid. This allowed for differentiation between native *virB* transcript and that expressed by the  $P_{BAD}$  promoter.

Under conditions of low levels of VirB protein (addition of glucose) both the promoters of *virB* and *virF* were only weakly active as was expected for these repressive conditions (30°C). However, with the increase in VirB protein production (addition of arabinose) both the *virB* and *virF* promoter were de-repressed. Expression from *virB* increased 5 fold upon addition of VirB protein, while *virF* expression increased 3 fold. Levels of expression from both promoters in an *hns* null mutant produced similar results to that of the VirB de-repressed promoters. This is consistent with the role of *virB* as an H-NS antagonist (Fig. 4.3 b).

As a negative control, an “empty”  $P_{BAD}$  vector was introduced into BS184 (wild type) and BS185 (*hns* mutant). This plasmid does not contain the *virB* gene. These strains were also grown under repressive conditions for expression (30°C), in the presence of

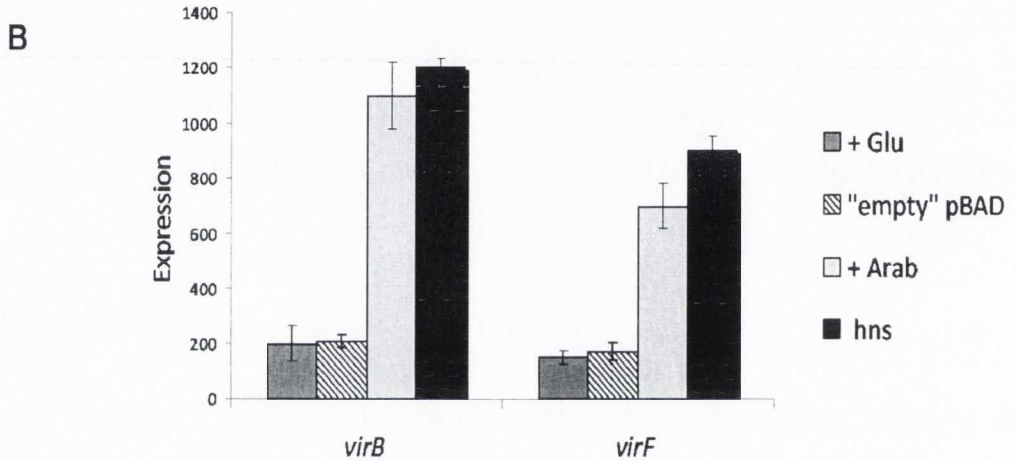
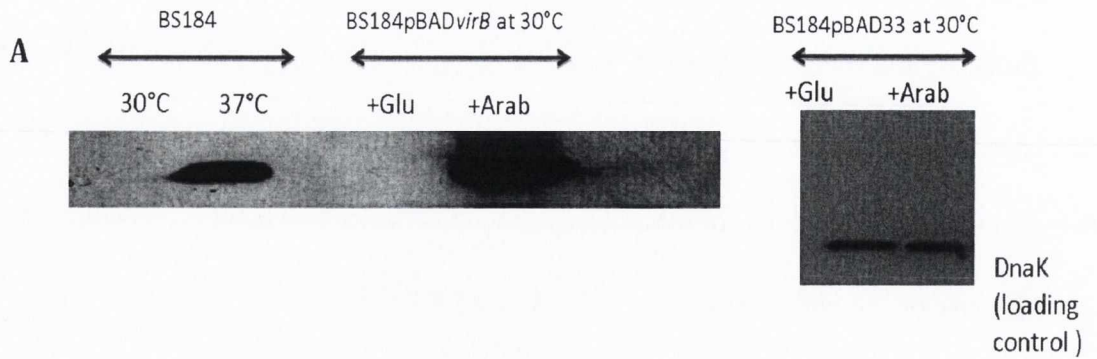


Figure 4.2. Artificial expression of VirB protein at 30°C can alleviate the repressive effects of H-NS at the *virB* and *virF* promoters. (A) Western blot detection of VirB protein in the wild type strain BS184 at 30°C and 37°C and in BS184 harbouring the arabinose-inducible pBAD*virB* plasmid at 30°C in the presence of glucose (P<sub>BAD</sub> promoter is repressed) or arabinose (P<sub>BAD</sub> promoter is induced). Also in BS184 harbouring an “empty” pBAD33 vector in the presence of arabinose and glucose, VirB protein cannot be detected despite the DnaK loading control being present. (B) qRT-PCR analysis of the transcript from native *virB* and *virF* promoter on the virulence plasmid in the presence and absence of VirB protein and in an *hns* null mutant, under repressive conditions of 30°C. As a negative control “empty” pBAD33 was used to account for any alterations addition of the sugars (arabinose/glucose) may have on expression

arabinose or glucose. Western blot analysis of these strains show that no VirB was detected (Figure 4.2 A), although the loading control DnaK is present. qRT-PCR analysis of these strains verified that without the VirB protein present, the *virB* and *virF* promoters cannot be de-repressed (Figure 4.2 B). These negative results act to control for any alterations the addition of sugars (arabinose and glucose) may have on expression.

#### **4.2.2 Auto-regulation of *virB* is dependent on a *cis*-acting VirB binding site.**

Further investigation of VirB autoregulation involved the construction of a *virB-gfp* reporter fusion. A region of 330 bp of the *virB* promoter was inserted upstream of the green fluorescent protein gene (*gfp*) to monitor expression. This created a construct, which incorporated the promoter, the H-NS binding sites spanning from -20 to +20, and the upstream region terminating at -310. The experiments were conducted under repressive conditions of a non-permissive temperature of 30°C. The P<sub>BAD</sub> vector system was used to manipulate the levels of VirB protein in the cell with the addition of either glucose or arabinose. This region of the *virB* promoter exhibited weak expression; however transcription increased 5 fold in the presence of the VirB protein (Figure 4.3 A). To monitor expression levels of a de-repressed *virB* promoter, the constructs were also monitored in an *hns* null mutant. As expected, *virB* levels increased 5-6 fold in the absence of H-NS protein despite the repressive temperature (Fig. 4.3 A). Shortening the promoter region to include only -80 to +190 abolishes the sensitivity of the *virB* construct to the presence of the VirB protein (Fig. 4.3 B). This suggests *virB* auto-

regulation is dependent on the region -310 to -80. This truncated *virB* construct retains the majority of the region bound by VirF and so has the ability to be activated upon removal of H-NS from the promoter. This was demonstrated by expression of the construct in an *hns* mutant which showed a 5-fold increase in promoter activity. These data suggested the possibility of a native binding site for the VirB protein in this area. Previous analysis of the *icsB*, *spa15*, and *virA* promoters of *Shigella sonnei* established a consensus binding site for VirB as being 5'-(A/G)(A/T)G(G)AAAT-3' that often appears as an inverted repeat. Bioinformatic analysis of this region of the *virB* promoter revealed a close match centered at -280. The putative site 5' – ATTGAAAT – 3' is a 7/8 match to the consensus. This site was mutagenized in accordance with previously published data related to the binding specificity of VirB. This created the mutated sequences 5' – GCCTGGGC – 3'. When monitored in the presence or absence of VirB protein at a repressive temperature (30°C) the VirB dependent up-regulation of the *virB* promoter was diminished, suggesting that this was indeed a binding site for VirB (Figure. 4.3 C).

Interestingly, insertion of a VirB binding site from *icsB* at position -310, immediately upstream of the putative native VirB binding site, inhibited this VirB dependent up-regulation (Figure. 4.3 E). When analyzed, constructs with a VirB binding site from *icsB* inserted upstream were unresponsive to the concentration of VirB in the cell.

Repression was only alleviated upon removal of H-NS protein, in an *hns* mutant. Insertion of the region to which VirB binds from the *icsB* promoter at position -80 however restored this VirB-dependent de-repression of the *virB* promoter

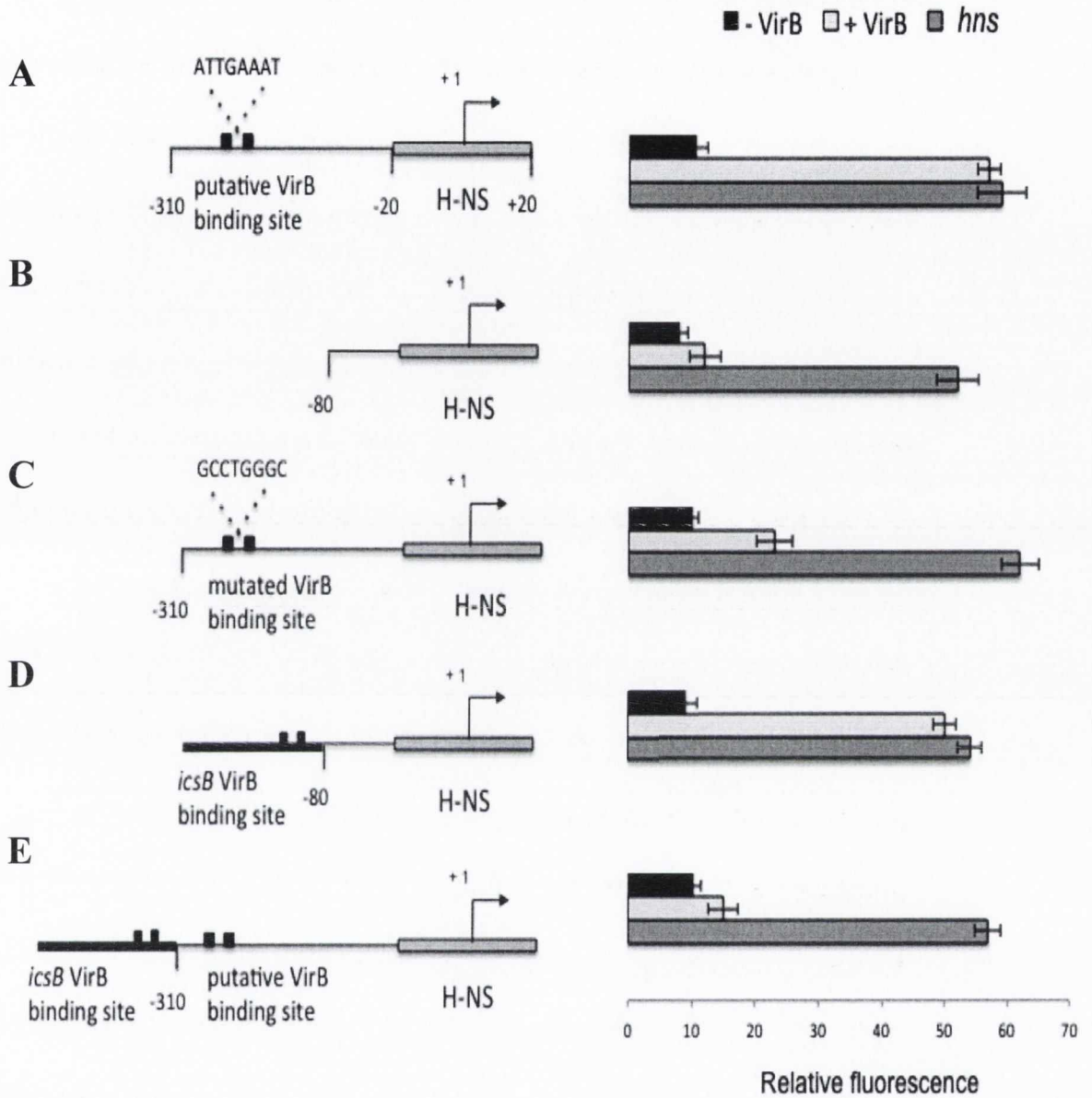


Figure 4.3 Autoregulation of *virB* is dependent on a *cis*-acting VirB binding site. Summaries of the structures of various derivatives of the *virB* promoter region are presented at left; levels of expression of the *virB-gfp* transcriptional reporter fusion are given at right. The native *virB* promoter encompassing positions -310 to +30 is shown in (A) the *virB* promoter from -80 to +30 shown in (B), *virB* promoter with mutations made to the putative VirB site (C), *virB* promoter with VirB binding site inserted at position -80 (D), *virB* promoter with VirB binding site inserted at position -310 (E).

(Figure. 4.3 D). This suggested that the VirB binding site at the *virB* promoter could be removed and functionally replaced by a VirB binding site native to another promoter. The data also implied that two consecutive VirB binding sites were detrimental to the ability of VirB to de-repress the *virB* promoter, as insertion of the VirB site from *icsB* immediately upstream of the native site resulted in no VirB dependency at the promoter. To investigate this phenomenon, a construct in which the binding site from *icsB* was inserted twice in tandem into the *virB* promoter region was tested in the presence or absence of the VirB protein. Similar to results obtained to construct (D) as this hybrid promoter was not de-repressed in the presence of VirB protein but only in the absence of H-NS (Figure 4.3).

#### **4.2.3 VirB de-repression of *virF* is facilitated by a *cis*-acting VirB binding site.**

To determine the effect of increasing the concentration of VirB protein on a repressed *virF* promoter, a *virF-gfp* transcriptional fusion was constructed (Table 2.2). The region of the *virF* promoter placed upstream of *gfp* spanned -380 to +64 and included all of the *virF* DNA bound by H-NS and FIS. Once again, the experiments were conducted at a repressive temperature (30°C) and the intracellular concentration of the VirB protein were set using the arabinose inducible P<sub>BAD</sub> promoter. This region of the *virF* promoter maintained H-NS-mediated repression at 30°C in the absence of VirB, however upon addition of VirB protein the promoter was up-regulated to levels similar to that obtained in a *hns* mutant (Figure. 4.4 A). Shortening the promoter region to remove DNA upstream of the H-NS binding site (-278) did not inhibit VirB mediated de-repression of

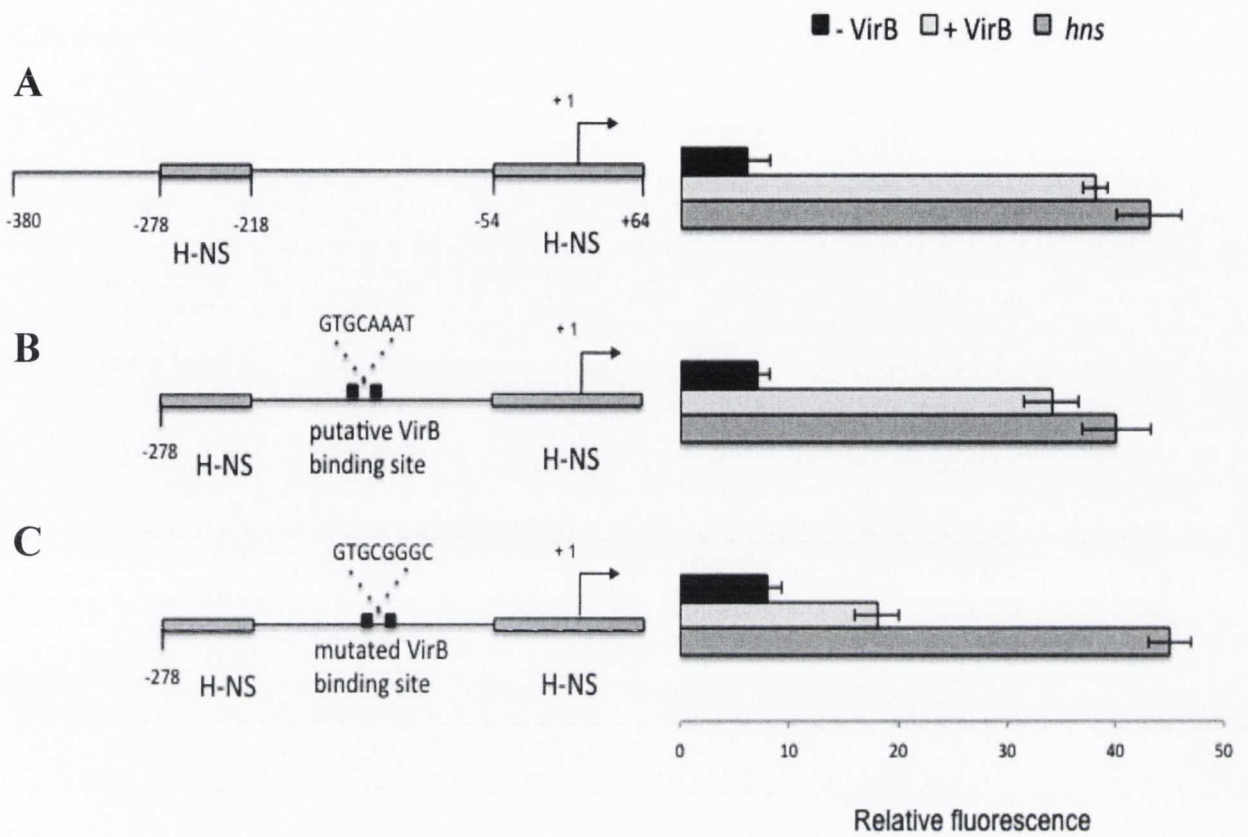


Figure 4.4. VirB-mediated de-repression of the *virF* promoter if facilitated by a *cis*-acting VirB binding site. Summaries of the structures of various derivatives of the *virF* promoter region are presented at left; levels of expression of the *virF-gfp* transcriptional reporter fusion are given at right. The native *virF* promoter encompassing positions -380 to +60 is shown in (A) the *virF* promoter from -280 to +30 shown in (B), *virF* promoter with mutations made to the putative VirB site (C)

the promoter (Figure. 4.4 B). This indicated that the potential VirB binding site either overlapped the H-NS binding sites or was located at a site within the 160 bp of DNA connecting the two H-NS sites. Inspection of this region revealed a close match to the consensus for VirB binding sites at position -134. The DNA sequence of the putative site, 5'- GTGCAAAT – 3', was a 7/8 match to the consensus sequence. This placed it almost exactly at the mid-point between the two H-NS sites. Mutation of this putative VirB binding site to the altered sequence 5'-GCTCGGGC–3' resulted in a reduced level of de-repression exerted by the VirB protein. Full de-repression was only detected in an *hns* mutant (Figure. 4.4 C).

#### **4.2.4 VirB interaction with the regulatory region of the *virB* and *virF* promoters.**

Bioinformatic analysis and assays of mutagenized promoters identified possible VirB binding sites in the regulatory regions of the *virB* and *virF* promoters.

DNase I footprinting experiments using an automated capillary DNA sequencer were performed to assess the possibility of specific binding of VirB protein to these regions of the *virB* and *virF* promoters. In this method, the DNA is amplified with a fluorescent dye, rather than a radioactive label, that can be easily detected with an automated fluorescent DNA analysis instrument that resolves oligonucleotides by capillary electrophoresis (Section 2.11). The *virF* and *virB* promoter regions were amplified using FAM labelled primers (Table 2.3) and incubated with or without purified VirB protein. These mixtures were then subjected to DNase I digestion. The reactions were stopped after 10 minutes and the fluorescently-labelled digested



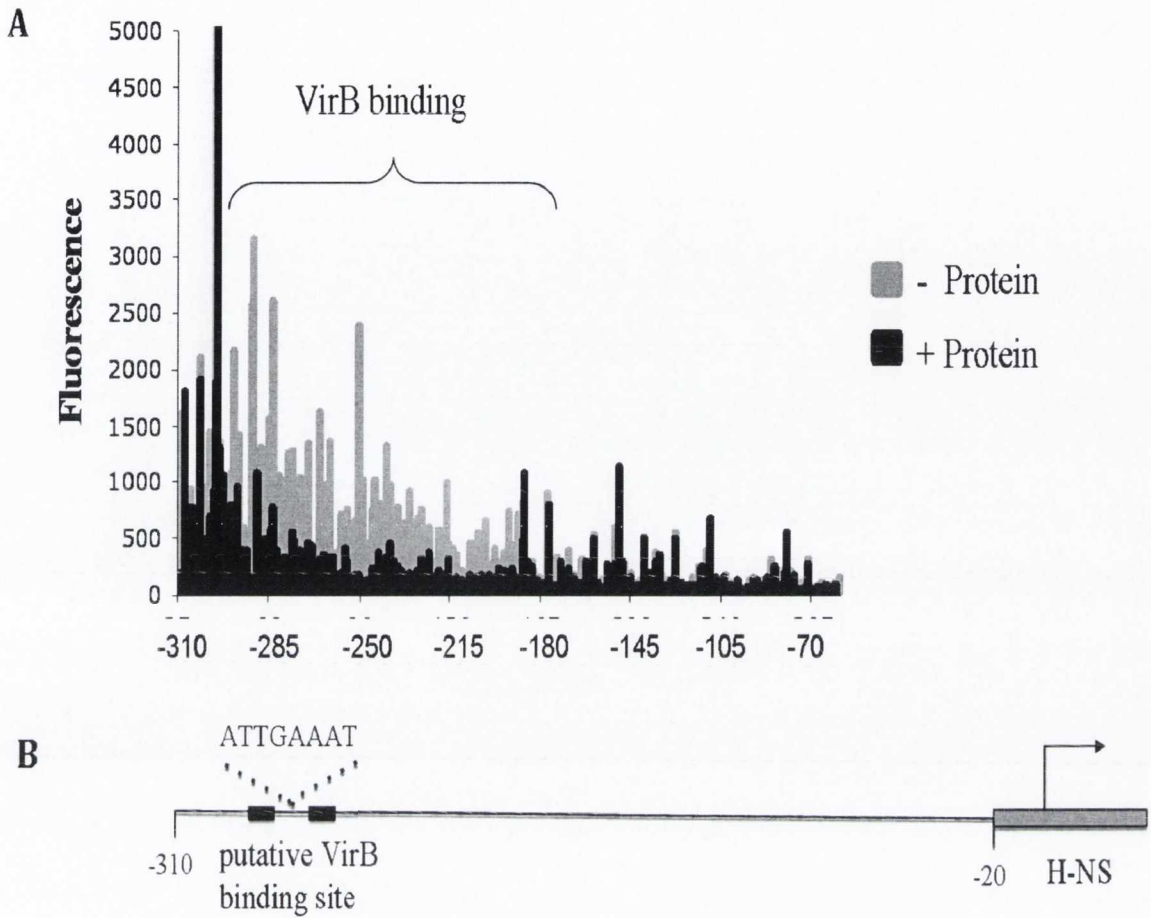


Figure 4.5. VirB requires a *cis*-acting sequence for efficient binding at *virB*. (A) Plot showing the protection pattern of the *virB* promoter region after digestion with DNaseI following incubation in the presence (black) or absence (grey) of VirB protein. The labelled VirB site shows an area with a significant difference in peak pattern. (B) Schematic representation of the promoter region analyzed is shown below the graph.

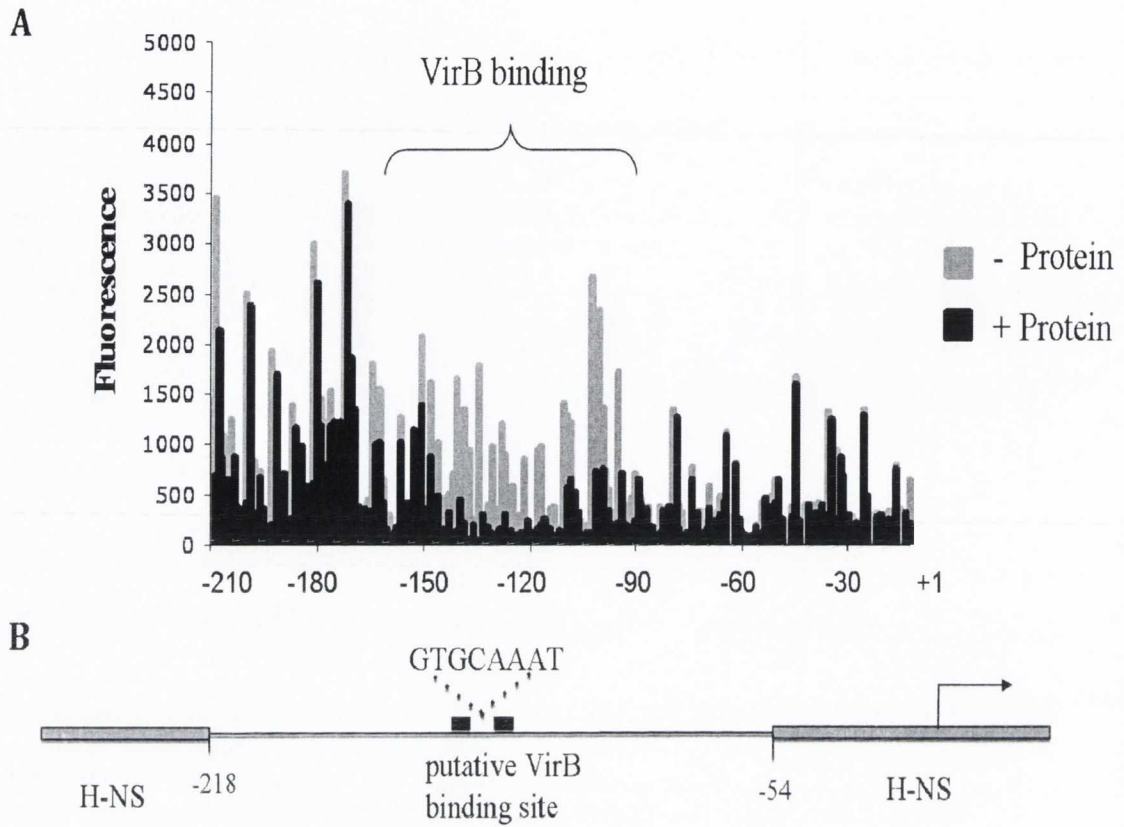


Figure 4.6. VirB requires a *cis*-acting sequence for efficient binding at *virF*. (A) Plot showing the protection pattern of the *virF* promoter region after digestion with DNaseI following incubation in the presence (black) or absence (grey) of VirB protein. The labelled VirB site shows an area with a significant difference in peak pattern. (B) Schematic representation of the promoter region analyzed is shown below the graph.

products were analysed. In the case of the *virB* regulatory region, an area of protection was detected between positions -280 and -180 (Figure 4.5 A). This location corresponded to the VirB binding site consensus sequence revealed by bioinformatics studies (Figure 4.5 B). For the *virF* region, an area of clearance associated with VirB binding was observed between positions -250 and -150, again corresponding to the location of the putative VirB binding site consensus sequence (Figure 4.6 A, B).

#### **4.2.5 VirB is required for the thermal induction of the *virB* and *virF* promoters.**

An increase in temperature, indicative of entry into the human host, is the key environmental signal resulting in an invasive phenotype in *S. flexneri* (Tobe *et al.*, 1991). The above data suggested that VirB could exert an effect at 30°C when over-expressed ectopically using an inducible promoter. It was possible that this VirB mediated de-repression of H-NS at the *virB* and *virF* promoters aided the up-regulation of these genes, and hence the virulence genes, following a thermal upshift from 30°C to 37°C.

To test this hypothesis, the wild type strain BS184 harbouring the *virB-gfp* fusion on the plasmid pZep08 and *virB* mutant strain CJD1018 harbouring the same *virB-gfp* fusion were grown overnight at 30°C. Cells were then diluted 1:100 and grown at 30°C until early exponential phase (OD<sub>600</sub> 0.1) when the cultures were divided and half shifted to a temperature of 37°C. Samples were taken at regular time points and the

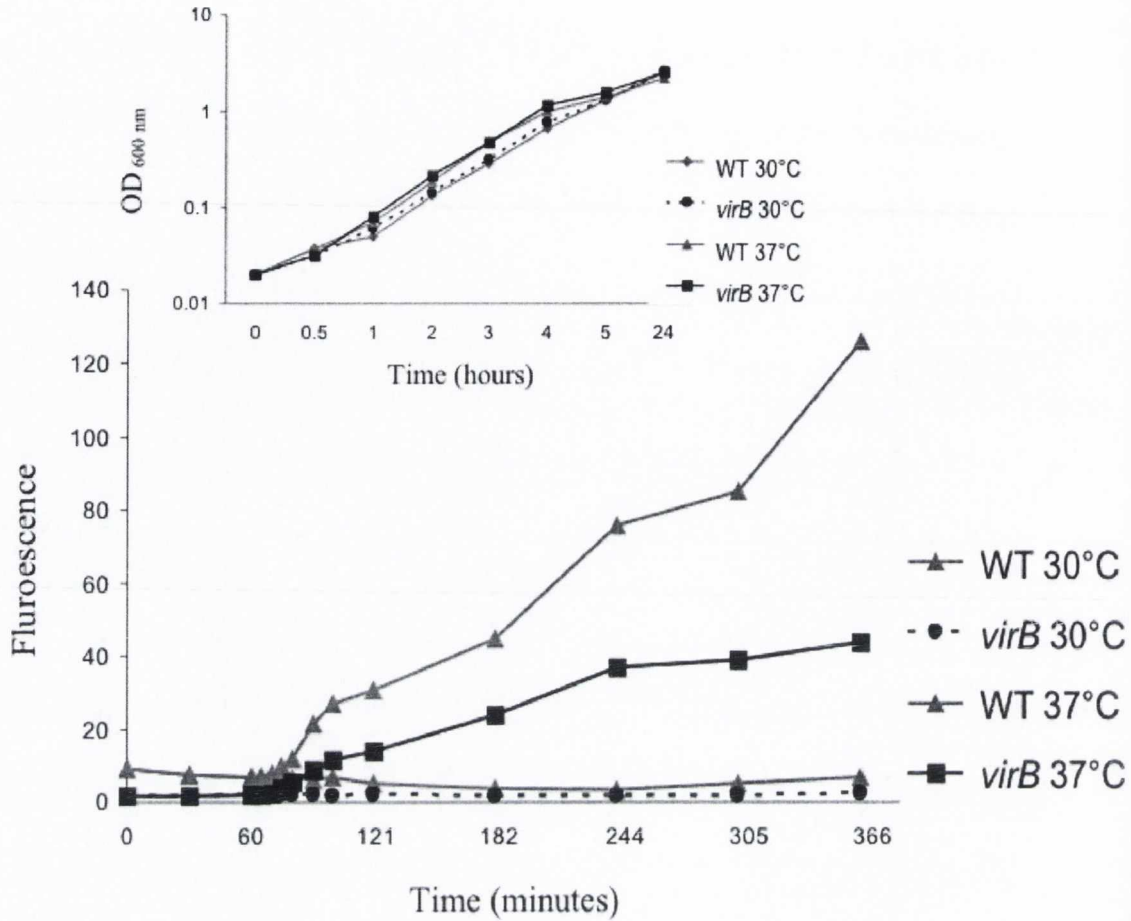


Figure 4.7. Thermal induction of the *virB* promoter. Expression of *virB-gfp* fusion in response to growth phase at 30°C or 37°C in a wild type (grey line) or *virB* mutant (dashed line) background. Samples were taken at regular intervals and measured for fluorescence. OD<sub>600</sub> data for each culture are represented in the graph above. The error bars are the standard deviation of three independent replicates.

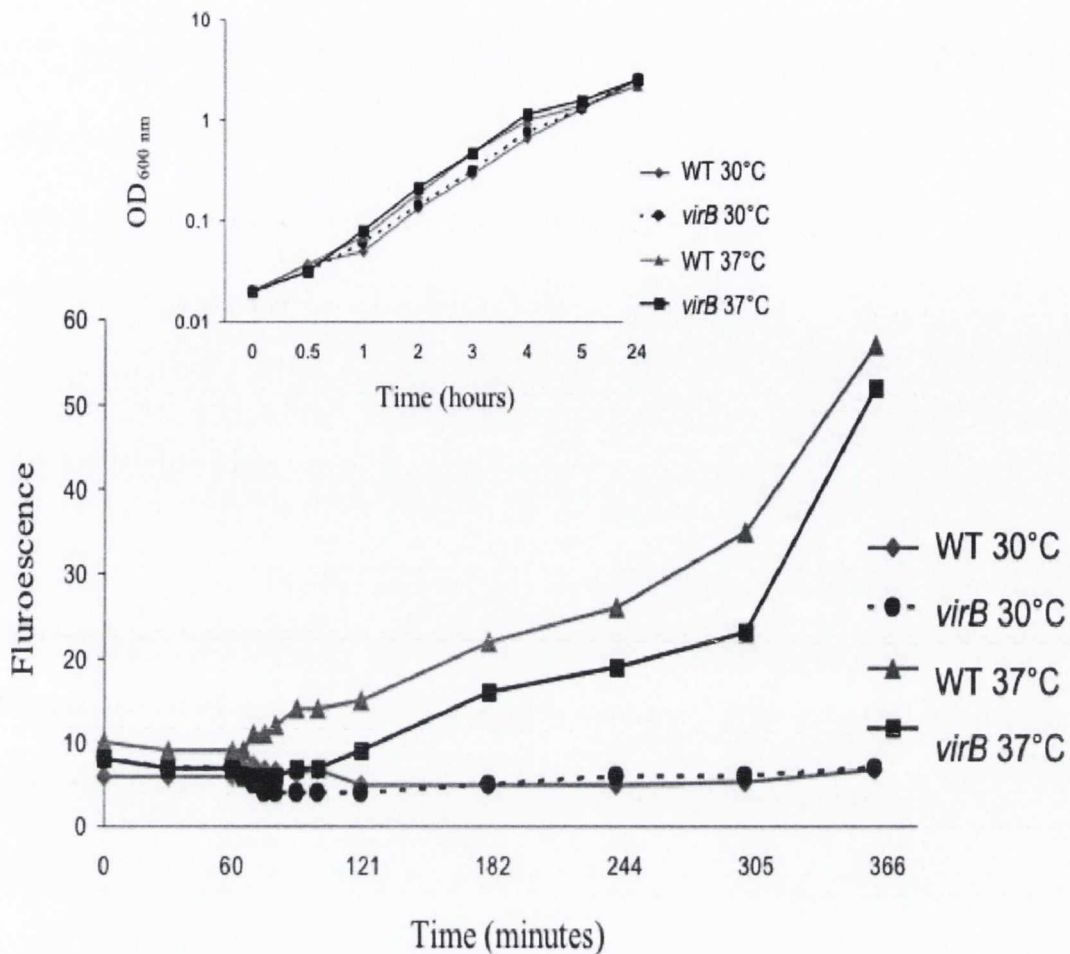


Figure 4.8 Thermal induction of the *virF* promoter. Expression of *virF-gfp* fusion in response to growth phase at 30°C or 37°C in a wild type (grey line) or *virB* mutant (dashed line) background. Samples were taken at regular intervals and measured for fluorescence. OD<sub>600</sub> data for each culture are represented in the graph above. The error bars are the standard deviation of three independent replicates

level of GFP fluorescence measured. It was found that in the *virB* mutant induction of the *virB* promoter was delayed and reduced compared to the induction of the *virB* promoter in a wild type background with a functional copy of the *virB* gene (Figure 4.7). Therefore VirB contributes to the full induction of the *virB* promoter in response to a thermal upshift, presumably by aiding in the removal of H-NS from the *virB* promoter region.

The experiments were also conducted with the *virF-gfp* promoter fusions in wild type and *virB* mutant backgrounds. The results showed that while the initial induction of the *virF* promoter was reduced in strains lacking the VirB protein, the ultimate level of *virF* transcription after 6 hours of growth was equal to that observed in the wild type. This suggested that while VirB may play a role in boosting the initial activation of transcription of *virF* in response to thermal upshift, this VirB dependency could be overcome by continued growth at the appropriate temperature (Figure. 4.8).

### 4.3 Discussion

Bacteria have evolved a diverse set of regulatory pathways that govern various adaptive responses to survive in rapidly changing environmental conditions (McAdams *et al.*, 2004). Feedback loops are common elements of cellular regulatory circuits (Cosentino *et al.*, 2007). Feedback regulation is critical for enzymatic control in archea, development in

eukaryotic organisms (Cavicchioli, 2007) and virulence in bacterial pathogens (Williams and Cotter, 2007) Positive feedback is thought to shape response timing to allow rapid expression of necessary genes when activated by certain stimuli (Shin *et al.*, 2006; Mitophanov *et al.*, 2010).

The regulatory cascade of *Shigella flexneri* is a model example of the integration of environmental cues to elicit the activation and assembly of essential virulence factors. While once thought to be a linear, step by step process, wherein *virF* activates *virB* which in turn activates the structural and effector protein genes, the data shown here offer a more dynamic view of transcriptional control. Quantitative Real Time PCR data demonstrate the ability of the VirB protein to up-regulate the H-NS repressed *virB* and *virF* promoters even during growth at the non-permissive temperatures of 30°C. These results were verified by plasmid-based experiments where varying lengths of the *virB* and *virF* promoters located upstream of the *gfp* reporter gene were monitored for VirB mediated H-NS de-repression.

Bioinformatic analysis of the *virB* and *virF* regulatory regions revealed in each case the presence of DNA sequences showing close matches to the consensus for the VirB binding site. Unlike the relatively straightforward regulation of the *proU* operon described in Chapter 3 which involved only repressor versus antagonist influences, both *virB* and *virF* regulation involve a more complex interplay between transcription factors.

The results from Chapter 3 suggest that the binding site for VirB must be positioned in the vicinity (within approximately 110 bp) of the H-NS-DNA repressive complex in order to exert an antagonistic effect. While this may hold true for a system in which removal of H-NS is sufficient to activate the promoter, the interaction between repressor and antagonist may be altered if additional transcription factors are required for promoter activity. In the case of the gene *virB*, the essential activator VirF binds immediately upstream of the H-NS binding sites (Figure 4.1). As VirF is absolutely necessary for transcription initiation at the *virB* promoter, evolutionary re-arrangements which have resulted in the insertion of a binding site for VirB at this location would be unfavourable. IHF, an architectural protein, binds upstream of the VirF binding site to a consensus sequence located at -175. IHF has also been shown to act positively on the transcription of *virB* (Porter and Dorman, 1997). The newly defined VirB binding site at *virB* is located upstream of the IHF site at position -280. Although seemingly at too great a distance from the H-NS binding site to displace it from the promoter, VirB at this location may act in co-operation with IHF to distort the local conformation of the DNA in order to undermine H-NS binding, allowing for activation by VirF.

The *virF* promoter experiences numerous influences, including those of H-NS and its antagonists, the positively acting FIS and IHF proteins. The newly-discovered putative VirB binding site is located at position -134, flanked by the two sites bound by H-NS. It is interesting to note that the consensus match to VirB binding sites is positioned midway between the two H-NS sites, with 80 bp of DNA extending either side of the VirB site. An obvious comparison may be made between the *virF* promoter and that of



*proU*, in which two spatially separated H-NS sites act co-operatively to form a repression nucleoprotein complex. In the case of *proU*, modifications to include a VirB binding site within the H-NS-DNA loop domain only partially de-represses the promoter (Figure 3.8). However, with the addition of the VirB binding site to *proU*, the distance between the interacting URE and DRE H-NS sites was increased to 270 bp. This allowed VirB-mediated de-repression of the adjacent URE but not full removal of H-NS at the DRE because of the distance constraints. In contrast to this, the VirB site located at *virF* is in close proximity to both H-NS binding sites, i.e. H-NS bound at the site overlapping the promoter and at the site further upstream (Figure 4.4). With VirB nucleation occurring at -134, and the protein known to coat about 100 bp of DNA (Turner and Dorman, 2007), this would result in VirB binding extending just to the cusp of H-NS occupied sites allowing de-repression to occur. It should be noted that two of the four sites bound by FIS are also located in this region. Whether the proteins are in direct competition for these sites remains unclear.

The relevance of this *virB* auto-regulation and positive feedback to *virF* becomes clear when the thermal induction kinetics of both promoters are observed. Without auto-regulation, the *virB* promoter has delayed responses to the environmental cue and maintains a reduced level of transcription. The *virF* promoter also displays a delay in activation, however this is eventually overcome, presumably by conformational changes in the DNA, which displace H-NS and permit activation by FIS. The intermediate regulatory step of *virB* expression in the regulatory cascade has been a matter for speculation. It is assumed that it provides a regulatory checkpoint prior to full

commitment to expression of the elaborate type III secretion system. However this additional regulation may impose a cost on the bacterium by postponing the assembly of the essential secretion system, while *virB* is de-repressed, transcribed and translated into a functioning antagonist of H-NS. Perhaps the purpose of positive feedback in the regulatory cascade is to ensure a transcriptional surge in *virB* expression, and to a lesser extent *virF*, to allow for a “jump start” in TTSS gene expression in response to appropriate environmental signals (Figure. 4.9).

The results also reveal the delicate nature of transcriptional control at these promoters. The artificial insertion of an additional binding site for VirB at the *virB* promoter, rather than enhancing its de-repressive effects, inhibited VirB from exerting a positive effect (Figure 4.3). This suggests that the DNA topological disturbances induced by VirB may make regions downstream unfavourable for further VirB binding. This highlights the importance of correct positioning of transcription factors with respect to one another at their target promoters.

Although the promoter regions of *virB* and *virF* revealed functioning VirB binding sites displaying close matches to the consensus nucleation sequence, neither region exhibited the extended nucleotide sequence similarity to the *parS* region of the P1 plasmid seen at the *icsB* promoter (Turner and Dorman, 2007). My original hypothesis assumes that this region of homology may have once originated in the vicinity of the *virB* gene and that this arrangement may have conferred auto-regulation. My findings however indicate that *virB* still retains this feature in the modern virulence plasmid. The most plausible

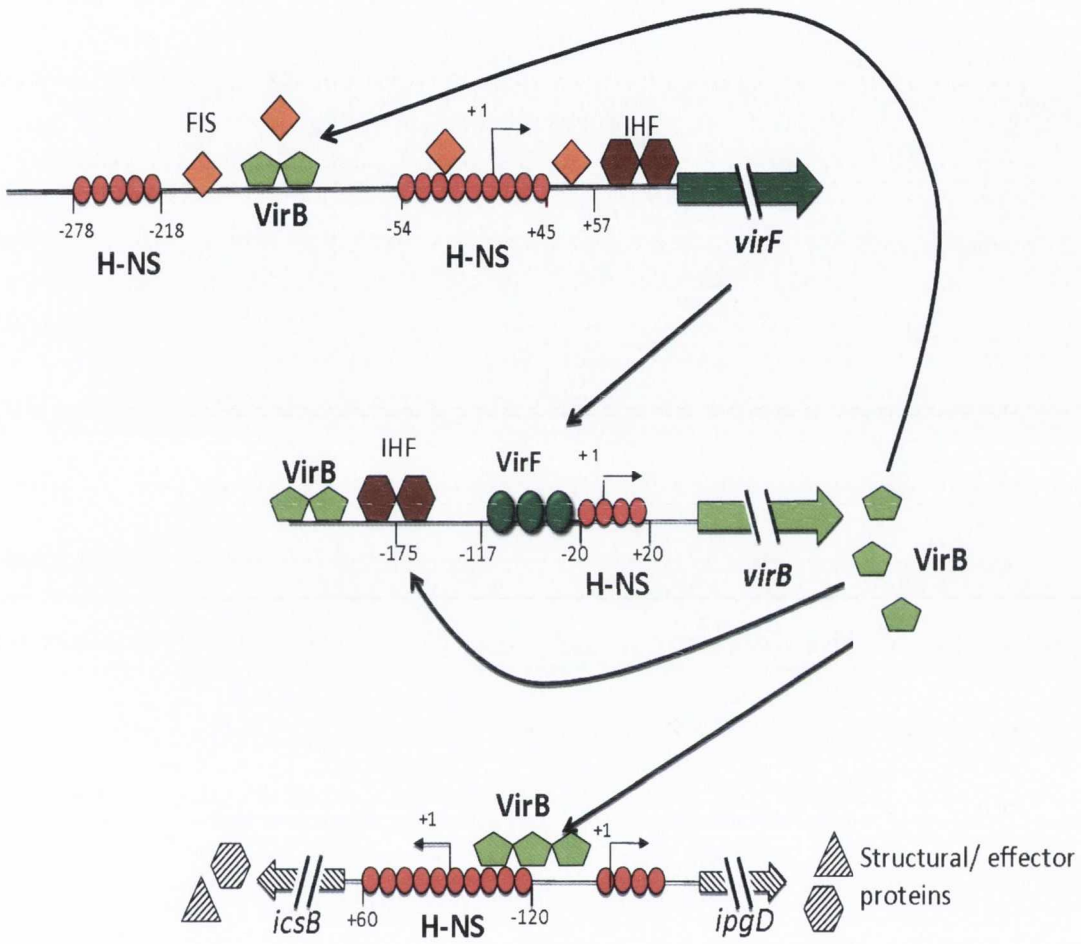


Figure 4.9. The revised schematic overview of the interplay of transcription factors in the virulence gene regulatory cascade. The VirB protein (green pentagons) binds to its own promoter acting displace H-NS and ensure maximal expression. The VirB protein also feedback back onto *virF*, binding to its promoter and facilitating H-NS removal.

explanation of these findings is that there is a selective benefit for VirB to exhibit positive feedback. This is apparent from the thermal induction experiments where expression is only maximal in the presence of VirB. Following disconnection of the VirB binding site from its gene and insertion upstream of *icsB*, through point mutation a close match to the VirB nucleation sequence evolved to restore positive feedback. So while a new regulatory switch emerged at *icsB*, subsequently an approximation to the ancestral switch was restored. It is tempting to speculate that as VirB became co-opted into a dedicated H-NS antagonist, mutations that allowed VirB-mediated derepression of other virulence genes, such as *icsP*, *virA*, and *spa15*, emerged.

The detailed dissection of the virulence regulatory cascade hints at the flexibility of bacterial gene-regulatory circuits and their capacity to adapt through evolution to meet the challenges of an ever-changing environment.

## **Chapter 5**

### **General Discussion**

## 5.1. General Discussion

Gene regulatory networks are considered the most important organizational level in the cell allowing signals from the cell surrounding environment to be integrated into coordinated activation and inhibition of genes (Crombach and Hogeweg, 2008).

Many lines of evidence suggest the plasticity of the bacterial genome, i.e. that its content and organization is highly changeable (McAdams *et al.*, 2004). Genes and parts of genes can move within and between genomes, and transposable regions of DNA that encode many genes are transferred within genomes and among cells by mobile genetic elements (Ochman *et al.*, 2000). The processes that facilitate this plasticity occur randomly with the usual rate of spontaneous point mutations in the range of  $10^{-9}$ – $10^{-10}$  mutations/cell/generation. This slow rate could be assumed that almost all bacteria yield identical offspring, however their relatively short generation times has allowed extensive exploration of possible bacterial genomic arrangements. For rapid bacterial evolution, new molecular functions and new regulatory pathways arise from the shuffling and disruption of operons imposed by horizontal gene transfer (HGT). Other large scale events such as large sections of the chromosome becoming duplicated, which can provide extra copies of a gene into a genome. These extra copies are considered an important source of “raw material” for new genes to evolve (Carroll *et al.*, 2005).

Point mutation is probably the main means by which bacteria achieve fine-tune wiring,

including adjustment of their kinetic parameters and individual DNA-binding sites. Although the DNA-binding transcription factors are drawn from a small set of ancient conserved families, their relative abundance often shows dramatic variation among different phylogenetic groups (Babu *et al.*, 2004). Their binding sites may undergo several types of mutations as well including duplication: a binding site is copied to the upstream region of a random gene in the genome hence introducing a new connection in the network. Similarly a deletion event may occur, whereby multiple connections may be lost in the resulting network. Insertion events may also arise in which large regions of DNA may insert in the vicinity of a transcription factor site, thereby displacing it from its original location (Crombach and Hogeweg, 2008).

Modular functions are frequently more highly conserved among organisms than the regulatory networks that control their expression. For example, the structure, organization and assembly of the complex machinery of the bacterial flagellum are highly similar even in distantly related bacteria (Aldridge *et al.*, 2002). However, the control of expression that determines when and where the flagella are constructed can differ greatly. The same is true for the type III secretion system. While the modular structure remains relatively constant and certain regulatory motifs are conserved, the networks that control expression can vary between organisms (Section 1.9). These examples illustrate the ability of bacterial networks to adapt to particular environmental niches, enabling opportunistic use of essentially conserved structures.

The *Shigella* plasmid-borne virulence system is subject to complex control, reflecting the need for the bacterium to restrict its expression to only appropriate niches with the host. The regulatory cascade of events operates on the backdrop of overwhelming transcriptional repression imposed by the H-NS protein on the chromosome. This repressive activity extends to every virulence gene promoter yet investigated including those of the transcriptional activators. Therefore many of the genetic switches regulating virulence are involved in antagonism of this H-NS-mediated repression.

The evolution and mechanisms of antagonizing H-NS repression has been investigated in detail. Although H-NS is implicated in the regulation of many “house-keeping” genes such as *proU*, recent studies have focused on a key role of H-NS in the selective silencing of large numbers of horizontally-acquired H-NS genes. At the heart of the xenogeneic silencing hypothesis is the assumption that the cell benefits initially from the downregulation of the transcription of horizontally acquired genes by the H-NS protein allowing them to become incorporated into the genome in a largely inert state, and presumably, an associated reduction in the competitive fitness of the bacterium (Dorman, 2007b; Lucchini *et al.*, 2006; Navarre *et al.*, 2007). However, genetic information acquired by horizontal transfer often contain elements beneficial to survival such as virulence factors, like type III secretion systems, and antibiotic resistance mechanisms. This suggests that the bacteria must evolve tools to counteract the transcription silencing activity of H-NS in ways that allow the genes to be expressed for the benefit the microbe. The regulatory systems recruited for this purpose have many distinct solutions to the problem of H-NS repression, recruiting antagonists from a wide



repertoire of transcription factors families. The dilemma of H-NS silencing serves to tell us much about the flexibility and evolvability of gene regulatory circuits in bacteria and their capacity for future development.

The *Shigella* virulence plasmid, thought to be horizontally acquired, is an interesting case of anti repression of H-NS, encoding it's own specific activators to counteract silencing. The VirB activator provides an unusual example of H-NS antagonism. VirB shows strong amino acid sequence homology to plasmid partitioning proteins such as ParB from plasmid/phage P1 and SopB from the F plasmid (Beloin *et al.*, 2002). The DNA sequence to which it binds resembles the parS sequences that are bound by the plasmid partitioning proteins (Taniya *et al.*, 2003; Turner & Dorman, 2007). In vitro transcription assays show that VirB does not act as a conventional transcription factor, recruiting RNA polymerase to the promoter and/or enhancing the rate of formation of open transcription complexes. Instead, VirB acts to antagonize the repressive activity of the H-NS protein. While VirB has no role in the partitioning of the modern virulence plasmid, it would seem to provide an *ad hoc* solution to H-NS repression. Due to the breakable nature of the H-NS repressive bridge complex, a DNA-bending protein or DNA-wrapping protein that competes for the same DNA sequence might make excellent antagonists.

By comparison with the closely related *E.coli*, it is apparent that many gene functions have become unnecessary through evolution. *Shigella* does not ferment lactose, presumably because the trait is of little use in the intercellular lifestyle. Similarly, since

*Shigella* uses actin-assisted motility to move intracellularly, flagella-based motility becomes redundant for much of its life cycle. A large number of genes are involved in production of the flagellum and associated machinery, and turning it off would conserve significant resources. This evolutionary adaptation is essential for the survival of the bacterium. Perhaps the existence of two fully functioning plasmid partitioning proteins, resulted in VirB becoming redundant in terms of its partitioning function however allowing it to become available for other task and re-deployed into a regulatory role.

This was facilitated by the fortuitous placement of an appropriate *cis*-acting binding site in the regulatory region of an H-NS repressed gene. These genetic rearrangements are analogous with events occurring routinely in bacterial, leading to the emergence of new regulatory switches. Clearly, the example of VirB shows it is important to consider proteins in the context of the molecular details and not to fall into the trap of assigning unique biological properties to homologous DNA-binding proteins.

The work here sought to investigate how easily this target specific antagonism of H-NS could evolve. The hypothetical evolutionary step, whereby H-NS was antagonized by judicious placement of a binding site for a DNA-binding and distorting protein, was replicated in the lab using various molecular techniques. By introducing the VirB binding site from *icsB* gene into the regulatory region of the H-NS repressed *proU* promoter, we successfully de-repressed expression under usually non-permissive low salt conditions. We showed this up-regulation was dependent upon direct binding of VirB, in an appropriate location with respect to H-NS binding. We propose that this

position is subject to the special constraints in which the VirB protein can only distort DNA in the vicinity of its binding site, hence requiring co-localization of VirB and H-NS on the DNA. This study served to support our initial hypothesis that, creation of new regulatory circuits, due to the chance placement of protein binding sites, can create novel regulatory switches, and perhaps lead to the co-option of unrelated proteins. However, the results also illustrate that fine-tuning would be required for this regulatory network innovation to be successful. Correct positioning of the binding sites for the antagonist and repressor is essential for efficient de-repression to occur. The URE has also been shown to be a multifunctional component of the regulation of *proU*. While acting as a repressor, acting co-operatively with the DRE to form a silencing nucleoprotein complex, it also acts to protect *proU* expression from read-through transcription from neighboring promoters, or promoters that may arise through evolution in the future. This additional “barrier” function of the upstream H-NS binding sites also portrays, like VirB, the dangers of assuming unique biological properties on DNA binding proteins.

The essential feature of autoregulation is the direct modulation of gene expression by the product of the same gene. Thus, a protein controls the rate of synthesis of additional copies of itself. Autoregulation is not a new concept in genetics; the phenomenon is newly appreciated as a mechanism of regulation of genes whose products perform critical cellular functions. Approximately 59% of transcription factors in *Escherichia coli* regulate the transcription rate of their own gene (Keseler *et al.*, 2005). This suggests that auto-regulation has one or several important functions. These functions include a

mechanism for amplification of gene expression (Kourilsky and Gros, 1974), regulation of proteins involved in DNA replication (Braun *et al.*, 1985), buffering the response of genes to environmental changes (Hagen and Magasanik, 1973), and maintaining a constant intracellular concentration of a gene product, independent of cell size, cell growth rate and changes in gene copy number (Sompayrac and Maaloe, 1973).

Auto-regulating transcription factors are typically regulated by other TFs too. In fact, 23 auto-regulating TFs in *E. coli* are known to respond to at least *two* additional regulators (Keselar *et al.*, 2005). Studies investigating the PhoP/PhoQ two-component system in *Salmonella* has shown positive feed loops to promoter transcriptional surges that jump-start the *Salmonella* virulence circuit (Shin *et al.*, 2006). It was shown to ensure a timely response to inducing environmental signals. The initial activation may allow the immediate establishment of a new phenotypic state. Strains deficient in this autoregulation are avirulent making it an essential feature of *Salmonella*'s ability to cause lethal infection.

The study of the regulation of *virF* and *virB* revealed additional, previously uncharacterized, connections in the regulatory network. Under repressive 30°C conditions, both the *virF* and *virB* promoter could be de-repressed by the over-expression of VirB protein *in trans*. The VirB protein was shown to bind directly to the promoter to initiate its antagonistic effect. This would suggest that the regulatory cascade whereby VirF activates VirB, which activates the virulence operons, once

thought to be a strictly linear process, possess a more dynamic mechanism of positive feedback from its intermediate regulator VirB.

While the relationship between repressor and anti-repressor at the modified *proU* promoters could be easily dissected, the regulation of *virF* and *virB* involves a complex interplay of multiple transcription factors acting at these promoters.

At the *virF* promoter, H-NS binds at two sites one overlapping the promoter and one further upstream. These two sites flank the region to which FIS binds, a transcriptional activator of *virF*. The putative VirB binding site is located almost exactly mid-point between the two H-NS filaments, i.e. within the bridge formation. Although the modified *proU* promoter which contained a VirB binding within the two H-NS site, only displayed partial VirB mediated de-repression, this is likely due to the differing architecture of the *virF* promoter differs. This places the site about 80 bp from the bordering H-NS proteins, and within reach of the repressor. Also, unlike *proU* which can still retain repression with the downstream H-NS bound, repression of the *virF* promoter is strictly dependent on the URE and DRE elements being intact and functional. Therefore, unlike *proU*, any disruption of one element would have detrimental consequences for overall H-NS repression of the promoter. The proposed model for the de-repression of the *virF* promoter would be; DNA distortion brought about by a thermal upshift, which acts to disrupt H-NS bridging which is possibly facilitated by the binding and DNA wrapping action of VirB. This is followed by, the binding of FIS, recruitment of RNA polymerase and transcription initiation.

At the *virB* promoter, studies have shown the removal of H-NS from the promoter involves the activator VirF which binds directly upstream of H-NS. IHF also imparts a positive effect binding upstream of VirF at the promoter. This study has now shown *virB* binds upstream of IHF at the regulatory region of *virB*. Although, the work here with *proU* would suggest VirB at this location is at too far a distance to exert its effects on H-NS, VirB could act to distort the DNA to create more favorable binding for the upstream IHF and VirF proteins, which together with topological changes in the DNA induced by a thermal upshift, acts to displace H-NS.

The coordinated positive regulation of promoter expression by an AraC-like protein and VirB is not unusual in regulation of the virulence genes in *Shigella*. VirB and VirF both act positively at the *icsP* to alleviate H-NS repression and activate the promoter (Wing *et al.*, 2004). Expression of *phoN2*, *ospF*, *ospC1* and *virA* was decreased in the *virB* mutant compared with the wild-type strain, and also found to be regulated by MxiE (an AraC-like protein) (Kane *et al.*, 2002). The co-regulation of a repressed promoter by an anti-repressor and a conventional transcription activator is a common theme at bacterial promoters.

Thermal induction of both the *virF* and *virB* promoters in the absence of VirB protein resulted in a loss of the initial surge of transcription post temperature change. This suggests that positive loops in the virulence cascade are necessary to ensure a timely response to environmental cues, and immediate expression. Perhaps this counteracts the

additional intermediate step in regulation of the virulence genes provided by VirB. While this “check point” reduces wasteful expression, it must also slow the response of the bacteria to induction signals. While *virF* promoter could eventually recover expression levels to that of wild type after several hours at 37°C, the *virB* promoter was permanently down regulated. A reduction in VirB protein in the cell would have a knock-on effect on the de-repression of the structural and effector genes, and hence on the construction of the type III secretion system. This indicates that autoregulation and positive feed-back are crucial for full virulence in *Shigella*.

The adaptability of VirB as an antirepressor is reflected in the number of promoters that this protein regulates on the *S. flexneri* virulence plasmid. Those binding sites that have been studied have shown similarity to the minimal *parS* element that is bound by ParB-like proteins. This is a portion of the full-length *parS*-like motif that is found between the *icsB* and *ipgD* promoters and is approximately equivalent to the region at the box 1/box 2 motif. Based on this study of *proU*, it can be hypothesized that insertion of the *icsB* operon disconnected the full *parS*-like VirB binding site from its gene on the virulence plasmid. This facilitated the co-option of VirB into a regulatory role as an HNS antagonist at the *icsB-ipgD* promoter region. Investigation of the regulation of *virB* and *virF* has indicated that this hypothetical evolutionary step resulted the regeneration of a partial binding site (containing box1/box2 motif but lacking the extended *parS*-like features) (Figure 1.4) sufficient for VirB binding at the *virB* promoter. This autoregulation was selected for as it promoted efficient de-repression of the gene.

Possibly through point mutation, other H-NS genes, including *virF*, acquired similar partial binding sites to assist in the de-repression their genes.

While the VirB binding site at the *icsB-ipgD* intergenic region displays extended homology to the full *parS* site on P1 or P7 (Figure 1.4), it can be argued that these similarities are not significant when compared to “partial” VirB binding sites located at other promoters. The identification of a functional VirB binding site at the *virB* promoter allows for alternative hypothetical evolutionary steps to be explored. It is possible that the VirB binding site at *icsB* did not originate from a ‘disconnection step’ but rather point mutations in the DNA that resulted in the formation of a VirB site that co-opted the VirB protein into a role at this promoter, which aided in de-repression. If we assume this is true, then the VirB site at *virB* (required for autoregulation) remained in the *virB* promoter region and was not removed and then regenerated. A full bioinformatic study of all VirB binding site’s on the virulence plasmid (showing homology to full or minimal *parS* sequences) would help in our understanding of how VirB came to be re-directed into a regulatory role. A search for similar sequences on the *Shigella* chromosome may also reveal some interesting results. If the gene on the virulence plasmid can co-opt VirB to remove H-NS from it’s promoters, then H-NS repressed chromosomal genes may also benefit from acquiring a VirB site.



## References

- Abe, A., Matsuzawa, T., Kuwae, A.** (2005) Type-III effectors: sophisticated bacterial virulence factors. *C R Biol* **328**(5):413-28.
- Abele, A. L., Friedman, S. A., Austin, S. J.** (1985) Partition of unit-copy miniplasmids to daughter cells. III. The DNA sequence and functional organization of the P1 partition region. *J Mol Biol* **20**;185(2):261-72.
- Adam, T., Arpin, M., Prévost, M. C., Gounon, P., Sansonetti, P. J.** (1995) Cytoskeletal rearrangements and the functional role of T-plastin during entry of *Shigella flexneri* into HeLa cells. *J Cell Biol* **129**(2):367-81.
- 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):627-35.
- Aldridge, P., Hugues, K. T.** (2002) Regulation of flagellar assembly. *Curr Opin Microbiol* **5**(2):160-5
- Al Mamun, A. A., Tominaga, A., Enomoto, M.** (1997) Cloning and characterization of the region III flagellar operons of the four *Shigella* subgroups: genetic defects that cause loss of flagella of *Shigella boydii* and *Shigella sonnei*. *J Bacteriol* **179**(14):4493-500.

**Al Mamun, A. A., Tominaga, A., Enomoto, M.** (1996) Detection and characterization of the flagellar master operon in the four *Shigella* subgroups. *J Bacteriol* **178**(13):3722-6.

**Aliprantis, A. O., Weiss, D. S., Radolf, J. D., Zychlinsky, A.** (2001) Release of Toll-like receptor-2-activating bacterial lipoproteins in *Shigella flexneri* culture supernatants. *Infect Immun* **69**(10):6248-55.

**Ashida, H., Toyotome, T., Nagai, T., Sasakawa, C.** (2007) *Shigella* chromosomal IpaH proteins are secreted via the type III secretion system and act as effectors. *Mol Microbiol* **63**(3):680-93.

**Babu, M. M., Luscombe, N. M., Aravind, L., Gerstein, M., Teichmann, S. A.** (2004) Structure and evolution of transcriptional regulatory networks. *Curr Opin Struct Biol* **14**(3):283-91.

**Badaut, C., Williams, R., Arluison, V., Bouffartigues, E., Robert, B., Buc, H., Rimsky, S.** (2002) The degree of oligomerization of the H-NS nucleoid structuring protein is related to specific binding to DNA. *J Biol Chem* **277**(44):41657-66.

**Ball, C. A., Osuna, R., Ferguson, K. C., Johnson, R. C.** (1992) Dramatic changes in Fis levels upon nutrient upshift in *Escherichia coli*. *J Bacteriol* **174**(24):8043-56.

**Battistuzzi, F. U., Feijao, A., Hedges, S. B.** (2004) A genomic timescale of prokaryote evolution: insights into the origin of methanogenesis, phototrophy, and the colonization of land. *BMC Evol Biol* **9**;4:44.

**Beloin, C., Dorman, C. J.** (2003) An extended role for the nucleoid structuring protein H-NS in the virulence gene regulatory cascade of *Shigella flexneri*. *Mol Microbiol* **47**(3):825-38.

**Beloin, C., McKenna, S., Dorman, C. J.** (2002) Molecular dissection of VirB a key regulator of the virulence cascade of *Shigella flexneri*. *J Biol Chem* **277**(18):15333-44.

**Berg, J., Willmann, S., Lässig, M.** (2004) Adaptive evolution of transcription factor binding sites. *BMC Evol Biol* **28**;4:42.

**Bernardini, M. L., Fontaine, A., Sansonetti, P. J.** (1990) The two-component regulatory system *ompR-envZ* controls the virulence of *Shigella flexneri*. *J Bacteriol* **172**(11):6274-81.

**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**(10):3867-71.

**Bernardini, M. L., Sanna, M. G., Fontaine, A., Sansonetti, P. J.** (1993) OmpC is involved in invasion of epithelial cells by *Shigella flexneri*. *Infect Immun* **61**(9):3625-35.

**Blatter, E. E., Ross, W., Tang, H., Gourse, R. L., Ebright, R. H.** (1994) Domain organization of RNA polymerase alpha subunit: C-terminal 85 amino acids constitute a domain capable of dimerization and DNA binding. *Cell* **78**(5):889-96.

**Bouet, J. Y., Funnell, B. E.** (1999) P1 ParA interacts with the P1 partition complex at *parS* and an ATP-ADP switch controls ParA activities. *EMBO J* **18**(5):1415-24.

**Bouffartigues, E., Buckle, M., Baudet, C., Travers, A., Rimsky, S.** (2007) H-NS cooperative binding to high-affinity sites in a regulatory element results in transcriptional silencing. *Nat. Struct. Mol. Biol* **14**:441-448.

**Bradford, M. M.** (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem* **72**:248-54.

**Braun, R. E., O'Day, K. and Wright, A.** (1985) Autoregulation of the DNA replication gene *dnaA* in *E. coli* K-12. *Cell* **40**:159.

**Brinkmann, V., Reichard, U., Goosmann, C., Fauler, B., Uhlemann, Y., Weiss, D. S., Weinrauch, Y., Zychlinsky, A.** (2004) Neutrophil extracellular traps kill bacteria.

*Science* **303**(5663):1532-5.

**Brown, N. L., Stoyanov, J. V., Kidd, S. P., Hobman, J. L.** (2003) The MerR family of transcriptional regulators. *FEMS Microbiol Rev* **27**(2-3):145-63.

**Buchrieser, C., Glaser, P., Rusniok, C., Nedjari, H., D'Hauteville, H., Kunst, F., Sansonetti, P., Parsot, C.** (2000) The virulence plasmid pWR100 and the repertoire of proteins secreted by the type III secretion apparatus of *Shigella flexneri*. *Mol Microbiol* **38**(4):760-71.

**Buckler, D. R., Anand, G. S., Stock, A. M.** (2000) Response-regulator phosphorylation and activation: a two-way street? *Trends Microbiol* **8**(4):153-6.

**Caron, J., Coffield, L. M., Scott, J. R.** (1989) A plasmid-encoded regulatory gene, *rns*, required for expression of the CS1 and CS2 adhesins of enterotoxigenic *Escherichia coli*. *Proc Natl Acad Sci U S A* **86**(3):963-7.

**Carroll, S. B., Grenier, J., Weatherbee, S. D.** (2005). From DNA to Diversity: Molecular Genetics and the Evolution of Animal Design. Second Edition. Oxford: Blackwell Publishing.

**Castellanos, M. I., Harrison, D. J., Smith, J. M., Labahn, S. K., Levy, K. M., Wing, H. J.** (2009) VirB alleviates H-NS repression of the *icsP* promoter in *Shigella flexneri*

from sites more than one kilobase upstream of the transcription start site. *J Bacteriol* **191**(12):4047-50.

**Cavicchioli, R., ed.** (2007). *Archaea: Molecular and Cellular Biology*. ASM Press, Washington, DC

**Chirgwin, J.M.** (1979) Isolation of biologically active ribonucleic acid from sources enriched in ribonuclease. *Biochemistry* **18**:5294–9.

**Choy, H. E., Adhya, S.** (1992) Control of gal transcription through DNA looping: inhibition of the initial transcribing complex. *Proc Natl Acad Sci U S A* **89**(23):11264-8.

**Collazo, C. M., Galán, J. E.** (1997) The invasion-associated type-III protein secretion system in *Salmonella*. *Gene* **192**(1):51-9.

**Cooper, V. S., and R. E. Lenski.** (2000) The population genetics of ecological specialization in evolving *Escherichia coli* populations. *Nature* **407**:736–739.

**Cosentino Lagomarsino, M., Jona, P., Bassetti, B., Isambert, H.** (2007) Hierarchy and feedback in the evolution of the *Escherichia coli* transcription network. *Proc Natl Acad Sci U S A* **104**(13):5516-20.

**Crombach, A., Hogeweg, P.** (2008) Evolution of evolvability in gene regulatory networks. *PLoS Comput Biol.* **4**(7)

**Day, W. A. Jr., Maurelli, A. T.** (2001) *Shigella flexneri* LuxS quorum-sensing system modulates virB expression but is not essential for virulence. *Infect Immune* **69**(1):15-23.

**De Geyter, C., Wattiez, R., Sansonetti, P., Falmagne, P., Ruyschaert, J. M., Parsot, C., Cabiliaux, V.** (2000) Characterization of the interaction of IpaB and IpaD, proteins required for entry of *Shigella flexneri* into epithelial cells, with a lipid membrane. *Eur J Biochem* **267**(18):5769-76.

**Dorman, C. J., Kane, K. A.** (2009) DNA bridging and antibridging: a role for bacterial nucleoid-associated proteins in regulating the expression of laterally acquired genes. *FEMS Microbiol Rev* **33**(3):587-92.

**Dorman, C. J., Ni Bhriain, N., Higgins, C. F.** (1990) DNA supercoiling and environmental regulation of virulence gene expression in *Shigella flexneri* *Nature*. **344**(6268):789-92.

**Dorman, C. J., Porter, M. E.** (1998) The *Shigella* virulence gene regulatory cascade: a paradigm of bacterial gene control mechanisms. *Mol Microbiol* **29**(3):677-84.

**Dorman, C. J.** (2006) DNA supercoiling and bacterial gene expression. *Sci Prog* **89**(Pt 3-4):151-66

**Dorman, C. J.** (2004) H-NS: a universal regulator for a dynamic genome. *Nat Rev Microbiol* **2**(5):391-400.

**Doyle, M., Fookes, M., Ivens, A., Mangan, M. W., Wain, J., Dorman, C. J.** (2007) An H-NS-like stealth protein aids horizontal DNA transmission in bacteria. *Science* **315**(5809):251-2.

**DuPont, H. L., Levine, M. M., Hornick, R. B., Formal, S. B.** (1989) Inoculum size in shigellosis and implications for expected mode of transmission. *J Infect Dis* **159**(6):1126-8.

**Durand, J. M., Björk, G. R.** (2003) Putrescine or a combination of methionine and arginine restores virulence gene expression in a tRNA modification-deficient mutant of *Shigella flexneri*: a possible role in adaptation of virulence. *Mol Microbiol* **47**(2):519-27.

**Durand, J. M., Dagberg, B., Uhlin, B. E., Björk, G. R.** (2000) Transfer RNA modification, temperature and DNA superhelicity have a common target in the regulatory network of the virulence of *Shigella flexneri*: the expression of the *virF* gene. *Mol Microbiol* **35**(4):924-35.



**Ebright, R. H.** (1993) Transcription activation at Class I CAP-dependent promoters.

*Mol Microbiol* **8**(5):797-802.

**Edwards, B. H.** (1999) *Salmonella* and *Shigella* species. *Clin Lab Med* **19**(3):469-87.

**Egile, C., d'Hauteville, H., Parsot, C., Sansonetti, P. J.** (1997) SopA, the outer membrane protease responsible for polar localization of IcsA in *Shigella flexneri*. *Mol Microbiol* **23**(5):1063-73.

**Escobar-Páramo, P., Giudicelli, C., Parsot, C., Denamur, E.** (2003) The evolutionary history of *Shigella* and enteroinvasive *Escherichia coli* revised. *J Mol Evol* **57**(2):140-8.

**Eustance, R. J. , Schleif, R. F.** (1996) The linker region of AraC protein. *J Bacteriol* **178**(24):7025-30.

**Falconi, M., Colonna, B., Prosseda, G., Micheli, G., Gualerzi, C. O.** (1998) Thermoregulation of *Shigella* and *Escherichia coli* EIEC pathogenicity. A temperature-dependent structural transition of DNA modulates accessibility of virF promoter to transcriptional repressor H-NS. *EMBO J* **17**(23):7033-43.

**Falconi, M., Prosseda, G., Giangrossi, M., Beghetto, E., Colonna, B.** (Involvement of FIS in the H-NS-mediated regulation of virF gene of *Shigella* and enteroinvasive *Escherichia coli*. *Mol Microbiol* **42**(2):439-52.

**Falkow, S.** (1997) The evolution of the pathogenicity in *Escherichia*, *Shigella* and *Salmonella*, p. 2723-2729. In F. C. Neidhardt et al. (ed.), *Escherichia coli* and *Salmonella*, vol. 2. ASM Press, Washington, DC.

**Filutowicz, M., Ross, W., Wild, J., Gourse, R. L.** (1992) Involvement of Fis protein in replication of the *Escherichia coli* chromosome. *J Bacteriol* **174**(2):398-407.

**Finkel, S. E., Johnson, R. C.** (1993) The Fis protein: it's not just for DNA inversion anymore. *Mol Microbiol* **7**(6):1023.

**Finlay, B. B., Ruschkowski, S., Dedhar, S.** (1991) Cytoskeletal rearrangements accompanying salmonella entry into epithelial cells. *J Cell Sci* **99** (Pt 2):283-96.

**Fletcher, S. A., Csonka, L. N.** (1995) Fine-structure deletion analysis of the transcriptional silencer of the *proU* operon of *Salmonella typhimurium*. *J. Bacteriol* **177**:4508-4513.

**Formal, S. B., Hornick, R. B.** (1978) Invasive *Escherichia coli*. *J Infect Dis* **137**(5):641-4.

**Freundlich, M., Ramani, N., Mathew, E., Sirko, A., Tsui, P.** (1992) The role of integration host factor in gene expression in *Escherichia coli*. *Mol Microbiol* **6**(18):2557-63.

**Funnell, B. E.** (1991) The P1 plasmid partition complex at parS. The influence of *Escherichia coli* integration host factor and of substrate topology. *J. Biol. Chem* **266**(22):14328-14337

**Galán, J. E. & Sansonetti, P. J.** (1996). Molecular and cellular bases of *Salmonella* and *Shigella* interactions with hosts. In *Escherichia coli* and *Salmonella: Cellular and Molecular Biology*, pp. 2757–2773. Edited by F. C. Neidhardt and others. Washington DC: American Society for Microbiology.

**Gallegos, M. T., Schleif, R., Bairoch, A., Hofmann, K., Ramos, J. L.** (1997) Arac/XylS family of transcriptional regulators. *Microbiol Mol Biol Rev* **61**(4):393-410.

**Garner MM, Revzin A.** (1981) A gel electrophoresis method for quantifying the binding of proteins to specific DNA regions: application to components of the *Escherichia coli* lactose operon regulatory system. *Nucleic Acids Res* **9**(13):3047-60.

**Golubeva, Y. A., Sadik, A. Y., Ellermeier, J. R., Slauch, J. M.** (2012) Integrating Global Regulatory Input Into the *Salmonella* Pathogenicity Island 1 Type III Secretion System. *Genetics* **190**(1):79-90.

**Gómez-Duarte, O. G, Kaper JB.** (1995) A plasmid-encoded regulatory region activates chromosomal *eaeA* expression in enteropathogenic *Escherichia coli*. *Infect Immun* **63**(5):1767-76.

**González-Gil, G., Bringmann, P., Kahmann, R.** (1996) FIS is a regulator of metabolism in *Escherichia coli*. *Mol Microbiol* **22**(1):21-9.

**Goosen, N., van de Putte, P.** (1995) The regulation of transcription initiation by integration host factor. *Mol Microbiol* **16**(1):1-7

**Gowrishankar, J., Manna, D.** (1996) How is osmotic regulation of transcription of the *Escherichia coli proU* operon achieved? A review and a model. *Genetica* **97**:363–378.

**Gowrishankar, J., Pittard, A. J.** (1998) Superimposition of TyrR protein-mediated regulation on osmoreponsive transcription of *Escherichia coli proU* in vivo. *J. Bacteriol* **180**:6743–6748.

**Hagen D. C. and Magasanik, B.** (1973) Isolation of the self-regulated repressor protein of the Hut operons of *Salmonella typhimurium*. *Proc. Natl. Acad. Sci. USA* **70**:808.

**Hale, T. L., Oaks, E. V., Formal, S. B.** (1985) Identification and antigenic characterization of virulence-associated, plasmid-coded proteins of *Shigella* spp. and enteroinvasive *Escherichia coli*. *Infect Immun* **50**(3):620-9.

**Hale, T. L.** (1991) Genetic basis of virulence in *Shigella* species. *Microbiol Rev* **55**(2):206-24.

**Haugen, S. P., Ross, W., Gourse, R. L.** Advances in bacterial promoter recognition and its control by factors that do not bind DNA. *Nat Rev Microbiol.* **2008** Jul;6(7):507-19.  
Epub **2008** Jun 3.

**Hautefort, I., Proenca, M. J., Hinton, J. C.** (2003) Single-copy green fluorescent protein gene fusions allow accurate measurement of *Salmonella* gene expression in vitro and during infection of mammalian cells. *Appl. Environ. Microbiol* **69**:7480–7491.

**Hayes, F., Austin, S. J.** (1993) Specificity determinants of the P1 and P7 plasmid centromere analogs. *Proc Natl Acad Sci U S A* **90**(19):9228-32.

**High, N., Mounier, J., Prévost, M. C., Sansonetti, P. J.** (1992) IpaB of *Shigella flexneri* causes entry into epithelial cells and escape from the phagocytic vacuole. *EMBO J* **11**(5):1991-9.

**Hueck, C. J.** Type III protein secretion systems in bacterial pathogens of animals and plants. *Microbiol Mol Biol Rev* **62**(2):379-433

**Hulton, C. S., Seirafi, A., Hinton, J. C., Sidebotham, J. M., Waddell, L., Pavitt, G. D., Owen-Hughes, T., Spassky, A., Buc, H., Higgins, C. F.** (1990) Histone-like protein H1 (H-NS), DNA supercoiling, and gene expression in bacteria. *Cell* **63**(3):631-42.

**Isberg, R. R.** (1991) Discrimination between intracellular uptake and surface adhesion of bacterial pathogens. *Science* **252**(5008):934-8.

**Ishihama, A.** (2000) Functional modulation of *Escherichia coli* RNA polymerase. *Annu Rev Microbiol* **54**:499-518.

**Jennison, A. V., Verma, N. K.** (2004) *Shigella flexneri* infection: pathogenesis and vaccine development. *FEMS Microbiol Rev* **28**(1):43-58.

- Jensen, R. B., Gerdes, K.** (1997) Partitioning of plasmid R1. The ParM protein exhibits ATPase activity and interacts with the centromere-like ParR-parC complex. *J Mol Biol* **269**(4):505-13.
- Jiang, Y., Yang, F., Zhang, X., Yang, J., Chen, L., Yan, Y., Nie, H., Xiong, Z., Wang, J., Dong, J., Xue, Y., Xu, X., Zhu, Y., Chen, S., Jin, Q.** (2005) The complete sequence and analysis of the large virulence plasmid pSS of *Shigella sonnei*. *Plasmid* **54**(2):149-59.
- Jin, Q., Yuan, Z., Xu, J., Wang, Y., Shen, Y., Lu, W., Wang, J., Liu, H., Yang, J., Yang, F., Zhang, X., Zhang, J., Yang, G., Wu, H., Qu, D., Dong, J., Sun, L., Xue, Y., Zhao, A., Gao, Y., Zhu, J., Kan, B., Ding, K., Chen, S., Cheng, H., Yao, Z., He, B., Chen, R., Ma, D., Qiang, B., Wen, Y., Hou, Y., Yu, J.** (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):4432-41.
- Kaer, K., Mätlik, K., Metsis, M., Speek, M.** (2008) Combination of native and denaturing PAGE for the detection of protein binding regions in long fragments of genomic DNA. *BMC Genomics* **9**:272.
- Kane, C. D., Schuch, R., Day, W. A. Jr., Maurelli, A. T.** (2002) MxiE regulates intracellular expression of factors secreted by the *Shigella flexneri* 2a type III secretion system. *J Bacteriol* **184**(16):4409-19.

**Kane KA, Dorman CJ.** (2011) Rational design of an artificial genetic switch: Co-option of the H-NS-repressed *proU* operon by the VirB virulence master regulator. *J Bacteriol* **193**(21):5950-60.

**Kato, A., Groisman, E. A.** (2008) The PhoQ/PhoP regulatory network of *Salmonella enterica*. *Adv Exp Med Biol* **631**:7-21.

**Keseler, I. M., Collado-Vides, J., Gama-Castro, S., Ingraham, J., Paley, S.** (2005) Ecocyc: a comprehensive database resource for *Escherichia coli*. *Nucl Acids Res* **33**: D334–337

**Kim, D. W., Lenzen, G., Page, A. L., Legrain, P., Sansonetti, P. J., Parsot, C.** (2005) The *Shigella flexneri* effector OspG interferes with innate immune responses by targeting ubiquitin-conjugating enzymes. *Proc Natl Acad Sci U S A* **102**(39):14046-51.

**Köhler, H., Rodrigues, S.P., McCormick, B. A.** (2002) *Shigella flexneri* Interactions with the Basolateral Membrane Domain of Polarized Model Intestinal Epithelium: Role of Lipopolysaccharide in Cell Invasion and in Activation of the Mitogen-Activated Protein Kinase ERK. *Infect Immun* **70**(3):1150-8.



**Korzheva N, Mustaev A, Kozlov M, Malhotra A, Nikiforov V, Goldfarb A, Darst SA.** A structural model of transcription elongation. *Science*. 2000 Jul 28;289(5479):619-25.

**Kotloff, K. L., Winickoff, J. P., Ivanoff, B., Clemens, J. D., Swerdlow, D. L., Sansonetti, P. J., Adak, G. K., Levine, M. M.** (1999) Global burden of *Shigella* infections: implications for vaccine development and implementation of control strategies. *Bull World Health Organ* 77(8):651-66

**Kourilsky, P. and Gros, F.** (1974) in Regulation of Gene Expression in Cultured Cells. Foga Conference Center Proceedings, No. 22, US Government Printing Office, Washington, D.C.

**Kweon, M. N.** (2008). Shigellosis: the current status of vaccine development. *Current Opinion in Infectious Diseases*, 21(3), 313-318.

**Lan, R., Alles, M. C., Donohoe, K., Martinez, M. B., Reeves, P. R.** (2004) Molecular evolutionary relationships of enteroinvasive *Escherichia coli* and *Shigella* spp. *Infect Immun* 72(9):5080-8.

**Lan, R., Lumb, B., Ryan, D., Reeves, P. R.** (2001) Molecular evolution of large virulence plasmid in *Shigella* clones and enteroinvasive *Escherichia coli*. *Infect Immun* 69(10):6303-9.

**Lan, R., Reeves, P. R.** (2002) *Escherichia coli* in disguise: molecular origins of *Shigella*. *Microbes Infect* 4(11):1125-32.

**Lang, B., Blot, N., Bouffartigues, E., Buckle, M., Geertz, M., Gualerzi, C. O., Mavathur, R., Muskhelishvili, G., Pon, C. L., Rimsky, S., Stella, S., Babu, M. M., Travers, A.** (2007) High-affinity DNA binding sites for H-NS provide a molecular basis for selective silencing within proteobacterial genomes. *Nucleic Acids Res* 35(18):6330-7.

**Le Gall, T., Mavris, M., Martino, M. C., Bernardini, M. L., Denamur, E., Parsot, C.** (2005) Analysis of virulence plasmid gene expression defines three classes of effectors in the type III secretion system of *Shigella flexneri*. *Microbiology* 151(Pt 3):951-62.

**Levine, M. M., Kotloff, K. L., Barry, E. M., Pasetti, M. F., Sztein, M. B.** (2007) Clinical trials of *Shigella* vaccines: two steps forward and one step back on a long, hard road. *Nat Rev Microbiol* 5(7):540-53.

**Lucchini, S., Rowley, G., Goldberg, M. D., Hurd, D., Harrison, M., Hinton, J. C.** (2006) H-NS mediates the silencing of laterally acquired genes in bacteria. *PLoS Pathog* 2(8):e81.

**Man, A. L., Prieto-Garcia, M. E., Nicoletti, C.** (2004) Improving M cell mediated transport across mucosal barriers: do certain bacteria hold the keys? *Immunology* **113**(1):15-22.

**Manna, D., Gowrishankar, J.** (1994) Evidence for involvement of proteins HU and RpoS in transcription of the osmoreponsive *proU* operon in *Escherichia coli*. *J Bacteriol* **176**(17):5378-84.

**Martin RG, Rosner JL.** The AraC transcriptional activators. *Curr Opin Microbiol.* **2001** Apr;4(2):132-7

**Matsumoto, H., Young, G. M.** (2009) Translocated effectors of Yersinia. *Curr Opin Microbiol* **12**(1):94-100.

**Maurelli, A. T., Sansonetti, P. J.** (1988) Identification of a chromosomal gene controlling temperature-regulated expression of *Shigella* virulence. *Proc Natl Acad Sci U S A* **85**(8):2820-4.

**Mavris, M., Sansonetti, P. J., Parsot, C.** (2002) Identification of the cis-acting site involved in activation of promoters regulated by activity of the type III secretion apparatus in *Shigella flexneri*. *J Bacteriol* **184**(24):6751-9.

**McAdams, H. H., Srinivasan, B., Arkin, A. P.** (2004) The evolution of genetic regulatory systems in bacteria. *Nat Rev Genet* **5**(3):169-78

**McKenna, S., Beloin, C., Dorman, C. J.** (2003) In vitro DNA-binding properties of VirB, the *Shigella flexneri* virulence regulatory protein. *FEBS Lett* **545**(2-3):183-7.

**McLeod, S. M., Xu, J., Cramton, S. E., Gaal, T., Gourse, R. L., Johnson, R. C.** (1999) Localization of amino acids required for Fis to function as a class II transcriptional activator at the RpoS-dependent *proP* P2 promoter. *J Mol Biol* **294**(2):333-46.

**McLeod, S. M., Xu, J., Johnson, R. C.** (2000) Coactivation of the RpoS-dependent *proP* P2 promoter by Fis and cyclic AMP receptor protein. *J Bacteriol* **182**(15):4180-7.

**Ménard, R., Dehio, C., Sansonetti, P. J.** (1996) Bacterial entry into epithelial cells: the paradigm of *Shigella*. *Trends Microbiol* **4**(6):220-6.

**Ménard, R., Sansonetti, P., Parsot, C., Vasselon, T.** (1994) Extracellular association and cytoplasmic partitioning of the IpaB and IpaC invasins of *S. flexneri*. *Cell* **79**(3):515-25.

**Ménard, R., Sansonetti, P. J., Parsot, C.** (1993) Nonpolar mutagenesis of the ipa genes defines IpaB, IpaC, and IpaD as effectors of *Shigella flexneri* entry into epithelial cells. *J Bacteriol* **175**(18):5899-906.

**Mengaud, J., Ohayon, H., Gounon, P., Mege, R-M., Cossart, P.** (1996) E-cadherin is the receptor for internalin, a surface protein required for entry of *L. monocytogenes* into epithelial cells. *Cell* **84**(6):923-32.

**Miller, J. H.** (1992) A short course in bacterial genetics. Cold Spring Harbor Laboratory Press, Plainview, N.Y.

**Mitrophanov, A.Y., Hadley, T. J., Groisman, E. A.** (2010) Positive autoregulation shapes response timing and intensity in two-component signal transduction systems. *J Mol Biol* **401**(4):671-80.

**Mitobe, J., Arakawa, E., Watanabe, H.** (2005) A sensor of the two-component system CpxA affects expression of the type III secretion system through posttranscriptional processing of InvE. *J Bacteriol* **187**(1):107-13.

**Miyazaki, W. Y, Orr-Weaver, T. L.** (1994) Sister-chromatid cohesion in mitosis and meiosis. *Annu Rev Genet* **28**:167-87.

**Moss, J. E., Fisher, P. E., Vick, B., Groisman, E. A., Zychlinsky, A.** (2000) The regulatory protein PhoP controls susceptibility to the host inflammatory response in *Shigella flexneri*. *Cell Microbiol* **2**(6):443-52.

**Müller, J., Oehler, S., Müller-Hill, B.** (1996) Repression of lac promoter as a function of distance, phase and quality of an auxiliary lac operator. *J Mol Biol* **257**(1):21-9.

**Munson, G. P., Scott, J. R.** (1999) Binding site recognition by Rns, a virulence regulator in the AraC family. *J Bacteriol* **181**(7):2110-7.

**Murakami, K. S., Masuda, S., Campbell, E. A., Muzzin, O., Darst, S. A.** (2002) Structural basis of transcription initiation: an RNA polymerase holoenzyme-DNA complex. *Science* **296**(5571):1285-90.

**Nagarajavel, V., Madhusudan S., Dole S., Rahmouni A. R., Schnetz K.** (2007). Repression by binding of H-NS within the transcription unit. *J. Biol. Chem* **282**:23622–23630.

**Nakayama, S., Watanabe, H.** (1998) Identification of *cpxR* as a positive regulator essential for expression of the *Shigella sonnei virF* gene. *J Bacteriol* **180**(14):3522-8.

**Nakayama, S., Watanabe, H.** (1995) Involvement of *cpxA*, a sensor of a two-component regulatory system, in the pH-dependent regulation of expression of *Shigella sonnei virF* gene. *J Bacteriol* **177**(17):5062-9.

**Navarre, W. W., McClelland, M., Libby, S. J, Fang, F. C.** (2007) Silencing of xenogeneic DNA by H-NS-facilitation of lateral gene transfer in bacteria by a defense system that recognizes foreign DNA. *Genes Dev* **21**(12):1456-71.

**Newell, D. G., Koopmans, M., Verhoef, L., Duizer, E., Aidara-Kane, A., Sprong, H., Opsteegh, M., Langelaar, M., Threlfall, J., Scheutz, F., van der Giessen, J., Kruse, H.** (2011) Food-borne diseases - the challenges of 20 years ago still persist while new ones continue to emerge. *Int J Food Microbiol* **30**;139.

**Nickels BE, Dove SL, Murakami KS, Darst SA, Hochschild A.** Protein-protein and protein-DNA interactions of sigma70 region 4 involved in transcription activation by lambda<sub>daI</sub>. *J Mol Biol.* **2002** Nov 15;324(1):17-34

**Nordström, K., Austin, S. J.** (1989) Mechanisms that contribute to the stable segregation of plasmids. *Annu Rev Genet* **23**:37-69.

**Ochman, H., Lawerence, J. G., Groisman. E. A.** (2000) Lateral gene transfer and the nature of bacterial innovation. *Nature* **405**(6784):299-304

**Ogawa, H., Nakamura, A., Nakaya, R.** (1968) Cinemicrographic study of tissue cell cultures infected with *Shigella flexneri*. *Jpn J Med Sci Biol* **21**(4):259-73.

**Okuda, J., Toyotome, T., Kataoka, N., Ohno, M., Abe, H., Shimura, Y., Seyedarabi, A., Pickersgill, R., Sasakawa, C.** (2005) *Shigella* effector IpaH9.8 binds to a splicing factor U2AF(35) to modulate host immune responses. *Biochem Biophys Res Commun* **333**(2):531-9.

**Opel, M. L., Aeling, K. A., Holmes, W. M., Johnson, R. C., Benham, C. J., Hatfield, G. W.** (2004) Activation of transcription initiation from a stable RNA promoter by a Fis protein-mediated DNA structural transmission mechanism. *Mol Microbiol* **53**(2):665-74.

**Osiecki, J. C., Barker, J., Picking, W. L., Serfis, A. B., Berring, E., Shah, S., Harrington, A., Picking, W. D.** (2001) IpaC from *Shigella* and SipC from *Salmonella* possess similar biochemical properties but are functionally distinct. *Mol Microbiol* **42**(2):469-81.

**Peng, J., Yang, J., Jin, Q.** (2010) Research progress in *Shigella* in the postgenomic era. *Sci China Life Sci.* **53**(11):1284-90.



**Perdomo, J. J., Gounon, P., Sansonetti, P. J.** (1994) Polymorphonuclear leukocyte transmigration promotes invasion of colonic epithelial monolayer by *Shigella flexneri*. *J Clin Invest* **93**(2):633-43.

**Perdomo, O. J., Cavaillon, J. M., Huerre, M., Ohayon, H., Gounon, P., Sansonetti, P. J.** (1994) Acute inflammation causes epithelial invasion and mucosal destruction in experimental shigellosis. *J Exp Med* **180**(4):1307-19.

**Pérez-Rueda, E., Collado-Vides, J.** (2000) The repertoire of DNA-binding transcriptional regulators in *Escherichia coli* K-12. *Nucleic Acids Res* **28**(8):1838-47

**Philpott, D. J., Yamaoka, S., Israël, A., Sansonetti, P. J.** (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):903-14.

**Pon, C. L., Stella, S., Gualerzi, C. L.** (2000) Repression of Transcription by Curved DNA and Nucleoid Protein H-NS: A Mode of Bacterial Gene Regulation. Madame Curie Bioscience Database.

**Porter, M. E., Dorman, C. J.** (2002) In vivo DNA-binding and oligomerization properties of the *Shigella flexneri* AraC-like transcriptional regulator VirF as identified by random and site-specific mutagenesis. *J Bacteriol* **184**(2):531-9.

- Porter, M. E., Dorman, C. J.** (1997) Differential regulation of the plasmid-encoded genes in the *Shigella flexneri* virulence regulon. *Mol Gen Genet* **256**(2):93-103.
- Porter, M. E., Dorman, C. J.** (1997) Positive regulation of *Shigella flexneri* virulence genes by integration host factor. *J Bacteriol* **179**(21):6537-50.
- Porter, M. E., Smith, S. G., Dorman, C. J.** (1998) Two highly related regulatory proteins, *Shigella flexneri* VirF and enterotoxigenic *Escherichia coli* Rns, have common and distinct regulatory properties. *FEMS Microbiol Lett* **162**(2):303-9.
- Prosseda, G., Fradiani, P. A., Di Lorenzo, M., Falconi, M., Micheli, G., Casalino, M., Nicoletti, M., Colonna, B.** A role for H-NS in the regulation of the virF gene of *Shigella* and enteroinvasive *Escherichia coli*. *Res Microbiol* **149**(1):15-25.
- Prosseda, G., Falconi, M., Giangrossi, M., Gualerzi, C. O., Micheli, G., Colonna, B.** (2004) The *virF* promoter in *Shigella*: more than just a curved DNA stretch. *Mol Microbiol* **51**(2):523-37.
- Pupo, G. M., Karaolis, D. K., Lan, R., Reeves, P. R.** (1997) Evolutionary relationships among pathogenic and nonpathogenic *Escherichia coli* strains inferred from multilocus enzyme electrophoresis and mdh sequence studies. *Infect Immun* **65**(7):2685-92.

**Radnedge, L., Davis, M. A., Austin, S. J.** (1996) P1 and P7 plasmid partition: ParB protein bound to its partition site makes a separate discriminator contact with the DNA that determines species specificity. *EMBO J* **15**(5):1155-62.

**Rajkumari, K., Kusano, S., Ishihama, A., Mizuno, T., Gowrishankar, J.** (1996) Effects of H-NS and potassium glutamate on sigmaS- and sigma70-directed transcription in vitro from osmotically regulated P1 and P2 promoters of *proU* in *Escherichia coli*. *J Bacteriol* **178**(14):4176-81.

**Raqib, R., Gustafsson, A., Andersson, J., Bakhiet, M.** (1997) A systemic downregulation of gamma interferon production is associated with acute shigellosis. *Infect Immun* **65**(12):5338-41.

**Rimsky, S., Zuber, F., Buckle, M., Buc, H.** (2001) A molecular mechanism for the repression of transcription by the H-NS protein. *Mol Microbiol* **42**(5):1311-23.

**Ross, W., Ernst, A., Gourse, R. L.** (2001) Fine structure of *E. coli* RNA polymerase-promoter interactions: alpha subunit binding to the UP element minor groove. *Genes Dev* **15**(5):491-506.

**Sansonetti, P. I.** (1999) *Shigella* plays dangerous games. *ASM News* **65**:611 ± 617.

**Sansonetti, P. J., Arondel, J., Fontaine, A., d'Hauteville, H., Bernardini, M. L.** (1991) OmpB (osmo-regulation) and *icsA* (cell-to-cell spread) mutants of *Shigella flexneri*: vaccine candidates and probes to study the pathogenesis of shigellosis. *Vaccine* **9**(6):416-22.

**Sansonetti, P.J., d'Hauteville H, Ecobichon C, Pourcel C.** (1983) Molecular comparison of virulence plasmids in *Shigella* and enteroinvasive *Escherichia coli*. *Ann Microbiol* **134A**(3):295-318.

**Sansonetti, P. J., Kopecko, D. J., Formal, S. B.** (1982) Involvement of a plasmid in the invasive ability of *Shigella flexneri*. *Infect Immun* **35**(3):852-60.

**Sansonetti, P. J., 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):581-90.

**Sansonetti, P. J.** (1994) *Shigella flexneri*: from in vitro invasion of epithelial cells to infection of the intestinal barrier. *Biochem Soc Trans* **22**(2):295-8.

**Savery, N. J., Lloyd, G. S., Kainz, M., Gaal, T., Ross, W., Ebright, R. H., Gourse, R. L., Busby, S. J.** (1998) Transcription activation at Class II CRP-dependent promoters: identification of determinants in the C-terminal domain of the RNA polymerase alpha subunit. *EMBO J* **17**(12):3439-47.

**Schleif, R.** (1996) Two positively regulated systems, *ara* and *mal*., p. 1300-1310. In F. C. Niedhardt (ed.), *Escherichia coli* and *Salmonella typhimurium*. Cellular and molecular biology, vol. 1. American Society for Microbiology, Washington D.C.

**Schlumberger, M. C., Hardt, W. D.** (2006) *Salmonella* type III secretion effectors: pulling the host cell's strings. *Curr Opin Microbiol* 9(1):46-54.

**Sergueev, K., Dabrazhynetskaya, A., Austin, S.** (2005) Plasmid partition system of the P1par family from the pWR100 virulence plasmid of *Shigella flexneri*. *J Bacteriol* 187(10):3369-73.

**Sheridan, S. D., Benham, C. J., Hatfield, G. W.** (1998) Activation of gene expression by a novel DNA structural transmission mechanism that requires supercoiling-induced DNA duplex destabilization in an upstream activating sequence. *J Biol Chem* 273(33):21298-308.

**Shin, D., Lee, J., Huang, H. & Groisman, E. A.** (2006). A positive feedback loop promotes transcription surge that jump-starts *Salmonella* virulence circuit. *Science* 314:1607–1609.

**Singer, A. U., Rohde, J. R., Lam, R., Skarina, T., Kagan, O., Dileo, R., Chirgadze, N. Y., Cuff, M. E., Joachimiak, A., Tyers, M., Sansonetti, P. J., Parsot, C.,**

**Savchenko, A.** (2008) Structure of the *Shigella* T3SS effector IpaH defines a new class of E3 ubiquitin ligases. *Nat Struct Mol Biol* **15**(12):1293-301.

**Soisson, S. M., MacDougall-Shackleton, B., Schleif, R., Wolberger, C.** (1997) Structural basis for ligand-regulated oligomerization of AraC. *Science* **276**(5311):421-5.

**Soisson, S. M., MacDougall-Shackleton, B., Schleif, R., Wolberger, C.** (1997) The 1.6 Å crystal structure of the AraC sugar-binding and dimerization domain complexed with D-fucose. *J Mol Biol* **273**(1):226-37.

**Sompayrac, L., Maalö e, O.** (1973) Autorepressor model for control of DNA replication. *Nature New Biol* **241**, 133.

**Stoebel, D. M., Free, A., Dorman, C. J.** (2008) Anti-silencing: overcoming H-NS-mediated repression of transcription in Gram-negative enteric bacteria. *Microbiology* **154**(Pt 9):2533-45.

**Suzuki, T., Sasakawa, C.** (2001) Molecular basis of the intracellular spreading of *Shigella*. *Infect Immun* **69**(10):5959-66.

**Tabuchi, A., Min, Y. N., Kim, C. K., Fan, Y. L., Womble, D. D., Rownd, R. H.** (1988) Genetic organization and nucleotide sequence of the stability locus of IncFII plasmid NR1. *J Mol Biol* **202**(3):511-25.

**Taniya, T., Mitobe, J., Nakayama, S., Mingshan, Q., Okuda, K., Watanabe, H.**

(2003) Determination of the InvE binding site required for expression of IpaB of the *Shigella sonnei* virulence plasmid: involvement of a ParB boxA-like sequence. *J Bacteriol* **185**(17):5158-65.

**Tobe, T., Yoshikawa, M., Sasakawa, C.** (1995) Thermoregulation of virB

transcription in *Shigella flexneri* by sensing of changes in local DNA superhelicity. *J Bacteriol* **177**(4):1094-7.

**Tobe, T., Yoshikawa, M., Mizuno, T., Sasakawa, C.** (1993) Transcriptional control

of the invasion regulatory gene virB of *Shigella flexneri*: activation by *virF* and repression by H-NS. *J Bacteriol* **175**(19):6142-9.

**Tobe, T., Nagai, S., Okada, N., Adler, B., Yoshikawa, M., Sasakawa, C.** (1991)

Temperature-regulated expression of invasion genes in *Shigella flexneri* is controlled through the transcriptional activation of the *virB* gene on the large plasmid. *Mol Microbiol* **5**(4):887-93.

**Tominaga, A., Lan, R., Reeves, P. R.** (2005) Evolutionary changes of the flhDC

flagellar master operon in *Shigella* strains. *J Bacteriol* **187**(12):4295-302.

**Tran Van Nhieu, G., Ben-Ze'ev, A., Sansonetti, P. J.** (1997) Modulation of bacterial entry into epithelial cells by association between vinculin and the *Shigella* IpaA invasin. *EMBO J* 16(10):2717-29.

**Tran Van Nhieu, G., Caron, E., Hall, A., Sansonetti, P. J.** (1999) IpaC induces actin polymerization and filopodia formation during *Shigella* entry into epithelial cells. *EMBO J* 18(12):3249-62.

**Travers, A., Muskhelishvili, G.** (2005) DNA supercoiling - a global transcriptional regulator for enterobacterial growth? *Nat Rev Microbiol* 3(2):157-69.

**Turner, E. C., Dorman, C. J.** (2007) H-NS antagonism in *Shigella flexneri* by VirB, a virulence gene transcription regulator that is closely related to plasmid partition factors. *J Bacteriol* 189(9):3403-13.

**Urbonavicius, J., Durand, J. M., Björk, G. R.** (2002) Three modifications in the D and T arms of tRNA influence translation in *Escherichia coli* and expression of virulence genes in *Shigella flexneri*. *J Bacteriol* 184(19):5348-57.

**Vasselon, T., Mounier, J., Prevost, M. C., Hellio, R., Sansonetti, P. J.** (1991) Stress fiber-based movement of *Shigella flexneri* within cells. *Infect Immun* 59(5):1723-32.



**Wake, R. G., Errington, J.** (1995) Chromosome partitioning in bacteria. *Annu Rev Genet* **29**:41-67.

**Wassef, J. S., Keren, D. F., Mailloux, J. L.** (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):858-63.

**Watanabe, E., Inamoto, S., Lee, M. H., Kim, S. U., Oguo, T., Mori, H., Hiraga, S., Yamasaki, M., Nagai, K.** (1989) Purification and characterization of the *sopB* gene product which is responsible for stable maintenance of mini-F plasmid. *Mol Gen Genet* **218**(3):431-6

**Watanabe, H., Arakawa, E., Ito, K., Kato, J., Nakamura, A.** (1990) Genetic analysis of an invasion region by use of a Tn3-lac transposon and identification of a second positive regulator gene, *invE*, for cell invasion of *Shigella sonnei*: significant homology of *invE* with ParB of plasmid P1. *J Bacteriol* **172**(2):619-29.

**Way, S. S., Borczuk, A. C., Dominitz, R., Goldberg, M. B.** (1998) An essential role for gamma interferon in innate resistance to *Shigella flexneri* infection. *Infect Immun* **66**(4):1342-8.

**Wei, J., Goldberg, M. B., Burland, V., Venkatesan, M. M., Deng, W., Fournier, G., Mayhew, G. F., Plunkett, G., Rose, D. J., Darling, A., Mau, B., Perna, N. T.,**

**Payne, S. M., Runyen-Janecky, L. J., Zhou, S., Schwartz, D. C., Blattner, F. R.**

(2003) Complete genome sequence and comparative genomics of *Shigella flexneri* serotype 2a strain 2457T. *Infect Immun* **71**(5):2775-86.

**Weinrauch, Y., Drujan, D., Shapiro, S. D., Weiss, J., Zychlinsky, A.** (2002)

Neutrophil elastase targets virulence factors of enterobacteria. *Nature* **417**(6884):91-4.

**Weinreich, M. D., Reznikoff, W. S.** (1992) Fis plays a role in Tn5 and IS50

transposition. *J Bacteriol* **174**(14):4530-7.

**Weiss, J., Victor, M., Kao, L., Elsbach, P.** (1985) Killing of Gram-negative bacteria by

neutrophils: role of O<sub>2</sub>-independent system in intracellular killing and evidence of O<sub>2</sub>-dependent extracellular killing. *Adv Exp Med Biol* **184**:35-50.

**Whitehead, N. A., Barnard, A. M., Slater, H., Simpson, N. J., Salmond, G. P.**

(2001) Quorum-sensing in Gram-negative bacteria. *FEMS Microbiol Rev* **25**(4):365-404.

**Williams, C. L. & Cotter, P. A.** (2007). Autoregulation is essential for precise

temporal and steady-state regulation by the *Bordetella* BvgAS phosphorelay. *J.*

*Bacteriol* **189**:1974–1982.

- Wing, H. J., Goldman, S. R., Ally, S., Goldberg, M. B.** (2005) Modulation of an outer membrane protease contributes to the virulence defect of *Shigella flexneri* strains carrying a mutation in the *virK* locus. *Infect Immun* **73**(2):1217-20.
- Winstanley, C., Hart, C. A.** (2001) Type III secretion systems and pathogenicity islands. *J Med Microbiol* **50**(2):116-26.
- van Hijum, S. A., Medema, M. H., Kuipers, O. P.** (2009) Mechanisms and evolution of control logic in prokaryotic transcriptional regulation. *Microbiol Mol Biol Rev* **73**(3):481-509,
- Vecchiarelli, A. G., Schumacher, M. A., Funnell, B. E.** (2007) P1 partition complex assembly involves several modes of protein-DNA recognition. *J Biol Chem* **282**(15):10944-52.
- Xu, J., Johnson, R. C.** (1995) Fis activates the RpoS-dependent stationary-phase expression of *proP* in *Escherichia coli*. *J Bacteriol* **177**(18):5222-31.
- 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):71-9.

**Zhi, H., Wang, X., Cabrera, J. E., Johnson, R. C., Jin, D. J.** (2003) Fis stabilizes the interaction between RNA polymerase and the ribosomal promoter *rrnB* P1, leading to transcriptional activation. *J Biol Chem* **278**(47):47340-9.

**Zychlinsky, A., Fitting, C., Cavaillon, J. M., Sansonetti, P. J.** (1994) Interleukin 1 is released by murine macrophages during apoptosis induced by *Shigella flexneri*. *J Clin Invest* **94**(3):1328-32.

**Zychlinsky, A., Kenny, B., Prevost, M. C., Holland, I. B., Sansonetti, P. J.** (1993) The ipaB gene of *Shigella flexneri* and macrophage-programmed cell death. *Infect Agents Dis* **2**(4):212-4.

**Zychlinsky, A., Prevost, M. C., Sansonetti, P. J.** (1992) *Shigella flexneri* induces apoptosis in infected macrophages. *Nature* **358**(6382):167-9

**Zychlinsky, A., Thirumalai, K., Arondel, J., Cantey, J. R., Aliprantis, A. O., Sansonetti, P. J.** (1996) In vivo apoptosis in *Shigella flexneri* infections. *Infect Immun* **64**(12):5357-65.